

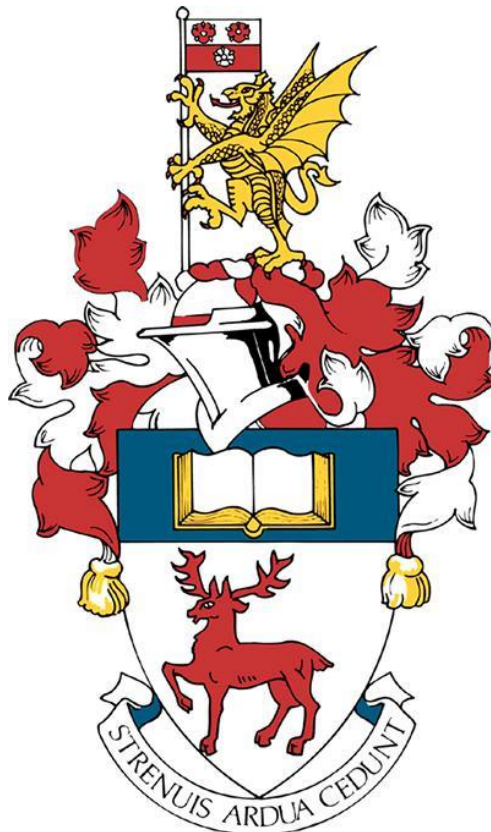
University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: B S Mackay (2021) “Labelling, Modelling, and Predicting Cell Biocompatibility using Deep Neural Networks”, University of Southampton, Optoelectronic Research Centre of the Zepler Institute, PhD Thesis, pagination.

Data: B S Mackay (2022) Data for the Labelling, Modelling, and Predicting Cell Biocompatibility using Deep Neural Networks. <https://doi.org/10.5258/SOTON/D2076>




University of Southampton

Faculty of Engineering and Physical Sciences

Optoelectronic Research Centre of the Zepler Institute

LABELLING, MODELLING, AND PREDICTING CELL BIOCOMPATIBILITY USING DEEP NEURAL NETWORKS

Benita S. Mackay

 0000-0003-2050-8912

Supervisory Team:

Ben Mills, Richard O. C. Oreffo, James A. Grant-Jacob,
Janos Kanczler & Robert W. Eason

Thesis for the degree of Ph.D.

DOI: 10.5258/SOTON/D2076

DECEMBER 25, 2021

EDITED JULY 15, 2022

UNIVERSITY OF SOUTHAMPTON
United Kingdom

ABSTRACT

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Optoelectronic Research Centre of the Zepler Institute

Thesis submitted for the degree of Doctor of Philosophy

Labelling, Modelling, and Predicting Cell Biocompatibility using Deep Neural Networks

by

Benita S. Mackay

With an approximate annual cost of £1.7 billion for the NHS, osteoporosis and osteoarthritis alone represent a major socio-economic burden to the UK. While bisphosphonates and non-opioids such as non-steroidal anti-inflammatory drugs are prescribed for osteoporosis, these types of drugs may lead to serious medical complications when high doses are taken for a long time or when someone is at an advanced age or in poor health, and may even accelerate cartilage destruction in osteoarthritis. Tissue engineering is a successful alternative or additional approach, but the use of grafts and implants is not without risk, specifically the risk of rejection and failure. Therefore, there is a growing need for innovative techniques to promote implant integration and reduce the failure rate of osteopathic intervention. Additionally, at the other end of the age scale, poor placental function can compromise fetal development and placental function directly determines fetal growth. Poor fetal growth is linked with higher chronic disease rates, so leads to an increased probability of health conditions in later life.

The interaction between fibroblasts, pericytes, and endothelial cells in placenta can improve understanding of how cell placement and behaviour affect placental health. To model and analyse these complex interactions, 3D visualisation is required, and 3D-labelling is therefore a necessity to the medical imaging field, which can take months of researcher-time. Similarly, the experimentation required in tissue engineer, from *in-vitro* to *ex-vitro*, can be equally lengthy and complex, as cell response to biochemical and biophysical cues remains poorly understood.

The three areas of research in this project include labelling cells for 3D visualisation, modelling cell response to biophysical cues, and predicting biocompatibility of tissue engineering scaffolds. The data used for these approaches includes 3D nanoscale-resolution images of placenta for labelling and images of 3D bioengineered scaffolds for biocompatibility analysis, which were provided by collaborators. Data also included images of stem cells cultured on topographically-varied surfaces, to analyse stem cell response to biophysical cues, which was collected for this project. This data was used to train multiple deep neural networks, with the goal of applying deep learning to label cells in placenta, and both model and predict cell responses to biophysical and biochemical cues.

It was found that deep neural networks can be used to replace labour-intensive manual labelling, with automated labels comparable, pixel-to-pixel, to manual labels by over 98% on average. The response of cells to physical cues can also be modelled by a deep neural network. With a probability of $P < 0.001$, it can therefore be used as a model, with potential implications for tissue structure development and tissue engineering. Deep learning can also be used to predict biocompatibility, which may act as a future replacement for animal models in place of traditional computer modelling. When predictions were compared with experimental results, images displayed excellent agreement.

Tissue engineering is an increasingly important area of regenerative medicine, partly due to aging populations around the world, combining cell biology, bioengineering and clinical research. However, focus on health from much earlier in life, such as through analysis on placental tissue, may also aid in overcoming the high chronic-disease levels in old age. Applying deep learning to regenerative medicine research may not only help increase efficiency of 3D-image processing, but also potentially help increase understanding of stem cell behaviour and reduce levels of necessary animal testing in research. In the future, application of deep learning to regenerative medicine could help increase quality of life in our later years.

TABLE OF CONTENTS

Abstract.....	1
Table of Figures	7
Research Thesis Declaration.....	11
Acknowledgements.....	13
Abbreviations and Acronyms	15
Mathematical Notation.....	19
1 Introduction.....	20
1.1 Project Motivation.....	21
1.2 Thesis Outline and Aims	22
1.3 Contributions and Publications	24
1.3.1 Summary of Contributions.....	24
1.3.2 Publications.....	24
First Author Publications Resulting From This Project :.....	24
Additional Output:	25
2 Background.....	30
2.1 Tissue Engineering.....	30
2.1.1 Stem Cells and Biomaterials.....	31
2.1.2 A Problem of Scope	33
2.2 Topography and Cells	34
2.2.1 Cell-Topography Interaction.....	35
2.2.2 Topographical Effects.....	38
2.3 Topography Creation.....	42
2.3.1 Femtosecond Laser-Machining and Avalanche Ionisation.....	43
2.3.2 Advantages of Laser Ablation	45
2.4 Biomaterials and Animal Testing.....	48
2.4.1 Biochemical Characteristics.....	48

2.4.2	CAM Assays	51
2.5	Imaging A Placenta in 3D	52
2.5.1	Placenta	52
2.5.2	Fibroblasts, Pericytes and Endothelial Cells.....	54
2.5.3	3D Imaging	54
2.6	Summary	56
3	Deep Learning.....	58
3.1	Introductory Mathematics of Deep Neural Networks	58
3.2	Concept of Deep Learning	61
3.2.1	Simplified Conceptual Mathematics of Deep Learning.....	67
3.3	U-Nets and W-Nets	74
3.4	Training a Neural Network: Example	76
4	Data: Collection and Composition.....	81
4.1	Placental SBFSEM Images – Data for Chapter 5	81
4.2	Stem Cell Response to Surface Topographies – Data for Chapter 6	81
4.2.1	Surface Topographies on Glass	82
4.2.2	Cell Culture and Imaging	86
4.2.3	Collected data composition.....	89
4.3	Stem cell response to Scaffolds – Data for Chapter 7.....	91
4.3.1	Data for <i>in vitro</i> Biocompatibility.....	91
4.3.2	Predicting <i>in vivo</i> Biocompatibility	93
5	Labelling 3-Dimensional Placental Tissue	97
5.1	Introduction	97
5.2	Hypothesis, Aims and Objectives	99
5.2.1	Hypothesis.....	99
5.2.2	Aims and Objectives	99
5.3	Method	100

5.3.1	Single Cell Labelling	100
5.3.2	Multi-Cellular Structures	103
5.3.3	Testing The Network	103
5.4	Results and Discussion.....	104
5.4.1	Analysis of Fibroblast Labelling.....	104
5.4.2	Reducing The Network Requirements for Pericyte Labelling.....	110
5.4.3	Multi-Cellular Labelling with Endothelial Cells	113
5.4.4	Labelling the Unknown.....	119
5.5	Summary	120
6	Modelling The Cell Response to Surface Topographies	123
6.1	Introduction	123
6.2	Hypothesis, Aims and Objectives	126
6.2.1	Hypothesis.....	126
6.2.2	Aims and Objectives	126
6.3	Method	127
6.3.1	Training the Neural Network	127
6.4	Network Outputs and Analysis	130
6.4.1	Introduction to Model Testing	130
6.4.2	Time and Density Independence.....	131
6.4.3	Topography and Density Dependence	133
6.4.4	Topographies for Alignment.....	136
6.4.5	Evaluating Versatility and Limitations	137
6.4.6	Statistical Validation.....	139
6.4.7	Network-Model Prediction	143
6.5	Summary	145
7	Predicting Biocompatibility of 3-Dimensional Scaffolds.....	147
7.1	Introduction	147

7.2	Hypothesis, Aims and Objectives	149
7.2.1	Hypothesis.....	150
7.2.2	Aims and Objectives	150
7.3	Method	151
7.3.1	Fluorescent Image Analysis.....	152
7.3.2	Blood Vessel Isolation and Analysis	153
7.4	Results and Discussion.....	156
7.4.1	Fluorescence Imaging Generation Analysis	158
7.4.2	CAM Assay Image Generation Analysis.....	161
7.4.3	Link-up of Multiple Networks	165
7.5	Summary	170
8	Conclusion	173
8.1	Chapter 5	173
8.2	Chapter 6	173
8.3	Chapter 7	174
8.4	A Note on Future Research	174
8.5	Research Caveats.....	174
8.6	Final Conclusion	175
	Appendix.....	177
	Stro-1 isolation method and protocol.....	177
	References.....	183
	Contacting The Author	192

TABLE OF FIGURES

Figure 1: The population of the world is getting older.20

Figure 2: Project outline and overarching premise of application of AI for future modelling applications.23

Figure 3: An illustration of the surface of a textured implant.....32

Figure 4: Elongation and alignment of human corneal epithelial cells parallel to nano-grating topography.36

Figure 5: Cell spreading and interaction with nanoscale topography.....39

Figure 6: Skeletal stem cell response to different nanotopographies.41

Figure 7: A schematic difference between processing of long pulse (μ s) laser machining with ultrashort (fs) laser machining.45

Figure 8: Scanning Electron Microscope (SEM) images of micromachining over a sphere through use of a Digital Micromirror Device (DMD) and computational correction.47

Figure 9: A table of different biophysical and biochemical cues, and the associated effects on cell behaviour.....49

Figure 10: SEM images of scaffold structures and associated stereo micrographs of calcium staining (Alizarin red) for MSC osteogenesis.....50

Figure 11: Two different methods of performing CAM assays resulting in skeletonised image for vessel analysis.51

Figure 12: An illustration of the placenta connected to a fetus and a photograph of a placenta postnatal outside of the womb.53

Figure 13: Process of Serial Block Face Scanning Electron Microscopy (SBFSEM) imaging.56

Figure 14: BNN (biological neural network) cells and ANN (artificial neural network) nodes have many similarities.62

Figure 15: Comparison of a simple neural network and a deep neural network.64

Figure 16: A U-Net is able to produce more realistic-appearing outputs than a CNN alone. .66

Figure 17: Simple NN architecture with a single hidden layer.....68

Figure 18: Gradient descent visualised with a 3D error landscape illustration.72

Figure 19: Visualisation of a matrix multiplication where a larger 2D array is transformed into a smaller 2D array via a filter.73

Figure 20: CNN utilise both a convolution and pooling layer combination, with image data transformed at each layer.74

Figure 21: An example W-Net consisting of two U-Net based convolutional and deconvolutional arcs. 3-75

Figure 22: A combination of testing images for the 24 hour cell prediction AI. Images A-C are 256×256 images of Stro1+ isolated SSCs and D-F are 512×512 images of the same cells. 77

Figure 23: Graph of L1 error of versions 6-15 of the Cell Predictor cGAN as a percentage of incorrect pixels in training data vs the step count of training models. 79

Figure 24: The path from a designed topography to associated cell response through experimental and deep learning methods..... 82

Figure 25: The laser-machining process for creating microtopographical glass substrates. ... 83

Figure 26: SEM images of laser-machined parallel lines in glass with varying line separation. 85

Figure 27: The fluorescent image capture of cells cultured on a laser-machined glass sample. 87

Figure 28: Skeletal Stem cell enriched population response to laser-machined topographies, with adhesion and alignment along the parallel lines. 88

Figure 29: A graph for comparing the mean L1 error of network output when the images of stem cells from a single patient are extracted from the training dataset and are instead used for testing. 90

Figure 30: Material biocompatibility prediction method for the in vitro data transformation. 91

Figure 31: Generating data input-output pairs for training a DNN to transform fluorescent images to CAM assay images. 94

Figure 32: Labelling SBFSEM-imaged placenta to gain 3D projections of pericytes. 98

Figure 33: The method of single cell automated labelling, which contains several steps. 101

Figure 34: The training of a neural network for the automated labelling of feature structures. 102

Figure 35: Comparison of automated and manually labelled SEM imaged placenta for three different z-stack positions. 105

Figure 36: Graph of multiple statistical analysis factors for comparing automated to manually labelled SBFSEM placenta. 106

Figure 37: The extracted data from Z-position 590, with a visual comparison. 108

Figure 38: Comparison of a projected 3D structure of a fibroblast in SBFSEM imaged-placenta from both automated and manual labelling. 109

Figure 39: Comparison of automated and manually labelled SEM imaged placenta for three different z-stack positions 110

Figure 40: A graph comparing the total corresponding pixels for network output when 1 in 2, 1 in 5 and 1 in 10 images are used for training. 111

Figure 41: A graph comparing the matching pixels from network output when 1 in 2, 1 in 5 and 1 in 10 images are used for training with a traditional computational method when an image 1 position away, 5 positions away, and 10 positions away is used..... 112

Figure 42: Illustration of the W-Net architecture. 114

Figure 43: Comparison of automated and manually labelled SEM imaged placenta for three different z-stack positions. 115

Figure 44: Graph of multiple statistical analysis factors for comparing automated to manually labelled SBFSEM placenta. 116

Figure 45: In-depth graph of corresponding pixels across the z-stack of SBFSEM imaged placenta. 117

Figure 46: The extracted data from Z-positions 40 and 298..... 118

Figure 47: Visual evaluation of unlabelled regions and unlabelled stacks not used in training the network..... 120

Figure 48: Circular nature of improvement in science, data, and AI models. 124

Figure 49: The deep neural network W-Net architecture consists of multiple convolutional layers and skip connections between encoder and decoder sections. 128

Figure 50: Changes to cell behaviour over time for different densities and surface patterning in experimental images. 131

Figure 51: Maintaining a blank topography, time-points and density were varied to evaluate time-density independence..... 132

Figure 52: Two topographical inputs were used to test the network output for varying densities..... 134

Figure 53: Experimental images of skeletal cells grown on two cross-parallel topographies. 135

Figure 54: The generated output (b, d, f, h) for each line separation varies depending on the input topography (a, c, e, g). 136

Figure 55: The limits to successfully generated output for extreme topographical input. 138

Figure 56: Statistical testing required analytical differences between target outputs and network generated outputs. 140

Figure 57: The network was tested for overfitting from a range of inputs that would produce a random, unpredictable, output.	142
Figure 58: Four images of cell response to laser-machined parallel lines at increasing line separation, from 1 μm to 14 μm , and a graph of minimum separation required for cell alignment with data points (blue) taken from network output.....	143
Figure 59: The cycle of improvement to science, data, and neural networks (NNs) can include many types of data.	148
Figure 60: Chapter 5 advances upon work accomplished in chapters 3 and 4, with more DNNs likely required.....	150
Figure 61: Illustration of two networks used in a series for biocompatibility prediction.....	151
Figure 62: Analysis of material biocompatibility prediction using pixel comparisons.	153
Figure 63: Processing of CAM assay images for comparative analysis of network output and target images.	155
Figure 64: Randomly selected network generated output and experimental target images for multiple scaffold types.....	156
Figure 65: Plot of average percentage difference in green and red pixel values for testing scaffolds and randomly generated images.	158
Figure 66: Plots of the mean green pixel value and mean red pixel value for each scaffold type.....	160
Figure 67: Plots of the mean quantity of vessels and mean number of junctions for each scaffold type.....	162
Figure 68: Plot of the mean lacunarity for each scaffold type and randomly generated images.	164
Figure 69: Two scaffold predictions and accompanying target images overlaid with the extracted blood vessels	165
Figure 70: Results of two networks used in a series for biocompatibility prediction.....	167
Figure 71: Using a DNN to go from scaffold information to CAM assay images without fluorescent images.	168
Figure 72: Outputs from a single network trained without in vitro data and a network trained in a series using in vitro data.....	169

RESEARCH THESIS DECLARATION

Print name:

Benita S. Mackay

Title of thesis:

Labelling, Modelling, and Predicting Cell Biocompatibility using Deep Neural Networks

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published.

Signature: Date: 10 Dec 2021

ACKNOWLEDGEMENTS

I would like to thank the following people and organisations, as without their input and support this project could have been accomplished:

- The Faculties of Engineering and Physical Science and Medicine at the University of Southampton, for guidance, education, and resources.
- EPSRC for funding my PhD position.
- My supervisory team of Drs Ben Mills, James Grant-Jacob, Janos Kanczler, and Professors Robert Eason and Richard Oreffo. I have relied on the continued support and expertise of each, and I am forever grateful. Dr Ben Mills and Professor Richard Oreffo, your thoughtful critiques and supervision during the last three years have improved my skillset both inside academia and out, and I will continue to take your lessons forward.
- Matthew Praeger and the researchers of the Bone and Joint Research Group and the AI Photonics Research Group. From technicians to postdocs to fellow students, I owe you special thanks for helping me overcome obstacles and inspiring me throughout.
- My collaborators, Professor Rohan Lewis and Dr Karen Marshall, without whom two thirds of these projects would not have been possible!
- Finally, Ma and Mattias. Even when life around the world was changing beyond recognition and times were more uncertain than ever, you stood by me. Thank you.

ABBREVIATIONS AND ACRONYMS

Term	Definition
2D	2-Dimensional
3D	3-Dimensional
AI	Artificial Intelligence A wide-ranging branch of computer science focussed with building smart machines capable of performing tasks that typically require human-like intelligence.
AM	Acetoxymethyl AM groups are used as an established tool in cell biology to deliver and trap exogenous optical probes inside cells.
ANN	Artificial Neural Network A computational model that mimics the way nerve cells work in the human brain.
AVC	Animal Visual Cortex Located at the occipital pole of the brain the visual cortex receives, integrates and interprets the information relayed from the eye via subcortical nuclei.
BNN	Biological Neural Network A network of neurons that are connected together by axons and dendrites. The connections between neurons are made by synapses.
CAM	Chorioallantoic Membrane A simple, highly vascularized extraembryonic membrane.
cGAN	Conditional Generative Adversarial Network An ANN architecture typically utilising a U-Net and CNN in combination.
CNN	Convolutional Neural Network An ANN architecture utilising convolutional layers. Typically used in image-related tasks.
CPU	Central Processing Unit
D	Discriminator

	An independent ANN used to discriminate between output and target data.
DMD	Digital Micromirror Device A DMD chip has on its surface several hundred thousand microscopic mirrors arranged in a rectangular array which correspond to the pixels in the image to be displayed. The mirrors can be individually rotated to an on or off state.
DNA	Deoxyribonucleic Acid
DNN	Deep Neural Network An ANN comprising hidden layers, which are layers between the input and output layers.
EBL	Electron Beam Lithography A technique for creating nanostructures that are too small to fabricate with conventional photolithography using a focused beam of electrons.
ECM	Extracellular Matrix A three-dimensional network consisting of extracellular macromolecules and minerals, such as collagen, enzymes, glycoproteins and hydroxyapatite that provide structural and biochemical support to surrounding cells.
EthD-1	Ethidium Homodimer-1 A high-affinity nucleic acid stain that is weakly fluorescent until bound to DNA and emits red fluorescence.
FA	Frame Advance
FD	Frame Difference
G	Glasgow coating
GPU	Graphics Processing Unit
hASC	Human Multipotent Adipose Stem Cell
HBMSC	Human Bone Marrow Stromal Cell
L1	Least Absolute Deviations
L2	Least Mean-Squared
Leaky ReLU	Leaky Rectified Linear Unit
micro-CT	Micro-Computed Tomography

	A high-resolution method of imaging in which CT is made on relatively small samples.
M	Mineral coating
MSC	Marrow Stromal Cell
NC3R	National Centre for the Replacement, Refinement and Reduction of Animals in Research
NES	Nintendo™ Entertainment System (© 2018 Nintendo)
NIBIB	National Institute of Biomedical Imaging and Bioengineering
NN	Neural Network See ANN
NSAID	Non-Steroidal Anti-Inflammatory Drug
NSQ	Near Square
N	Nylon scaffold
OCN	Osteocalcin A small (49-amino-acid) noncollagenous protein hormone found in bone and dentin.
OPN	Osteopontin An extracellular structural protein and an organic component of bone.
PCR	Polymerase Chain Reaction
Plain	P
P-value	Probability Value
ReLU	Rectified Linear Unit
RGB	Red, Green and Blue
RNN	Recurrent Neural Network A type of ANN that typically is applied to sequential or time series data.
SBFSEM	Serial Block Face Scanning Electron Microscope/Microscopy A method of generating high resolution 3D images from small samples.
SEM	Scanning Electron Microscope/Microscopy A method of generating high resolution 2D images from small samples.

T	Titanium scaffold
αMEM	Alpha Minimum Essential Medium Used in the <i>in-vitro</i> culturing of cells.
λ	Mean Lacunarity Lacunarity, from the Latin lacuna, meaning "gap" or "lake", is a specialized term in geometry referring to a measure of how patterns, especially fractals, fill space, where patterns having more or larger gaps generally have higher lacunarity.

MATHEMATICAL NOTATION

The following mathematical notational conventions are used throughout this thesis:

Scalar	Italicised or regular upper or lower case letter	A, a, A, a
Vector	Bold lower case letter	a
Matrix	Bold italicised upper case letter	A
Single matrix layer	Bold italicised upper case letter with square bracketed superscript denoting the layer i	$A^{[i]}$
Single matrix point	Bold italicised upper case letter with square bracketed superscript denoting the layer i and row j	$A^{[i][j]}$
Transposed matrix	Bold italicised upper case letter with superscript T	A^T
Hadamard product operator		\circ
Function	The function a in terms of b	$a(b)$
Differential of a function	The differentiation of a function a	da
Partial differential	The partial differentiation of a function a in terms of b	$\frac{\partial a}{\partial b}$

1 INTRODUCTION

Osteoarthritis affects up to 1 in 5 people and causes immense long-term pain and disability, and the prevalence of suffering from osteoarthritis is on the rise due to obesity and older populations, from 12.7% to 17.4% over 2003 to 2013 in Finland and a 66% increase predicted in Australia over the next thirty years [18-20]. Joint replacement can provide an effective treatment for this multifactorial condition, yet outcomes can be poor and prosthesis lifespans are limited. Tissue engineering, specifically using a patient's own cells *ex vivo* to generate and grow replacements for damaged areas within the body, could provide a way to replace damaged bone through the extraction, *ex vivo* optimisation and *in vivo* transplantation of marrow-derived skeletal stem cells [21].

Regenerative medicine is a field with the potential to revolutionise modern healthcare, with particular ramifications for the treatment of increasingly aged populations, Figure 1. Advanced age is known to be associated with increased morbidities and, despite the fact that over the last 50 years we have improved life expectancy considerably, quality of life in latter years of life has not seen matched expectations [22, 23]. Osteoarthritis now affects up to 1 in 5 people, causing chronic pain and physical disability, with evidence showing the risk is rising every year in multiple countries [24, 25].

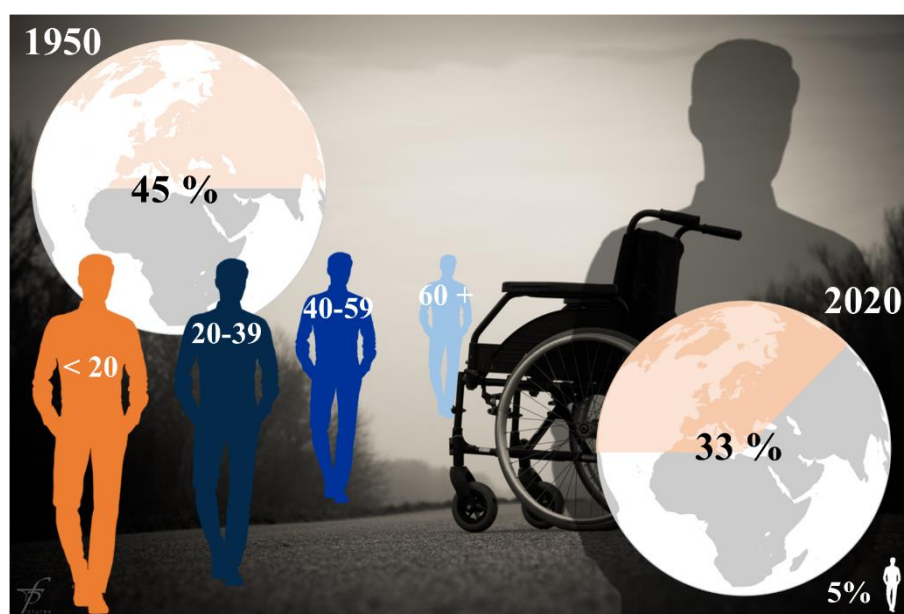


Figure 1: The population of the world is getting older.

Data from the UN Pew Research Centre shows that 45% of the population were less than 20 years old in 1950s compared to only 33% in 2020. The height of the people represents the proportion of the population in that age group in 2020, with 5% shown in white on the bottom right. The proportion of people over sixty is now over 10%, which equates to up to a billion people at risk of age related diseases and medical conditions across the planet. Background image has no Copyright under CC0 1.0 Universal.

Tissue engineering is a field which is both complex and time-intensive. It can range from *in-vitro* to *in-situ* experimentation, utilising both specially-designed bioengineering innovations and the body's own healing responses to improve patient outcomes, such as application of stem cells.

Stem cell control is a key multifaceted aim within tissue engineering, as successful bone repair based on stem cell therapies require conditions for stem cell proliferation without differentiation for applicable scales, and also identification and isolation of the target cells for *in vivo* bone regeneration. Adult skeletal stem cells, often called mesenchymal stem cells, are found in blood vessels in bone marrow, so can be isolated from human bone marrow stromal cells (HBMSCs), and have proven differentiation and regenerative potential and capabilities [26-28]. However, there is still limited understanding in how stem cells respond to environmental cues, and therefore stem cell control is difficult [21, 29-31].

A key challenge to furthering understanding into stem cell behaviour and control is the time and resources required to experiment the complex nature of stem cell response. Therefore, the objective to this thesis is to reduce the time and resources required for tissue engineering research through application of computational methods to the regenerative medicine field, with a focus on modelling stem cell response.

1.1 PROJECT MOTIVATION

Harnessing topographical cues offers a promising technique to control stem cell fate and function. This can develop into an ability to use a patient's own stem cells for targeted generation of bone or tissue within the body, leading to improved joint repair. Cells respond to the shape of their environment due to changes in contact guidance, cell spreading and contact inhibition (cell overlap) [32]. Cell behaviour can therefore be influenced through the precise engineering of surfaces and use of surface-directed nanobiotechnologies [33, 34].

Dalby et al. demonstrated that certain nanopatterns can significantly influence stem cell behaviour [35], while Heath et al. have demonstrated laser machining techniques for fabrication resolution below the wavelength scale and accurate machining of 3D surfaces, which could be applied to implants [36, 37]. For microtopographies, advances in AI have shown the potential for precise control of laser-based fabrication [38] and these techniques can potentially be combined with previous sub-wavelength machining work to create accurate and reproducible nanotopographies.

Introduction

Rather than developing an model solely based on the topographical response of stem cells, the project branched into modelling for larger multi-cellular systems, with potential applications in both the modelling of placental development and the analysis of cell response to biochemical alongside biophysical cues, with impacts on lifelong health and regenerative medicine respectively. Such models could aid in targeted medicine of the future, where not only could conditions such as osteoarthritis be improved through patient-specific treatments and management, but such conditions may also be predicted (and maybe one day prevented) from birth.

Additionally, projects involving cells of various types, not limited to stem cells and their derivatives, provide information on where limits or potential complications with computer-aided research may arise. The use of computational assistance in research is an important addition to regenerative medicine because the cell response to surrounding topographies and materials is too complex, and not yet well enough understood, to easily understand or predict. Similarly, models on placental development are still in their infancy, with implications on life-long health still to be accurately determined [39].

1.2 THESIS OUTLINE AND AIMS

There are two main areas to this thesis. The first is to limit time spend on labour-intensive tasks within the regenerative medicine field. The second is to limit the experimentation requires to determine stem cell responses in tissue engineering, specifically. As the second is an especially complex areas, this will be split into two sections. The thesis closely follows an overarching outline shown in Figure 2.

The three aims are to determine whether:

1. Unlabelled SBFSEM sections can be automatically labelled to identify and label specific cell types without labelling undesired features on unseen SBFSEM sections.
2. A model can be developed with a limited dataset of roughly 200 images to simulate cell behaviour for derivation of novel variables.
3. CAM assay images for a scaffold of a specific material and coating combination can be generated that is indistinguishable form an experimentally obtained CAM assay image.

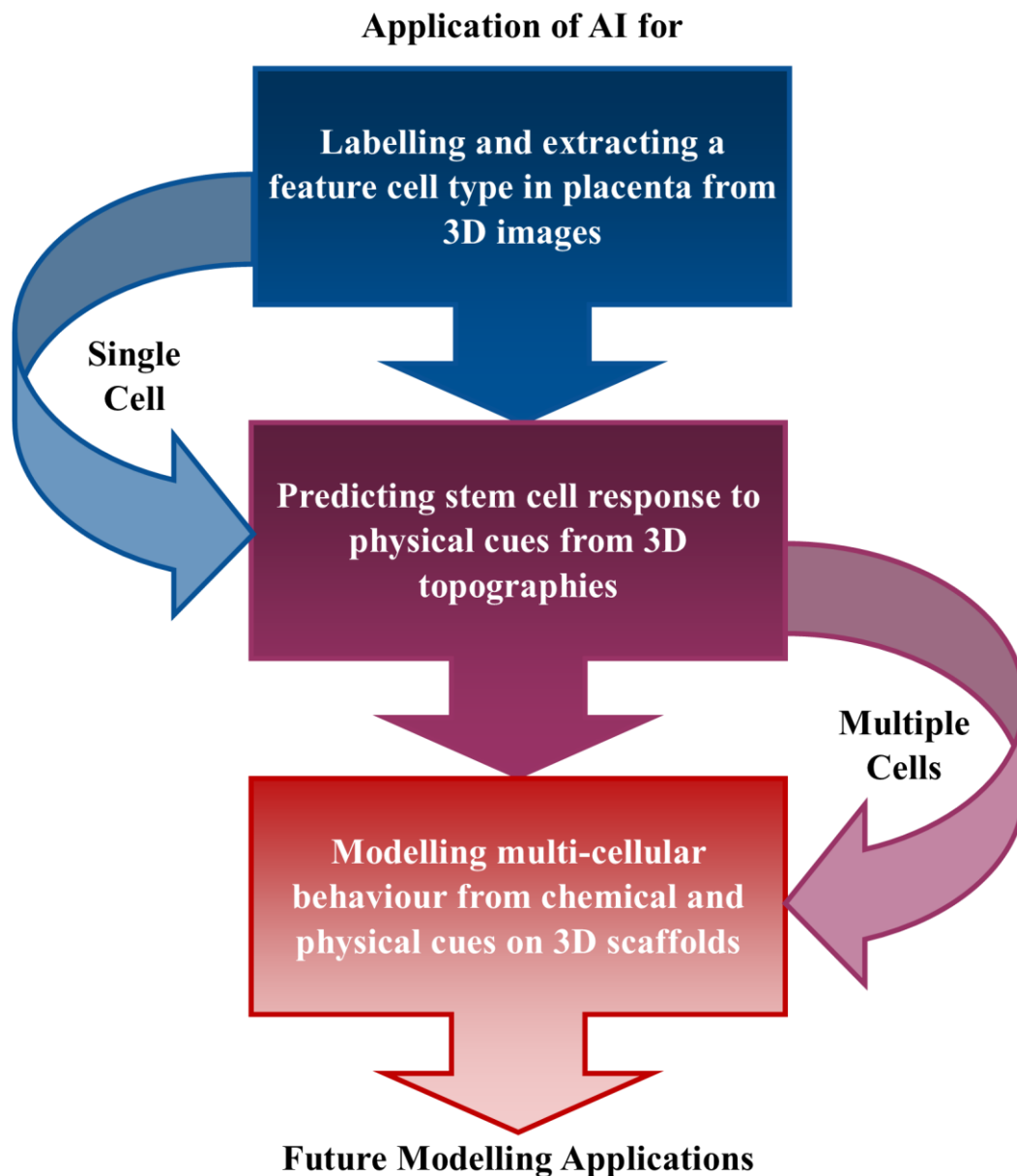


Figure 2: Project outline and overarching premise of application of AI for future modelling applications.

To achieve the overarching premise of application of AI for future modelling applications, there are multiple steps involved, of which three are covered in the project. Initially, AI is applied for labelling and extracting a feature cell. Next, cell behaviour in response to topography is predicted, ranging from a single cell to a dozen per image. Finally, a multi-cellular system containing stem cells is modelled when given unseen combinations of physical and chemical cues. The complexity therefore rises throughout this project from static tasks involving single cells to time-dependent predictions of potentially thousands of cells.

The first relates to automating labelling and extracting a targeted cell type from 3D images, which is covered in Chapter 5. Automation would allow for researchers to spend less time on menial tasks and lead to faster image analysis and a shorter path to scientific discovery.

Additionally, automation of cell labelling would show that computational methods have the potential to distinguish between cells and acellular material, and between cell types, in complex images.

Introduction

The second, covered in Chapter 6, relates to generated statistically likely cell positions in response to unseen topographical cues to detect new cell behaviours. The third, covered in Chapter 7, advances on this work and, instead of investigating predictions of a few cells, an entire cellular system on 3D scaffolds is modelled, with combinations of different chemical and physical cues to influence cell behaviour over time.

Chapter 2 will cover the background to each of the areas covered above. Chapter 3 will introduce deep learning as a tool to accomplish the above aims, and Chapter 4 will detail the data collection and composition for training deep learning models. Chapter 8 concludes on the project as a whole.

1.3 CONTRIBUTIONS AND PUBLICATIONS

1.3.1 SUMMARY OF CONTRIBUTIONS

1. Automated labelling of cellular structures in SBFSEM-imaged placental tissue that can match manually labelled images by over 95%.
2. Determining the minimum line separation of laser-machined lines on glass surfaces resulting in stem cell alignment of $11.7 \pm 1.3 \mu\text{m}$.
3. Developing a method for potentially reducing necessary levels of animal testing in tissue engineering experimentation

1.3.2 PUBLICATIONS

All listed works are peer-reviewed book chapters, conferences, or journal articles. Listing is grouped into publication type and in chronological order rather than order of significance. All first author publications result from this work. Additional publications are either partially resulting from this work or have shared methodology applied to both this work and the relevant publication, developed by the contributing authors. Potential publications under review or in preparation have not been listed for brevity. For an accurate citation count, please follow this link to the relevant [Google Scholar citation record](#).

FIRST AUTHOR PUBLICATIONS RESULTING FROM THIS PROJECT :

Book Chapter

- 2021 **Mackay, Benita Scout**, Grant-Jacob, James, Eason, R.W., Lewis, Rohan and Mills, Benjamin (2021) [Deep Learning for the Automated Feature Labelling of 3-Dimensional Imaged Placenta](#). In, *Biomedical Engineering Systems and Technologies: 13th International Joint Conference, BIOSTEC 2020, Valletta*,

Malta, February 24–26, 2020, Revised Selected Papers. 1 ed. Springer, p. 93.
(doi:10.1007/978-3-030-72379-8_6).

Journal Articles

- 2021 **Mackay, Benita Scout**, Marshall, Karen, Margaret, Grant-Jacob, James, Kanczler, Janos, Eason, Robert W., Oreffo, Richard and Mills, Benjamin (2021) The future of bone regeneration: integrating AI into tissue engineering. *Biomedical Physics and Engineering Express*, 7 (5), [052002]. (doi:10.1088/2057-1976/ac154f).
- 2020 **Mackay, Benita Scout**, Praeger, Matthew, Grant-Jacob, James, Kanczler, Janos, Eason, Robert, Oreffo, Richard and Mills, Benjamin (2020) Modeling adult skeletal stem cell response to laser-machined topographies through deep learning. *Tissue and Cell*, 67, [101442]. (doi:10.1016/j.tice.2020.101442).

Conference Articles and Presentations

- 2020 **Mackay, Benita**, Blundell, Sophie, Etter, Olivia, Xie, Yunhui, McDonnell, Michael, David Tom, Praeger, Matthew, Grant-Jacob, James, Eason, R.W., Lewis, Rohan and Mills, Benjamin (2020) Automated 3D labelling of fibroblasts and endothelial cells in SEM-imaged placenta using deep learning. In *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2 BIOIMAGING*. Science and Technology Publications, Lda. pp. 46-53. (doi:10.5220/0008949700460053).
- 2019 **Mackay, Benita, Scout**, Blundell, Sophie, Lewis, Rohan, Etter, Olivia, Xie, Yunhui, McDonnell, Michael, David Tom, Praeger, Matthew, Grant-Jacob, James, Eason, Robert and Mills, Benjamin (2019) Automated 3D labelling of fibroblasts in SEM-imaged placenta using deep learning. *The Southampton Machine Intelligence Showcase 2019, University of Southampton, United Kingdom*. 22 Oct 2019. 1 pp.

ADDITIONAL OUTPUT:

Conference Articles and Presentations

Introduction

- 2022 (Partially resulting from this work)
Mills, Benjamin, Grant-Jacob, James, **MacKay, Benita, Scout**, Lewis, Rohan and Sengers, Bram (2022) Modelling of 3D placental cell features using deep learning. In *BIOIMAGING2022*.
- 2020 Mills, Benjamin, Grant-Jacob, James, Jain, Saurabh, Xie, Yunhui, **Mackay, Benita, Scout**, McDonnell, Michael, David Tom, Praeger, Matthew, Loxham, Matthew, Richardson, David and Eason, Robert (2020) Particulates sensing using optical fibres and deep learning. *SPIE Photonics West, The Moscone Center, United States. 01 - 06 Feb 2020*.
- 2019 Grant-Jacob, James, Xie, Yunhui, **Mackay, Benita, Scout**, McDonnell, Michael, David Tom, Praeger, Matthew, Loxham, Matthew, Eason, Robert and Mills, Benjamin (2019) Bioaerosol sensing using deep learning. *IOP - Physics Meets Biology, United Kingdom. 08 - 11 Sep 2019*.
- Mills, Benjamin, McDonnell, Michael, David Tom, Heath, Daniel, Xie, Yunhui, Grant-Jacob, James, **Mackay, Benita, Scout**, Praeger, Matthew and Eason, Robert (2019) Deep learning for 3D modelling of multiple pulse femtosecond ablation. In *15th International Conference on Laser Ablation (COLA) 2019*.
- Grant-Jacob, James, Xie, Yunhui, McDonnell, Michael, David Tom, Eason, Robert, Loxham, Matthew, **Mackay, Benita, Scout**, Mills, Benjamin, Heath, Daniel and Praeger, Matthew (2019) Deep learning for real-time air and marine particle pollution sensing. *Clean Air Futures: working together for cleaner air, , Southampton, United Kingdom. 22 Mar 2019*.

Mills, Benjamin, Heath, Daniel, Grant-Jacob, James, Xie, Yunhui, **Mackay, Benita, Scout** and Eason, Robert (2019) Image-based monitoring of high-precision laser machining via a convolutional neural network. In *Photonics West*.

Xie, Yunhui, Heath, Daniel, Grant-Jacob, James, **Mackay, Benita, Scout**, McDonnell, Michael, David Tom, Praeger, Matthew, Eason, Robert and Mills, Benjamin (2019) Smart lasers for manufacturing of the future. *FEPS IoT Showcase*. 05 Apr 2019.

Grant-Jacob, James, Xie, Yunhui, **Mackay, Benita, Scout**, McDonnell, Michael, David Tom, Praeger, Matthew, Loxham, Matthew, Eason, Robert and Mills, Benjamin (2019) Super-resolution imaging using deep learning. *IOP - Physics Meets Biology, United Kingdom*. 08 - 11 Sep 2019

Journal Articles

2020 McDonnell, Michael, David Tom, Grant-Jacob, James, Xie, Yunhui, Praeger, Matthew, **Mackay, Benita, Scout**, Eason, R.W. and Mills, Benjamin (2020) Modelling laser machining of nickel with spatially shaped three pulse sequences using deep learning. *Optics Express*, 28 (10), 14627-14637. (doi:10.1364/OE.381421).

2019 Grant-Jacob, James, **Mackay, Benita, Scout**, Xie, Yunhui, Heath, Daniel J, Loxham, Matthew, Eason, Robert and Mills, Benjamin (2019) A neural lens for super-resolution biological imaging. *Journal of Physics Communications*. (doi:10.1088/2399-6528/ab267d).

Xie, Yunhui, Heath, Daniel J, Grant-Jacob, James, **Mackay, Benita, Scout**, McDonnell, Michael, David Tom, Praeger, Matthew, Eason, Robert and Mills, Benjamin (2019) Deep learning for the monitoring and process control of femtosecond laser machining. *Journal of Physics: Photonics*, 1 (3), 1-10, [3]. (doi:10.1088/2515-7647/ab281a).

Grant-Jacob, James, Jain, Saurabh, Xie, Yunhui, **Mackay, Benita, Scout**, McDonnell, Michael, David Tom, Praeger, Matthew, Loxham, Matthew, Richardson, David, Eason, Robert and Mills, Benjamin (2019) Fibre-optic based particle sensing via deep learning. *Journal of Physics: Photonics*.

Mills, Benjamin, Heath, Daniel, Grant-Jacob, James, **Mackay, Benita, Scout** and Eason, Robert (2019) Neural networks for predictive laser machining capabilities. In *Emerging Digital Micromirror Device Based Systems and Applications XI*. SPIE.. (doi:10.1117/12.2507375).

Grant-Jacob, James, Xie, Yunhui, **Mackay, Benita, Scout**, Praeger, Matthew, McDonnell, Michael, David Tom, Heath, Daniel J, Loxham, Matthew, Eason, Robert and Mills, Benjamin (2019) Particle and salinity sensing for the marine environment via deep learning using a Raspberry Pi. *Environmental Research Communications*, 1 (1). (doi:10.1088/2515-7620/ab14c9).

Grant-Jacob, James, **Mackay, Benita S.**, Baker, James, Xie, Yunhui, McDonnell, Michael D.T., Heath, Daniel, Praeger, Matthew, Eason, Robert W. and Mills, Ben (2019) Patterned nanofoam fabrication from a variety of materials via femtosecond laser pulses. *Materials Sciences and Applications*, 186-196. (doi:10.4236/msa.2019.103015).

2018 Heath, Daniel, **Mackay, Benita S.**, Grant-Jacob, James, Xie, Yunhui, Oreffo, Richard, Eason, Robert and Mills, Benjamin (2018) Closed-loop corrective beam shaping for laser processing of curved surfaces. *Journal of Micromechanics and Microengineering*, 28 (12), 1-7. (doi:10.1088/1361-6439/aae1d5).

Heath, Daniel, Grant-Jacob, James, Xie, Yunhui, **Mackay, Benita, Scout**, Baker, James, Eason, Robert and Mills, Benjamin (2018) Machine learning for 3D simulated visualization of laser machining. *Optics Express*, 26 (17), 21574-21584. ([doi:10.1364/OE.26.021574](https://doi.org/10.1364/OE.26.021574)).

Grant-Jacob, James, **Mackay, Benita S.**, Baker, James, Heath, Daniel J., Xie, Yunhui, Loxham, Matthew, Eason, Robert W. and Mills, Ben (2018) Real-time particle pollution sensing using machine learning. *Optics Express*, 26 (21), 27237-27246. ([doi:10.1364/OE.26.027237](https://doi.org/10.1364/OE.26.027237)).

2 BACKGROUND

Application of AI in regenerative medicine was reviewed in Biomedical Physics and Engineering Express [40].

The scope of this thesis is broad, covering imaging placental tissue, modelling Stro-1 enriched progenitor cells, and predicting scaffold biocompatibility.

With an approximate annual cost of £1.7 billion for the NHS, and costing between 1.0% and 2.5% of gross domestic burden internationally, osteoporosis and osteoarthritis represent a major socio-economic burden [41, 42]. Current approaches to alleviate suffering include medication, bone grafts and implants. The former solution is unsustainable as, while bisphosphonates and non-opioids such as non-steroidal anti-inflammatory drugs (NSAIDs) are prescribed for osteoporosis, NSAIDs can lead to serious medical complications when high doses are taken for a long time or when the patient is elderly/in poor health, and may even accelerate cartilage destruction in osteoarthritis [43-45]. A recent public health report showed 5% of UK citizens are prescribed opioids [46], which have limited long-term benefit and severe issues, including contradictory evidence of pain reduction and the numerous associated risks of opioid use [47]. The use of grafts and implants are not without risk, specifically the risk of rejection and failure, hence there is a growing need for innovative techniques to promote implant integration and reduce the failure rate of osteopathic intervention.

Section 2.1 will expand on the concept of tissue engineering further, 2.2 will focus on topographies and cell response to them, 2.3 will detail possible methods for creating these topographies, 2.4 will branch into scaffolds and animal testing in tissue engineering, and 2.5 will cover an alternative side to regenerative medicine, with emphasis on analysis of placental health, which can affect life-long patient outcomes.

2.1 TISSUE ENGINEERING

Tissue engineering is defined by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) as “an interdisciplinary and multidisciplinary field that aims at the development of biological substitutes that restore, maintain, or improve tissue function” [48]. It is a branch of regenerative medicine, with current research focussed on utilising the natural healing response of the body to regenerate bone, helped by bioactive materials [49]. In the past, bio-inert materials were frequently used in joint replacement as a treatment for osteoarthritis, but outcomes were inconsistent. Even implants not rejected by the body did not

provide an optimal treatment, as a prosthesis lifespan is typically less than 15 years. Fatigue and wear eventually make the implant unfit for purpose, and recipients of knee joint replacements often have to limit their sport activity and heavy lifting to prolong the lifespan, impacting their quality of life [50]. One end goal is bio-fabrication of bone at a scale viable for clinical use, but complexities in resolving vascularisation and necrosis issues remain unsolved [51]. A middle ground, where biomaterials encourage the body to regenerate bone *in-situ*, faces similar challenges, but recent technological advances being introduced to tissue engineering research, such as AI, may result in successful patient-specific advances [52].

2.1.1 STEM CELLS AND BIOMATERIALS

Organs and tissue, including bone, can be repaired and regenerated by the body but this healing response is limited by a variety of factors, including infection, compromised blood supply, and the scale of the trauma- and disease-induced defect. Biomaterials and bone grafts can promote osteoinduction, osteoconduction and osteogenesis, therefore can induce bone healing. Bone healing is a large focus area of tissue engineering, as bone is fast becoming the most transplanted tissue after blood (due in part to the increasing aging populations across the world) [51, 53]. These biomaterials must combine bone regenerative properties with the strength of bone, and also be biodegradable so new bone growth can replace the biomaterial unimpeded. Regrown bone and temporary bone scaffolds must both perform a wide range of tasks, from producing red and white blood cells to storing minerals, which impacts overall health and mobility of the body [26, 54-60].

Heterogeneity exists within skeletal stem-progenitor cell fractions, leading to significant differences in differentiation potential along the osteo-chondro-adipogenic lineages, characteristic of the tri-potent primitive mesenchymal stem cell population, where there is variation in cell stem-progenitor capacity [27]. Lack of specific markers is hampering successful isolation, which is further hindered by mismatches between *ex vivo/in vivo* mismatch in multi-lineage potential [61]. The use of Stro-1 as a stem cell marker is a popular example as, while multiple studies use Stro-1 for the enrichment of HBMSC populations, only a few examined and demonstrated Stro-1 reactivity in the tissues from which the cells originated, such as in sub-endothelial positions near sinusoids (a type of blood vessel within bone marrow) [29, 40, 62]. However, the evidence demonstrating Stro-1 selected populations resulting in fibroblasts, adipoblasts, myoblasts and osteoblasts, shows the necessity of further

Background

marker research, and the validity of using Stro-1 enriched HBMSC populations for *in vivo* biomaterial research [63].

HBMSC populations lose proliferative and, crucially, multipotential ability the longer they are cultured, but the driving mechanisms for this are difficult to overcome [61], which leads to challenges in stem-progenitor cell expansion. However, *in vivo* environments face challenges, too. The safety and performance of *in vivo* biomaterials are difficult to predict and monitor, with issues including inflammation, rejection, and possibly cancer, with evidence mounting against textured silicon-based implants. Unlike with smooth implants, textured implants are linked to non-Hodgkins lymphoma, resulting in a mass global recall due risks of anaplastic large cell lymphoma, Figure 3, [64-66]. Importantly, understanding is still deficient, and the relationship between textures and other biophysical cues with multi-factorial multi-cellular environment (all *in vivo* environments) is complex, both in the short term and long term relative to the life of an implant.

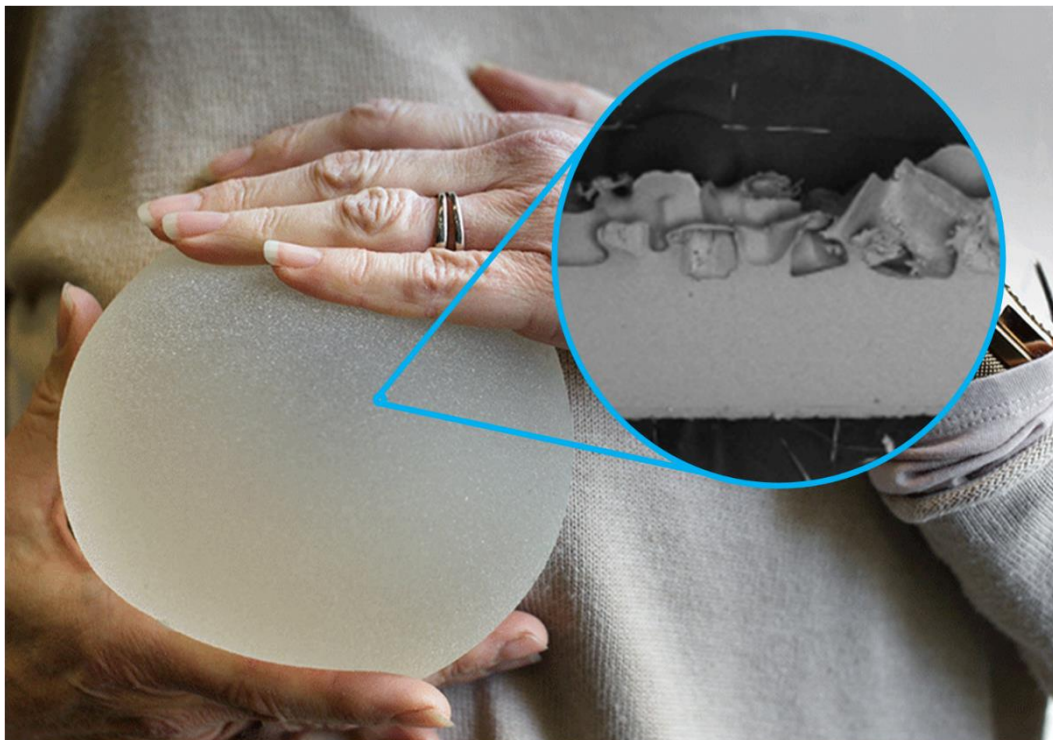


Figure 3: An illustration of the surface of a textured implant.

Textured breast implants were found to be linked to a type of non-Hodgkins lymphoma, with those with the most textured surfaces more likely to result in cancerous side effects during the lifetime of the implant. Smooth implants, made with the same materials, were not linked to non-Hodgkins lymphoma. These textured implants have since been banned . © Searsie, iStock, usage under iStock Standard License (Ref 2082231647) .

Healing is a complex multi-cellular environment with dynamic processes of migration, proliferation, and differentiation. This includes coordination of multiple cell types performing different tasks depending on the location, variety, and duration of trauma, which is a key

focus of tissue engineering [67]. Understanding of natural cues within the body during the immune response has resulted in Extracellular Matrix (ECM) mimicking scaffolds capable of manipulating the immune system for tissue regeneration stimulation but these are yet to be sufficiently scalable, reproducible, and inexpensive for clinical application [68, 69].

However, it shows how crucial understanding of the biochemical and biophysical cues within the body are to successful biomaterial development.

There are a variety of cues to consider, from biophysical features such as structure and topography, to biochemical features, such as beneficial cytokines and ions. Surface topography alone has been found to alter cell behaviour through both intracellular and intercellular signalling, affecting cytokine secretion [70], proliferation and differentiation [71], density and spreading [72, 73], and cell adhesion [74, 75]. Each feature, and the cumulative effect, must be considered when designing a new biomaterial [76]. Especially when the properties of bone can vary patient-to-patient, with factors including race, sex, age, and position in the body [77].

2.1.2 A PROBLEM OF SCOPE

Originally considered little more than a passive frame for cells, the ECM is currently thought to be a crucial actor in the modification and maintenance of cell behaviour. Biophysical cues modulate the ECM-cell interaction to create new material-cytoskeleton links, where surface topographies instruct cells via modifications to cell adhesion and, subsequently, morphology [40]. Therefore, cell spreading, migration, proliferation, and differentiation can be modified. Biomaterial topography is an important feature, with an emerging consensus that topography is as effective as biochemical cues and signals in controlling cell fate and function [78-80]. Many reviews cover how different physical characteristics alter the behaviour of cells [60, 76, 78], especially in the nanoscale [12, 16, 70, 81-83].

The complexity of cell response is such that, with one single area (topography), there are virtually infinite unique biophysical feature combinations to investigate. Biophysical features such as, roughness, ordering, and shape are a non-exclusive list of examples. These can be further divided into, for example, average height deviation amplitude, density of summits, texture aspect ratio, skewness, and grating (regularly spaced collection of essentially identical, parallel, elongated elements) ratio. A grating can include topographical features such as line separation, line thickness, grating length, increased disorder, and intersecting angles. Taking a small fraction of the available scales, investigating a range of sizes from 100

Background

μm to 100 nm with 10 nm precision, would increase the number of potential experiments by an order of magnitude. When two features are combined, for example line thickness and line separation, the number of unique combinations is increased by several orders of magnitude. If all the features listed are combined, just at the scales mentioned, the count of unique combinations becomes larger than 1×10^{30} . The fastest part in the cell culture- imaging-analysis chain is cell imaging with an automated fluorescent microscope. This could acquire both a bright field and fluorescent image on a single topography every millisecond, therefore taking 30 quintillion years without any repetition (substantially longer than the current approximation of the age of the universe of around 13.7 billion years [84]) to capture each topography [40].

However, just as a human does not need to see every feature to recognise a pattern, investigating every unique combination is superfluous. In another field, an AI-based model did not need to see every laser machined structure possible to determine the 3D depth profile of an unseen laser beam profile [13]. Only from the varied training set, the laser-machining model was also able to determine further useful laser-machining parameters (such as laser power and write-speed) without any underlying physical processes encoded (no physics was taught to the AI model before exposure to training data). Therefore, it is worth investigating how an AI-based model can be applied to regenerative medicine for the same purpose – determining useful biophysical features without needing underlying scientific understanding.

2.2 TOPOGRAPHY AND CELLS

Previous studies have shown topography can be used to influence cell behaviour, with preliminary results showing nanotopographies altering proliferation and promoting osteogenesis in partially disordered nanotopographies [70]. However, these studies have been relatively limited in both the scale and number of nanopatterns investigated, sticking with either uniform groove/ridge patterns or pit patterning of varied disorder. Nanogratings typically enhance cell adhesion and lower proliferation, while nanoposts and nanopits can reduce cell attachment and produce varied effects on proliferation [30]. To comprehensively test which nanotopographies are most effective (i.e. gratings, pits, randomised, or patterning not yet investigated) using traditional methods of culturing cells on virtually all patterns, with consequent analysis, would take an unfeasible amount of time. To elaborate, if only gratings and pits were used, that is two experiments. However, the variation in grating could include line separation, line thickness, grating length, increased disorder, different angles, etc., and the pits could be circular, triangular, rectangular, square or hexagonal (or any other shape).

That is now 10 experiments and a non-exhaustive list. Investigating a range of sizes from 100 μm to 100 nm with 10 nm precision would increase the experiment number to 1×10^3 . When two variations are combined, for example line thickness and line separation, the experiment count is at least 1×10^6 . If all the variations previously listed are combined, the experiment count becomes larger than 1×10^{30} . Just considering the imaging of topographies, the fastest part in the cell culture- imaging-analysis chain, if an automated fluorescent microscope can take a brightfield-fluorescent image pair on a single topography every millisecond, imaging alone would take 30 quintillion years (substantially longer than the current approximation of the age of the universe of around 13.7 billion years [84]).

Recent advances in machine learning give the possibility that, when given sufficiently varied training data, a relatively small training data-set can result in accurate predictions [13]. Training AI on varied, yet limited, nanopatterned surfaces and the subsequent cell response could result in predictions of cell response on untested nanotopographies. This approach could then lead to the determination of optimal topographies without the need for extremely lengthy experimentation and analysis. For example, while imaging all of the combined topographies would take longer than the age of the universe, combining only two of those features and imaging at one a second could be done in a period of weeks, rather than millennia.

For generation of varied topographical surface patterns on substrates for subsequent cell growth, femtosecond laser machining offers both the speed and the precision required for sub-micron scale surface ablation, with the versatility of laser machining also offering the possibility of directly machining implants [36]. Due to the sensitivity of this non-linear process, small levels of experimental noise can result in inferior fabrication quality, which could prevent necessary scaling up of nanotopography production. However, recent developments in real-time laser machining monitoring through artificial intelligence provide a direct solution.

2.2.1 CELL-TOPOGRAPHY INTERACTION

Cells respond to the shape of their environment due to changes in contact guidance, cell spreading and contact inhibition [85] (Figure 4). Cell behaviour can therefore be influenced through the nanoscale engineering of surfaces and use of surface-directed nanobiotechnologies [86, 87]. Variations to surface nanotopography has been demonstrated

Background

to affect cell adhesion [74, 75], density and spreading [72], cytokine secretion (important to cell signalling) [71], proliferation and differentiation [16, 70].

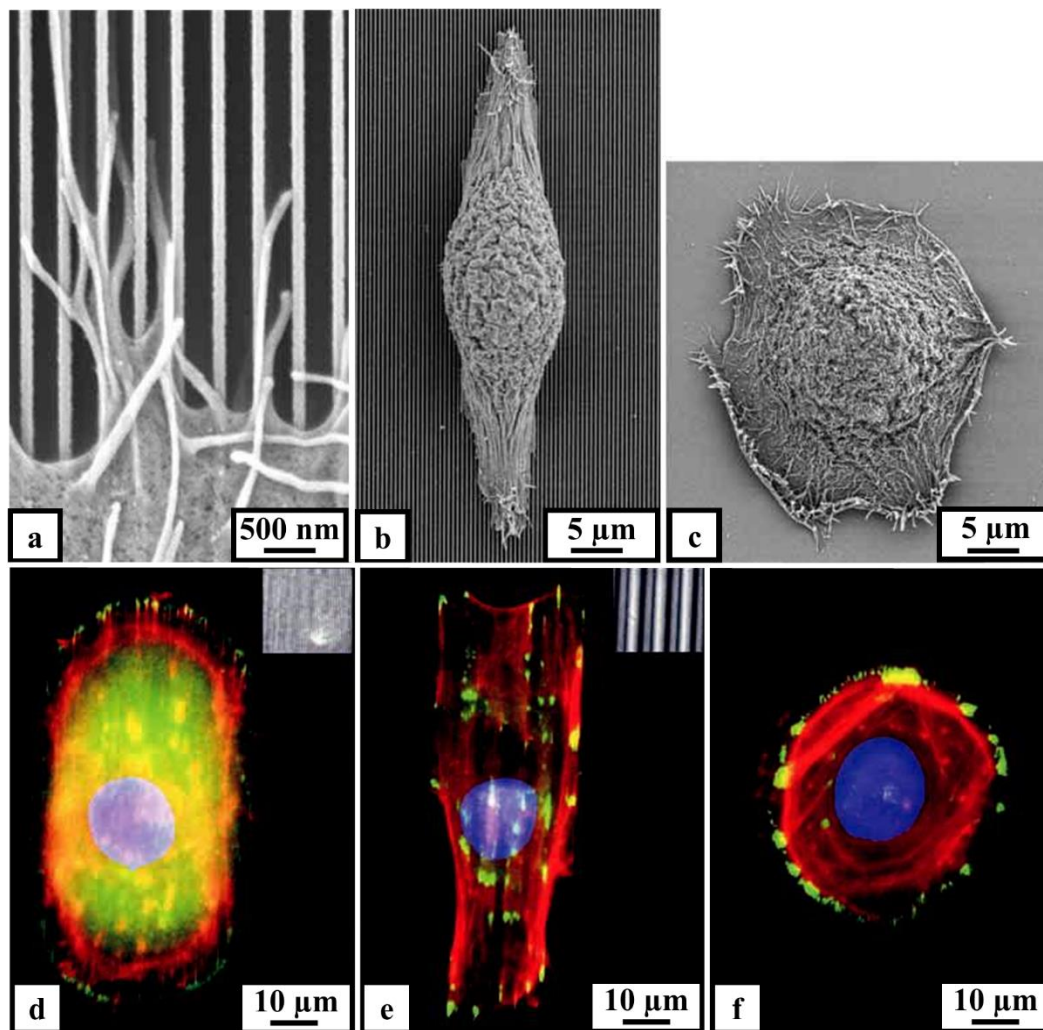


Figure 4: Elongation and alignment of human corneal epithelial cells parallel to nano-grating topography. (a-c) SEM images of cells cultures with lined gratings and on a flat surface: (a) Filopodia extend along a 400 nm pitched nano-grating, where pitch is the sum of the groove and ridge width, with some lamellipodia at the cell edge protrude into the gratings while others bridge the grating (bottom left); (b) An epithelial cell aligned and elongated parallel to the 70 nm wide ridges, with a depth of 600 nm, on a silicon oxide substrate; and (c) There is no visible elongation when an epithelial cell is cultured on a smooth silicon oxide substrate. (d-f) Fluorescent images of cells stained for actin (red), vinculin (green), and the nucleus (blue) with a reflection image of the substrate on the top right for illustrative purposes: (d) An epithelial cell cultured on the same 400 nm pitch as seen above, with large expression of vinculin, a membrane cytoskeletal protein involved in cell adhesion; (e) An epithelial cell on a grating of 4000 nm pitch with 1900 nm ridges at a depth of 600 nm, with less vinculin throughout the cell; and (f) An epithelial cell on a smooth silicon oxide substrate, where vinculin is concentrated at the cell edge. Grating inserts in (d, e) are to scale with accompanying cells. All images modified from Teixeira et al. with permission from The Company of Biologists Limited [9].

Observed since 1911, contact guidance is responsible for the alignment of cells with topographies of comparable size to the cells under observation [85]. Grooves, both of magnitude close to cell size (typically 5 to 10 μm) and smaller, are capable of orientating cells, where both the mechanical properties of the cytoskeleton and durability of cellular

attachment is affected by topographical variations for multiple materials, including silica and titanium.

At the sub-micrometre range, surface topographies can mimic cell adhesion surfaces such as other cells and proteins in the ECM. Low-voltage scanning electron microscopy has shown the ECM has a complex topography at cellular and subcellular level, with many features of the order of 50 to 100 nm. At nanoscale range, it may be possible to target receptor-driven pathways to directly control the response of cells [16]. For example, grating surface topographies with an ultrafine period of 260 nm of varying depths between 100 and 400 nm resulted in changes in cell behaviour, as seen in Figure 4, although it is highly dependent on cell type and cell-cell interactions – different cells interact in different ways to the same topography, with some cells being more responsive than others [88].

While exploration into topographical effects has grown, there have been many parallel studies on the addition of bioactive ion chemistry to titanium implants due to the potential for improved osteogenic capacity [89-96]. When the surface of a titanium implant is modified with calcium ions, chemical and nanotopographical alterations accelerate platelet activation, moderating the early phase healing of bone around the implants [97]. Strontium ions have also been used to modify titanium, benefitting bone healing [98, 99]. However, there was difficulty in concluding whether the nanotopography or surface ion chemistry was dominant for osteogenic improvement. When nanorough titanium surfaces with and without ion modification were used to culture murine bone marrow stromal cells (MSCs) and human multipotent adipose stem cells (hASCs), strontium-modified titanium was more potent than other surface factors for osteogenesis of MSCs [96], yet nanorough topographies were shown to be not as osteoconductive as nanotopographies with more ordered patterning [35, 100-102]. The introduction of ions into the body also risks toxic and potentially dangerous long-term side effects, which have not yet been sufficiently studied [103]. To compare how the combination of nanotopography and surface chemistry effects cells, a 2016 compared the effects of nanotopography with a variety of surface chemistries and found nanopatterning resulted in changes to macrophage behaviour independent of the surface chemistry [104]. This result shows strong evidence of the possibility to control stem cells through the patterning of surfaces on which the cells are cultured, without the need for potentially harmful substrate doping.

Background

2.2.2 TOPOGRAPHICAL EFFECTS

Cell interactions with various surface topographies impact both the directionality and speed of cell spreading [105], with proteins in the ECM hypothesised to be responsible for much cell signalling [106]. The most common way for proteins to interact is through the precise matching of one rigid surface and another, and the communication between cells is mediated mainly through extracellular signal molecules. Extracellular signals are relayed through plasma membrane receptors to the interiors of the cells, allowing recognition of adjacent cells and extracellular structures [107].

Organised ECM proteins modulate contact guidance, an essential component in cell migration regulation through the response of cells to micrometre and sub-micron scale surfaces [108]. As each cell is programmed to respond to specific combinations of extracellular signals, the key challenge is to determine the exact surface topographies that will induce desired cell behaviour, such as differentiation (the cell changes from one type to a typically more specialised type), or proliferation (the cell divides and the overall number of cells increases), for the target cell. This is difficult, as cell and topography interactions result in differing effects within a singular cell type due to the direct effect of the topography combined with the physiochemical properties of the surface substrate [106].

There is a minimum scale at which cell guidance is no longer applicable [109]. When the morphological behaviour of fibroblasts cultured on smooth and grooved nanotopographies were investigated, a cut off of 100 nm width and 75 nm depth was established after 4 hours for visible cell response. However, after 24 hours this reduced to groove depths of 35 nm. These limits are similar to the smaller sized features of an ECM, giving further weight to proteins of the ECM being responsible for important cell signalling.¹

Studies have shown for filopodia (thin cytoplasmic projections that extend beyond the leading edge of lamellipodia in migrating cells), and partially for lamellipodia (a cytoskeletal protein actin projection on the leading edge of the cell which propels the cell), that features as small as 10 nm generate a response [111] (Figure 5). Filopodia are extended from the main body of a cell and probe the surrounding environment. Cell movement then occurs in the direction of the contact through retraction of attached filipodia. Below this scale, contact

¹ The basement membrane of the corneal epithelium of the rhesus macaque was found to be composed of a varying nanotopography consisting of elevations ranging from 76 to 379 nm and pores ranging from 22 to 216 nm with a mean interpore distance of 87 nm¹¹⁰. Abrams, G.A., et al., *Nanoscale topography of the basement membrane underlying the corneal epithelium of the rhesus macaque*. Cell Tissue Res., 2000. **299**: p. 39-46..

guidance does not exist. Individual collagen fibrils have diameters typically 20-100 nm and commonly form larger aggregates [112, 113]. As cell guidance occurs below 100 nm, there is likely a combination of factors affecting cell behaviour: Simply mimicking cell surface topographies or protein sizes of the ECM may not be the most efficient nanotopography for cell control.

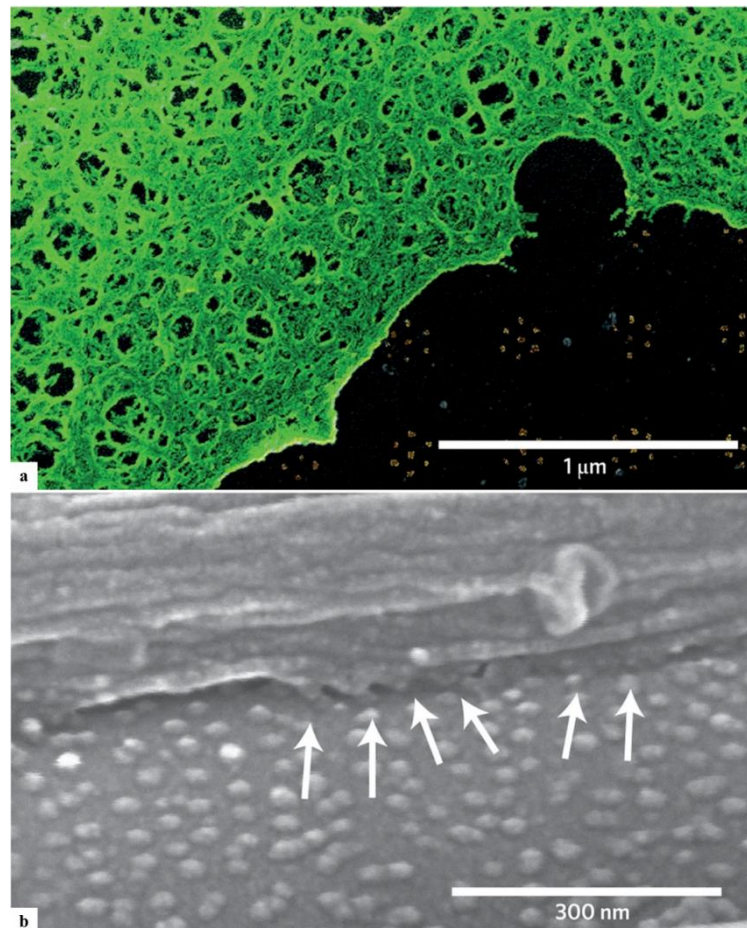


Figure 5: Cell spreading and interaction with nanoscale topography. Features are small enough in the nanotopography to be similar in size to individual cell receptors, specifically integrins (nanoscale) responsible for adhesion (microscale). (a) A heptameric arrangement of ligand clusters showing that four integrins are needed to be gathered, forming an adhesion, to result in cell spreading. (b) Nanopodia (arrows), nanoscale filopodia (cell membrane projections containing integrins) responding to nanotopography, which can result in whole-cell response. Image reused from Dalby et al. with permission from Springer Nature [16].

Real-time polymerase chain reaction (PCR), a technique in molecular biology to evaluate the amplification of targeted DNA using fluorescent signalling, of human adipose tissue-derived mesenchymal stem cells show an interaction between nanotopographical cues and microRNA activity, which lead to osteogenic differentiation [114]. However, disparities between cell types, surface types and fabrication methods, and specific capabilities of research laboratories involved, lead to great differences and sometimes contradictions in topographical response [115]. While there is no universal hypothesis relating to the mechanism for the effect of cell

Background

nanotopography interactions on cell behaviour [106], it is speculated that changes to cytoskeletal tension in response to topography could change interphase nucleus organisation, therefore influencing the gene expression profile of the cell, with work utilising gene array techniques showing that nanotopographies effect MSC genome expression [100, 116-123]. Further studies are needed on more complex and randomised patterns to best investigate cell response to topographies.

2.1.2 INFLUENCING ONTOGENESIS

The differentiation of MSCs into desired cell lineages can be done in a variety of ways, with topographical methods having the potential to influence MSC differentiation into adipogenic, chondrogenic and osteoblastic cells [100, 102, 124]. MSCs are thought to differentiate to osteoblastic cells through adhesion elongation and increased intracellular tension, which leads to activation of osteo-specific transcription factors, such as RUNX2, through phosphorylation.

Controlling this differentiation can prevent MSCs differentiating into soft tissues, which leads to repeat surgical intervention for weight-bearing implants in orthopaedic procedures or low joint mobility in non-weight-bearing implants, currently a problem with topographically-unaltered titanium [102]. The most common mode of implant failure is aseptic loosening, accounting for 24.3 and 29.8% of all failures in total hip and total knee arthroplasties respectively [125, 126]. A correctly topographed implant would encourage differentiation leading to mature osteoblasts, allowing the production of osteoid - essential for new bone formation [100].

Nanorough surfaces have been shown to produce an accelerated osteogenic differentiation-promoting effect in hASCs [96], as the nanosurface has a larger surface area and superior wettability, which is beneficial to osteogenesis due to inducing earlier cellular response from bone-forming cells through an increased absorption of cell adhesion proteins [89, 92, 95].

When surfaces are adapted with disordered nanoscale patterns, human MSCs can be stimulated to produce bone mineral without added osteogenic supplements, where limited disorder to nanotopography is beneficial to osteogenesis [35, 100, 102]. Using electron beam lithography (EBL), increasingly disordered nanotopographies were utilised to compare the effect of disordered nanotopography on MSCs including a uniform square array, adapted square arrays (where individual nanopits were ± 20 nm, near square (NSQ) 20, and ± 50 nm, NSQ 50, from uniform central positions) and completely randomised nanopit positioning.

Bone-specific ECM proteins Osteopontin (OPN) and Osteocalcin (OCN) were detected in NSQ 20 and increasingly for NSQ 50 cultured MSCs after 21 days via immunofluorescence but was not visible for the control (un-adapted), wholly random or uniform nanotopographies, as seen in Figure 6. Bright-field images further concluded that mature bone nodules containing mineral were found in NSQ 50 while MSCs on the control had a fibroblastic morphology after 28 days [100].

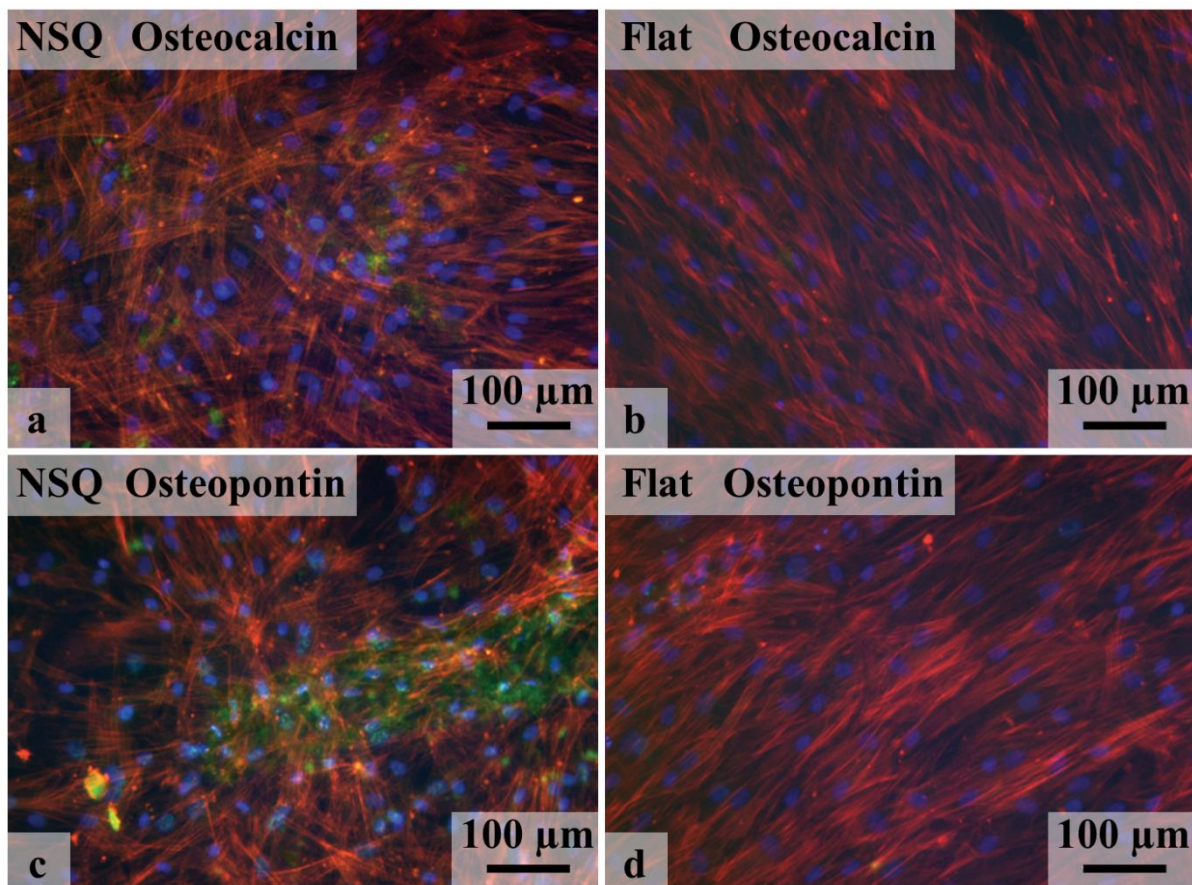


Figure 6: Skeletal stem cell response to different nanotopographies. Skeletal stem cells were cultured for 28 days on (a, c) partially randomised square array nanopits (NSQ) or (b, d) planar controls (FLAT). Stem cells were stained with actin (red) and nuclei (blue) for cell positioning and quantity, and immunofluorescence stained for bone cell markers osteocalcin (a-b) and osteopontin (c-d). Images reproduced from McNamara et al. under Creative Commons Attribution License open access (© 2010 Laura E. McNamara et al.) [12].

Controlled nanodisorder clearly demonstrates an efficient way to control MSC differentiation. In highly ordered surfaces, adhesions were forced to form in the limited areas between pits, as they were unable to form directly over the features. In disordered patterns, large super-mature adhesions could form. Previous studies have described adhesion-promoting fibronectin microscale shapes as beneficial to osteogenesis because MSCs are encouraged to spread, so larger adhesions are observed and consequently osteogenesis is promoted - MSCs differentiate towards osteoblasts [117-120]. Disordered nanotopographies

Background

result in areas where large adhesions can form through increased integrin gathering but also areas where it is harder for adhesions to develop. This allows for the formation of fewer but larger adhesions in osteoblasts than in, for example, fibroblasts [35]. Without such nanotopographical adaptations, the response of MSCs in bone marrow to a titanium implant commonly leads to production of soft tissue, resulting in possible failure of orthopaedic implants due to micromotion or poor joint repair.

These results are significant as they confirm the control of stem cells can be achieved through nanotopographies in this specialised case of nanopits of increasing disorder. However, the huge number of features needed for full exploration, understanding and optimisation of surface topographies for stem cell control requires AI support to limit the time otherwise required for full manual experimentation and analysis.

2.3 TOPOGRAPHY CREATION

There are numerous ways to generate bioactive materials that utilise topographical biophysical cues, with multi-dimensional methods (such as 3D-printing and hydrogel networks) and surface-directed methods (including as lithography and laser-machining) described and compared in many reviews [49, 60, 76-78, 81, 127, 128]. To keep the time needed for data collection to a manageable level within the scope of this PhD, focus was given to surface topographies, rather than branching into more depth-dependent designs. While using AI negates the requirement to fully-investigate all feature combinations, a varied dataset is required for initial training to prevent overfitting, and to aid the network in interpolation of unseen variations in future testing and application. Imaging cells in 3D can be performed using a network, as shown by the Serial Block-Face Scanning Electron Microscopy (SBFSEM) imaging in chapter 5, but these methods require implementation of additional networks and result in a substantial increase in complexity. Additionally, the success of using a network to predict cell response to 2D-cues should be demonstrated before moving on to more complex tasks, including how 3D scaffolds can impact biocompatibility investigated in chapter 7, to determine whether a network is capable of predicting such a multifaceted environment and whether the currently available GPU and computing power are sufficient.

Methods for surface topography fabrication can be largely segregated into two groups – those which produce highly controlled topographies, such as electron beam lithography (EBL) and photolithography, and those which produce unordered or difficult-to-control topographies,

such as polymer mixing, phase separation, colloidal lithography and chemical etching. Photolithography is an established method for generating microtopographies. The near-UV beams create highly customisable designs at relatively low cost, which are repeatable, as the same mask can be reused numerous times without risk of deterioration. However, while sub-diffraction scales are possible, the techniques are not yet as cheap and versatile as required for the amount of unique data-inputs deep learning typically requires – the cost of creating a unique mask for hundreds of unique topographies, vital for a varied dataset, makes photolithography an inefficient method [129-132]. The cost of EBL is a similar limitation, with advanced scanners costing up to £35 million and a set of advanced masks costing up to £2 million [133]. Training a DNN requires multiple variants of topographies, therefore multiple masks would be required and the subsequent cost of using EBL unrealistic. While many topographical patterns could be created on a single mask, this would not allow for iterative changes to the dataset after pilot experimentation. A more flexible approach was required but switching to faster and more inexpensive methods such as colloidal lithography would not provide the control needed over resulting surface topographies to generate repeated detailed patterning on command. Additionally, randomised surface nanotopographies from polymer demixing, electrospinning, phase separation, particulate leaching, chemical etching, and 3D-printing create a variety of interactions which, in some instances, may lead to toxicity [103], and poorly designed topographies may even result in cancer [64-66]. As a highly versatile, cost-effective, and easily controlled method for generating surface topographies, laser-machining was investigated in greater detail.

2.3.1 FEMTOSECOND LASER-MACHINING AND AVALANCHE IONISATION

Explored over several decades, femtosecond laser sub-micron laser-machining is particularly suited for micron to sub-micron scales given pulsed lasers can create peak intensities much higher than achievable by continuous wave lasers [134]. This is partially caused from the interaction of femtosecond laser pulses with matter, leading to optical breakdown with an energy transfer potentially much faster than thermal effects [135, 136]. Additionally, multiphoton absorption can cause laser-machining regions smaller than the focal region, resulting in an ability to machine features as small as 20 nm in highly optimised experimental configurations (far below the diffraction limited near-UV photolithography), further aiding sub-micron laser-machining [137]. Furthermore, as this process is highly non-linear, the repeatability for micron and smaller scales is challenging.

Background

Femtosecond laser pulses interact with matter differently to long pulse and continuous wave light, providing advantages for topography generation. One difference is the time-scale of energy transfer from laser pulse to electrons – most thermal effects proceed after the pulse has passed through the ablation area, with the heat-affected zone resulting from a femtosecond pulse generally less than $\sim 1 \mu\text{m}$ [37]. The other difference is the peak of the pulse, as intensities achievable with femtosecond pulses are sufficient for multiphoton absorption, which enables machining in materials that are optically transparent at low intensities, such as glass. Linear ionisation typically dominates in free electron generations for continuous or long pulse laser-machining [138]. However, the intensities of femtosecond machining can lead to avalanche ionisation, multiphoton ionisation, and tunnel ionisation. As these ionisations are relatively independent of initial defects of the target material [139], femtosecond machining is both possible and reproducible at micron and sub-micron scales, if all laser-machining parameters are stable so as to reduce non-linear effects, with laser-machining areas further reduced by multiphoton absorption (providing an effectively smaller focus) [140].

To fully understand why femtosecond laser-machining is advantageous, avalanche ionisation is broken down into more detail. For femtosecond laser pulses, photo-ionisation at the leading edge of the pulse can provide a “seed” electron for avalanche ionisation during the remainder of the pulse, resulting in at least one electron already in the conduction band of machined material [141]. Avalanche ionisation involves this seed electron absorbing several laser photons until the energy exceeds the conduction band minimum by more than the bandgap energy. This free electron then impact-ionises other electrons, leaving multiple electrons at the conduction band minima [142]. While the laser field is present, this process repeats, creating an exponential increase of electron density and increasing the plasma frequency of the electrons towards the frequency of the laser radiation. At critical plasma density, the plasma absorbs laser energy rapidly through free-carrier absorption (the process of an already excited carrier electron absorbing a photon, resulting in excitation to a higher energy state) [141]. In semiconductors, these excited electrons eventually become scattered by lattice vibration or impurities, but the duration of the laser pulse is much shorter than the electron-lattice relaxation time. Therefore, heat conduction through lattices is negligible because absorption is completed before the lattice changes, resulting in a significant non-equilibrium state between the lattices and electrons [139]. Consequently, (Figure 7) while some remaining energy below the optical penetration depth diffuses into the sample and may cause emission

of particles and droplets due to thermal boiling, thermal damage such as micro-cracks, melting and heat-affected-zone are minimal [143, 144].

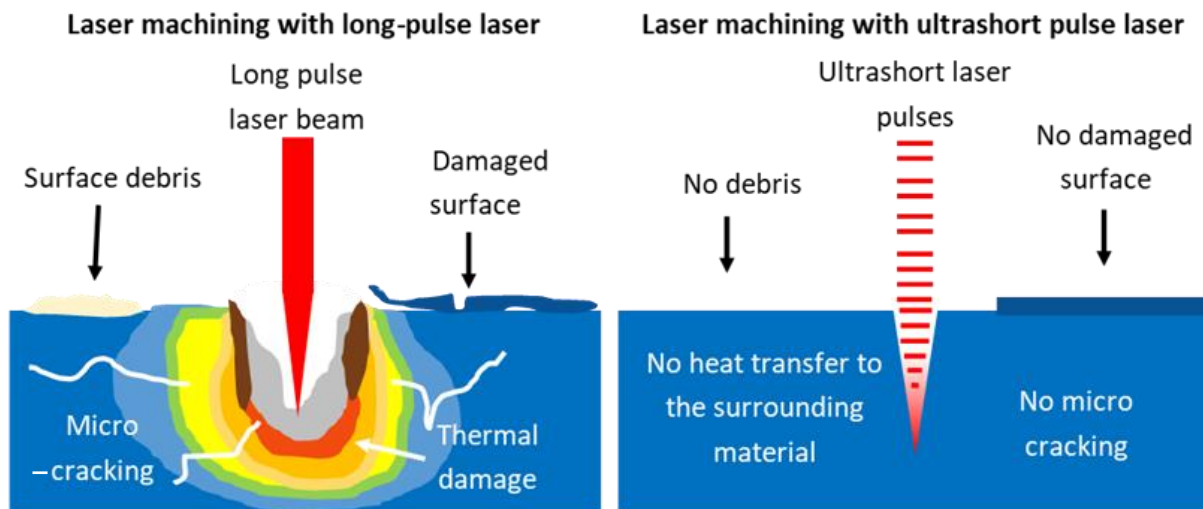


Figure 7: A schematic difference between processing of long pulse (μs) laser machining with ultrashort (fs) laser machining. Based on figure from Miyachi & Jenoptik [3], this is an idealised scenario and, in non-ideal experimental set-ups of ultrafast pulse lasers, debris, surface damage, etc. can still occur, but are relatively negligible in comparison to long pulse machining.

High precision ablation is producible on a wide range of materials, including (but not limited to) metals [140], ultra-hard materials [145], polymers [146], dielectrics [147], semiconductors [148], two-dimensional substrates [149], transparent materials [138], and biological materials [150]. This is due to previously mentioned multiphoton absorption, and subsequent non-linear dependence on intensity. Consequently, the higher peak power of femtosecond pulses results in a higher probability of the process occurring. Femtosecond machining therefore allows for precise machining of small features at high resolution on multiple materials, allowing the potential for micro and nanotopographies to be produced on glass, cell culture plastic, bone, or directly onto implants (e.g. titanium).

2.3.2 ADVANTAGES OF LASER ABLATION

An advantage to laser-machining over, for example, EBL for nanotopographies, is the possible flexibility of production. The standard approach for decreasing the topography fabrication time, when using laser machining, is to increase the area machined per laser pulse, therefore reducing the time spent moving the laser/sample. While beam shaping through a digital micromirror device (DMD), a type of intensity spatial light modulator, can produce a small laser-machined surface area in a relatively short time period [151], it would be time intensive to use such a method over a large surface area, such as a full implant. One way to achieve this is to utilise an intrinsic property of light – interference. In 2003, fabrication of

Background

periodic nanostructures through beam splitting with diffracted elements was manipulated when laser-machined on polycarbonate samples [152]. Over a decade later, this technique was further advanced with interference from six laser beams [153]. Six beams forming a six-sided period have the ability to produce a large variation of interference patterns through alterations in phase shift and amplitude among the beams [154]. These controlled interference patterns could be used in an ordered lattice of multiple uniform topographical features, similar to those produced through standard EBL but in a shorter time period, as lengthy mask-creation is not required for each new design. Controlled disorder can be introduced to these interference structures through small “random” alterations to one or more beams as to alter the pattern compared to the previous ablating pulse. Although the work in this thesis focuses on single-beam ablation as variation was prioritised over mass-production, laser-machining could have the potential to be as efficient at larger scale as lithography under optimised conditions when flexibility and introduction of disorder is necessary. The ability for a 1 fs pulse length laser-machining system to create a variety of 2D and 3D designs makes it a superior choice for neural network training data fabrication [155].

Another advantage to laser-machining is the ability to ablate topographies directly onto non-flat surfaces, including surgical implants, or samples with geographical imperfections. However, laser-machined topographical features ablated over curved surfaces, such as a typical titanium hip replacement implant, could result in skewed and non-uniform topographies that produce undesired cell behaviour. While a 3D sample may be rotated and spatially adjusted to provide an orthogonal surface to the laser beam, complex curved surfaces will ultimately skew or stretch the projected beam profile and complete orthogonality across the whole beam shape it not possible. Recent laser-machining techniques provide a way to counter these potential aberrations via closed-loop adaptive beam shaping (Figure 8) [36], where mask transformations applied to a DMD provide distortion-free laser beam shaping [156]. By introducing a closed-loop feedback mechanism, variation to the surface at different positions can be quickly detected and the spatial intensity and shape of the laser-beam modified [36]. When used alongside iterative corrective steps, DMD alterations can also adjust for spatial inaccuracies introduced by i.e. movement of positioning stages. Additionally, closed-loop correction with the utilisation of a CNN aids development of real-time correction to laser-machining [38], which has the potential for the laser-machining of design-perfect topographies, irrespective of material quality and uniformity.

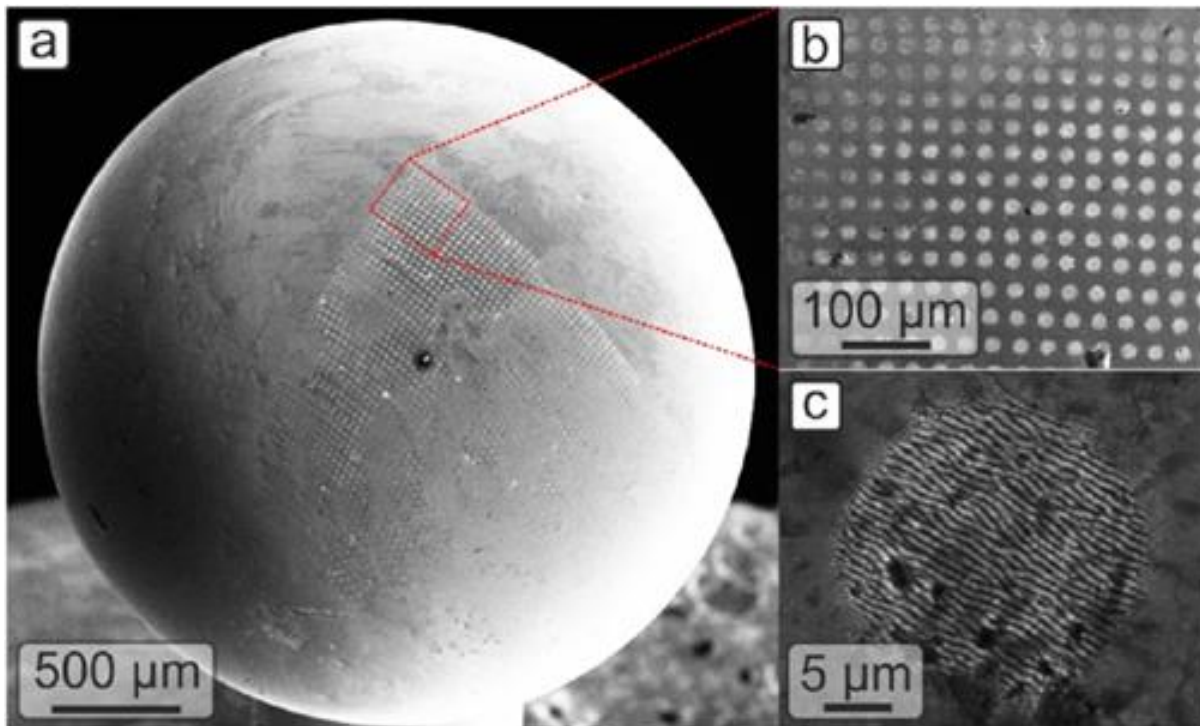


Figure 8: Scanning Electron Microscope (SEM) images of micromachining over a sphere through use of a Digital Micromirror Device (DMD) and computational correction.

The uniformity of the circles across the sphere was the result of computationally-calculated corrections to the beam spatial intensity profile, achieved through alterations to the pattern on a DMD, depending on the location of the beam-surface interface. Images reproduced from Heath et al. under Creative Commons Attribution License open access [13] (© 2018 Optics Express, published by The Optical Society.)

However, there are several challenges involved in laser-machining of this nature. As a result of the highly non-linear processes involved in laser ablation, where small fluctuations in laser output power can result in large effects at a topographical level, some areas may end up laser-machined and others not, possibly producing a patchy and non-uniform topography. This can be overcome, such as through CNN-monitoring of laser-machining [38], yet this reduces the overall speed and productivity of laser-machining, especially over multiple topographies. Consequently, laser-machining often takes time and monitoring of the resultant topographies can be labour intensive. Another challenge is in obtaining the optimal parameters for laser-machining – such as the correct combination of laser write-speed (how quickly the beam moves across the substrate or vice versa), DMD mask shape/equivalent focussed spot size (affecting the shape, size and resultant power of the laser-beam at the substrate) and laser power. These challenges combine when acquiring the threshold power (the minimum laser power required to ablate a specific substrate for a set write-speed and DMD shape) required for precision laser-machining due to differences in material properties.

2.4 BIOMATERIALS AND ANIMAL TESTING

In situ testing involves a highly complex and dynamic environment, with many intertwined features in the bioengineered composite, alone. There are currently no traditional computational methods capable of accurately simulating *in situ* testing, partly due to the relationship between features (or even what all of the features are) not being fully understood. Therefore, highly-likely statistical predictions via universal approximators (neural networks), are a potential alternative.

2.4.1 BIOCHEMICAL CHARACTERISTICS

A combination of biophysical and biochemical features results in a wide range of biomaterial properties, many of which have been detailed by Li *et al* [1], including improved proliferation and targeted differentiation (Figure 9), and Gaharwar *et al* [60]. Using a wide range of manufacturing techniques, from additive manufacturing processes such as 3D printing to ablation processes such as 2D laser femtosecond laser machining, biomaterials can be tailored for targeted cell behaviour [11, 128].

Biochemical characteristics can include the chemical structure of the material, the introduction of inflammatory-suppressing agents, reprogramming factors, cytokines, growth factors, and various minerals and ions. Different biochemical cues, combined with different biophysical cues, can result in a varied cell responses [15] (Figure 10). Changes in biochemical cues, such as increased release of signalling ions (altering the local microenvironment), can be achieved as the biomaterial breaks down [60].

Category		Effects on behaviors	References
Biophysical	Stiffness	Proliferation	Whitehead et al., 2018
		Spreading	Engler et al., 2006; Wingate et al., 2012
		Differentiation	Du et al., 2011; Shih et al., 2011; Higuchi et al., 2013; Sun et al., 2018
	Stress relaxation	Proliferation	Cameron et al., 2011; Chaudhuri et al., 2016
		Spreading	Cameron et al., 2011; Chaudhuri et al., 2016
		Differentiation	Cameron et al., 2011, 2014; Darnell et al., 2017
	Topography	Proliferation	Zhu et al., 2005; Jahani et al., 2012
		Spreading	McBeath et al., 2004
		Differentiation	McBeath et al., 2004; Kilian et al., 2010; Tang et al., 2010
	Pore size and porosity	Paracrine	Li et al., 2019
		Proliferation	Mygind et al., 2007; Zhang et al., 2016; Zhao et al., 2021
		Differentiation	Mygind et al., 2007; Kasten et al., 2008; Matsiko et al., 2015
		Adhesion	Murphy et al., 2016; Zhang et al., 2016
Biochemical	Growth factor and derivatives	Migration	O'Brien et al., 2007
		Proliferation	Rodrigues et al., 2010; Sahoo et al., 2010; Khojasteh et al., 2016
		Differentiation	Simmons et al., 2004; Park et al., 2009; Rodrigues et al., 2010; Sahoo et al., 2010; Re'em et al., 2012; Khojasteh et al., 2016; Rao et al., 2020; Zhu et al., 2020
		Paracrine	Tamama et al., 2010
	Small bioactive molecules (nitric oxide)	Spreading	Fan et al., 2007; Lam and Segura, 2013
		Adhesion	Fan et al., 2007; Qu et al., 2010; Khojasteh et al., 2016
		Immunomodulation	Yao et al., 2015
		Paracrine	Berardi et al., 2011
		Differentiation	Berardi et al., 2011
	Small bioactive molecules (oxygen level)	Adhesion	Xing et al., 2013
		Proliferation	Xing et al., 2013
		Proliferation	Zhou et al., 2014
		Differentiation	Zhou et al., 2014; Tong et al., 2016
	Genetic regulators (cDNA)	Paracrine	Tong et al., 2016
		Differentiation	Huang et al., 2005; Meinel et al., 2006; Yau et al., 2007; Lu et al., 2013
		Proliferation	You and Nam, 2013
		Migration	You and Nam, 2013
Genetic regulators (siDNA)	Paracrine	Moon et al., 2014	
	Differentiation	Nagane et al., 2010; Jia et al., 2014	
	Differentiation	Bauer et al., 2017	
Combined strategy	Mechanical property and polypeptides	Differentiation	Ge et al., 2009
	Strain and biochemical extract factors	Differentiation	Yao et al., 2020
	3D microenvironment and biochemical extract factors	Differentiation	Yao et al., 2020

Figure 9: A table of different biophysical and biochemical cues, and the associated effects on cell behaviour. A collection of biophysical and biochemical cues result in a range of effects on MSCs including proliferation, differentiation and adhesion. Combining both types of cues, such as 3D microenvironments and with biochemical extract factors, can result in beneficial differentiation. Reprinted from [1] under Open Access [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) © 2021 Li, Liu, Zhang, Yao, Enhejirigala, Li, Song, Wang, Duan, Yuan, Fu and Huang.

Background

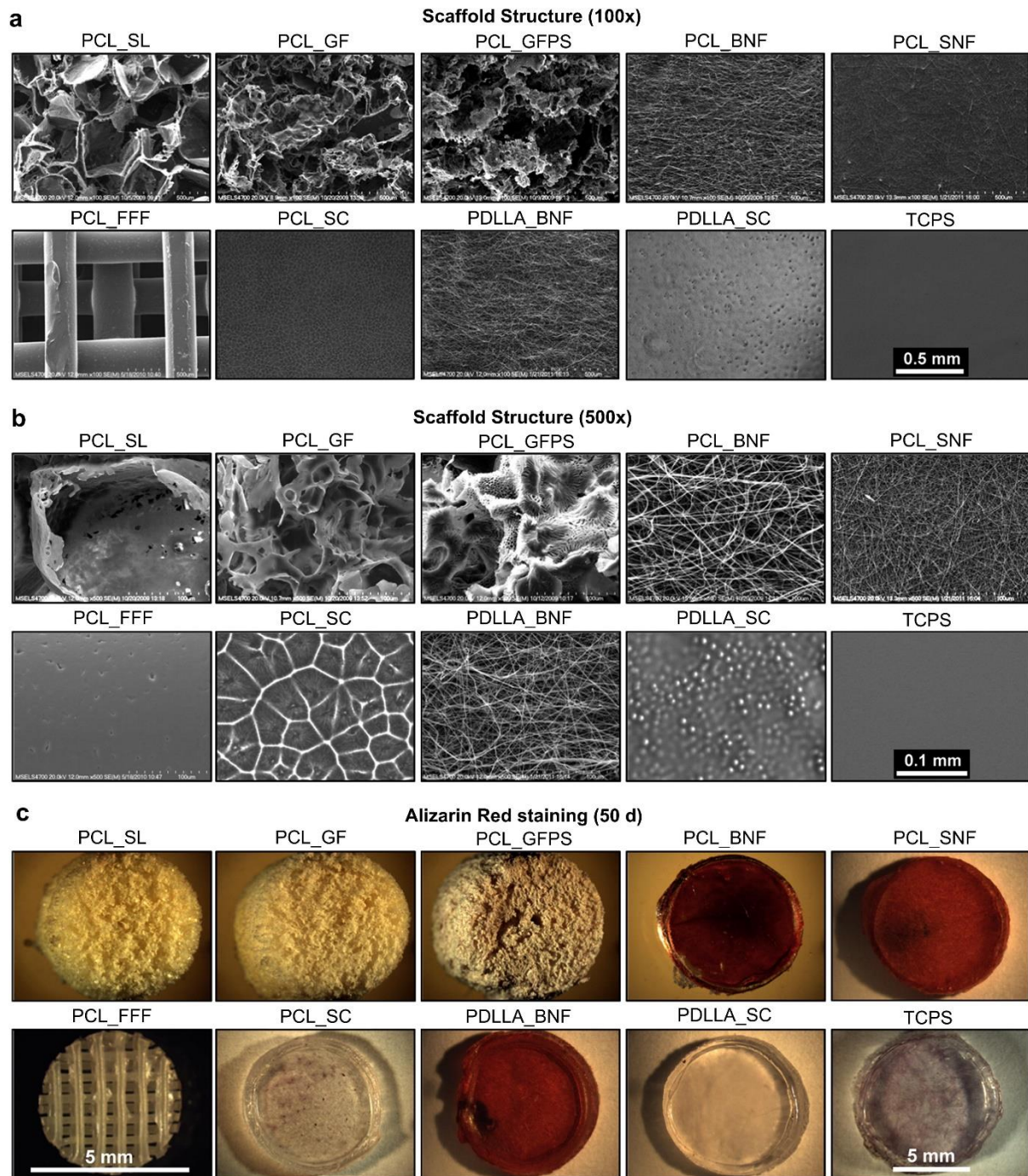


Figure 10: SEM images of scaffold structures and associated stereo micrographs of calcium staining (Alizarin red) for MSC osteogenesis.

(a) SEM images at 100x magnification and (b) 500x magnification of scaffold structures, with the scalebar on TCPS valid for all images in that section. (c) Stereo micrographs at 18x magnification, except PCL_FFF at 32 x magnification, of calcium staining (Alizarin red), which indicates via red staining the scaffolds where there is MSC osteogenesis due to the presence of calcium. Reprinted with permission from [15] © 2021 Elsevier.

2.4.2 CAM ASSAYS

Transitioning from cell culture-based testing to clinical application, it can be necessary to first utilise animal experimentation. Animal experimentation, such as CAM assays (Figure 11), is an essential component in the progression advancing promising biomaterials, because it is difficult to predict the outcome of *in vivo* experiments based on *in vitro* data. AI is a potential tool for predicting the outcome without the need for extensive animal experimentation [157].

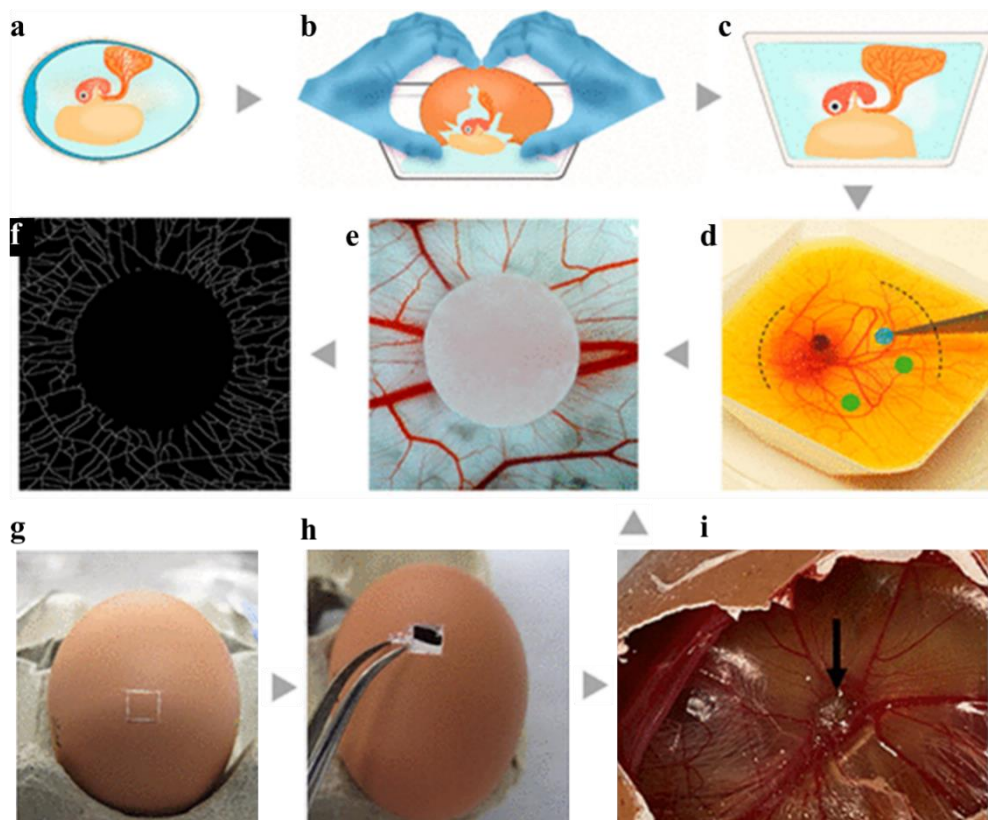


Figure 11: Two different methods of performing CAM assays resulting in skeletonised image for vessel analysis. (a-d) the process of *ex ovo* (shell-less) CAM assay preparation. (a) an illustration of an egg, (b) cracking the egg shell, (c) starting the *ex ovo* culture, and (d) implantation of different scaffolds, shows as green and blue circles. (g-i) the process of *in ovo* CAM assay preparation with (g, h) a window made in the egg for scaffold implantation and (i) a collagen sponge (black arrow) with an angiogenic response. (e) the blood vessels are imaged (clearly visible after a white contrast is injected). (f) the blood vessels are isolated and skeletonised computationally for analysis. (a-f) adapted with permission from [14]. © 2019 American Chemical Society. (g-i) adapted from [17] under Open Access [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) © Marshall KM, Kanczler JM, Oreffo RO.

Once optimal compositions have been identified, perhaps through AI using methodology introduced in chapter 6, they must be tested thoroughly before application in clinical settings. The CAM assay, an *ex vivo* preliminary model, is frequently used to determine biocompatibility and angiogenic/anti-angiogenic effects of biomaterials. These are applied to the vascular CAM, which has also been used to grow tissues and tumours for applications such as oncology research. Reviewed over the years, including recently by Marshall et al

Background

[17], there is increasing interest in the application for bone tissue engineering to screen materials prior to *in vivo* studies. The angiogenic response of the CAM can be quantified using the Chalkley Scoring method, with dots on a microscope eyepiece being positioned over an observed sample and the dots counted when they overlie a blood vessel. Due to the variability and subjectivity of this method, calculations using computer modelling, which removes the need for a Chalkley graticule eyepiece for the microscope, have been investigated to allow calculation of a microvessel area estimate for solid tumour samples [158]. A photograph of CAM blood vessels can be taken and manipulated using computer software programs to delineate the blood vessels, allowing for easier automated counting, as shown in a recent publication by Manger et al [159].

2.5 IMAGING A PLACENTA IN 3D

This project contains multiple areas of regenerative medicine and, in addition to tissue engineering, includes 3D structures in placental tissue. Manually labelling images of placenta in 3D is a time-intensive task. However, labelling is required to develop 3D models of structures within placenta. As previously mentioned, automation of cell labelling would not only decrease the time spent labelling 3D images, but also show that computational methods have the potential to distinguish between cells and acellular material, and between cell types, in complex images.

2.5.1 PLACENTA

Forming the interface with the mother, the placenta is a fetal organ which feeds the fetus, cleans its blood and, to support the pregnancy, secretes hormones which adapt the mother's physiology (Figure 12) [39]. Therefore, poor placental function can compromise fetal development and placental function directly determines fetal growth. As fetal development is the foundation for future growth, poor fetal growth is a predisposition to perinatal and postnatal diseases, and is important for health in later life – poor fetal growth is linked with higher chronic disease rates irrespective of age [160, 161].

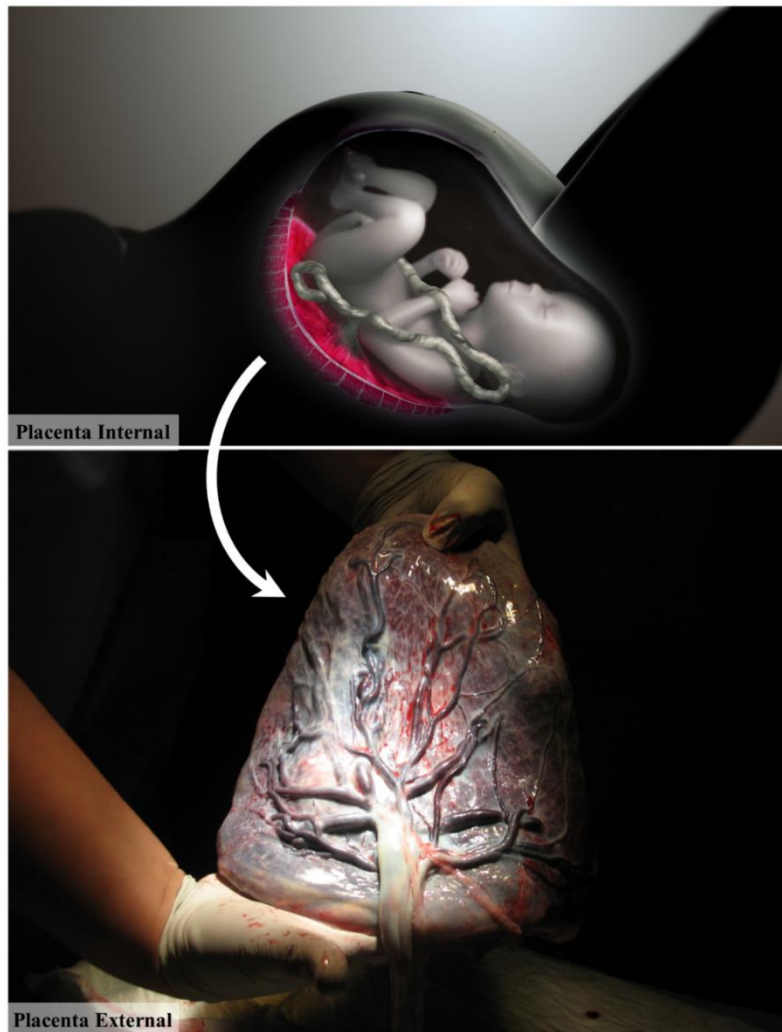


Figure 12: An illustration of the placenta connected to a fetus and a photograph of a placenta postnatal outside of the womb. The internal placenta illustration shows how the placenta connects the fetus to the mother, through an intricate network of blood vessels. The external placenta shows the size of the placenta and the scale of the vascular network connecting the umbilical cord to the placenta. A white arrow connects the illustrated placenta to the photograph of a real placenta. Both images are reused under Creative Commons licenses. © Scientific Animations Inc. [CC BY-SA 4.0](#). © Greger Ravik "Tree of Life" [CC BY 2.0](#).

There are many areas involved in investigating placental function, including understanding of its structure, physiology and biochemistry. Growth restricted fetuses have comparatively abnormal placentas, with alterations to structure observed, and reduced transport capacity (impacting availability of oxygen etc.). The relationship between the changes in structure and reduction of transport capacity is still undefined, and yet advances in 3D imaging provide insight into how structure determined function. To date, this work has been limited by scarcity of good tools for 3D study and, although overall structure of the placenta is relatively well understood, cellular anatomy is less so, especially at the ultrastructural level. For comprehensive understanding of placental function, spatial relationships of all placental cell types must be investigated, such as fibroblasts, pericytes and endothelial cells [162, 163].

Background

2.5.2 FIBROBLASTS, PERICYTES AND ENDOTHELIAL CELLS

All cells chosen for labelling in this chapter are significant for understanding development of the fetus and the impact of the placenta on lifelong health. Typically, alongside mesenchymal and mesenchymal-derived cells within the placenta villous core stroma, fibroblasts are found in the ECM between trophoblasts and fetal vessels. They maintain the ECM, on which the trophoblast layer resides and provides the scaffold for blood vessel growth. The number and function of trophoblasts affects placental function and development throughout pregnancy, which directly impacts risk of chronic disease and general health throughout life [39, 164]. Another important cell type in placental function are pericytes, which are found wrapped around the endothelial cells. Endothelial cells are of critical importance, as they form the walls of blood vessels, regulating the exchanges between the bloodstream and surrounding tissue, and endothelial signalling organises both growth and development of connective tissue cells (the surrounding layers of the blood-vessel wall) [165]. Pericytes communicate with endothelial cells via physical and paracrine signalling (cell-to-neighbouring-cell communication), and are thought to control vasculature stability and permeability, alongside providing physical or endocrine support for endothelium and, in some tissues, mediate vasoconstriction (slowing or even blocking blood flow through constriction of blood vessels) [2, 166, 167]. Together, the quantity, density, positioning, and structure of these cells give valuable insight into multiple features, such as nutrient transfer or toxicity within the placenta. Once understanding of these cells and their interactions is more advanced, models of the placenta could be developed, allowing for increased medical advancement. 3D structure is vital and therefore 3D imaging needs to be used, rather than conventional 2D approaches to cell study through visual analysis.

2.5.3 3D IMAGING

3D imaging has the potential to obtain complex spatial information about the placental structures, including villous and vascular branching. Additionally, associations between structures can be expressed and analysed, with possibilities of determining structures that are not readily visualised from 2D imaging techniques, such as thin stellate processes of stromal fibroblasts embedded within surrounding features [168]. Studying the anatomy of the placenta has previously relied on such 2D images, which were supported by SEM and vascular casting data for limited 3D interpretation [161, 163], and underlying structures within the placenta were developed from 2D images using stereology [162]. However, these approaches provide limited information about placental structure, such as the intricate spatial

relationship between cells, or whether there is regional heterogeneity in structure or even protein localisation [160].

Both micro-computed tomography (micro-CT), where X-rays at multiple angles are used to construct a microscale resolution, and SBFSEM, which is capable of producing nanoscale resolution, have been used successfully as 3D imaging techniques for biological tissue [169]. SBFSEM imaging starts by embedding a contrast-enhanced tissue in plastic or resin and taking a standard (2D) SEM image, then a thin slice of tissue is removed from the top of the block using an ultramicrotome slicing, which creates a new surface (typically 50 nm below the previous surface, although slices can be as thin as 10 nm). This process is repeated, with another 2D image taken of the new surface before it is cut away to reveal another surface (Figure 13).

The final step computationally joins the multiple SEM images into a full SBFSEM stack, for full 3D visualisation and manipulation, where x-y resolution is limited by the SEM imaging set-up (such as scanning speed), and the z resolution is limited by the depth of the ultrathin slices removed at each step [170].

Through displaying the 3D structures of different cells, and the interaction between cell structures and their surrounding environment (including the positional interplay between cells), SBFSEM imaging has been used to identify novel structures and develop deeper understanding of cellular function within the human placenta [168, 171, 172]. While application of conventional image segmentation for 3D medical image data can be combined to more quickly extract important feature information, these approaches require intensive data processing or a distinctive contrast in features within the tissue sample, which can be difficult and sometimes impossible to obtain in SBFSEM images of tissue [173, 174].

Unlike with some cell types and structures, which can be computationally rendered into 3D relatively easily through optimised heavy metal staining, rendering most tissue samples into 3D from a SBFSEM stack is time-consuming, and therefore an expensive task [175]. For a desired structure, for example a single fibroblast cell, to be extracted from an SBFSEM stack without any loss of data and resolution, each SEM section needs to be manually labelled. With each 2D image, of which there can be thousands in a single SBFSEM stack, taking several minutes to an hour (depending on factors including the size and resolution of the image, computational set-up, and manual skill) to manually label, a SBFSEM stack can take a month, if not longer, of dedicated labelling to complete. Furthermore, the complexity of

Background

SBFSEM images means existing defined image processing algorithms are incapable of achieving accurate computer automation. However, deep learning is a rapidly advancing field and, consequently, may provide a way for more time-efficient labelling without the time requisites of manual extraction of a 3D feature from a SBFSEM stack.

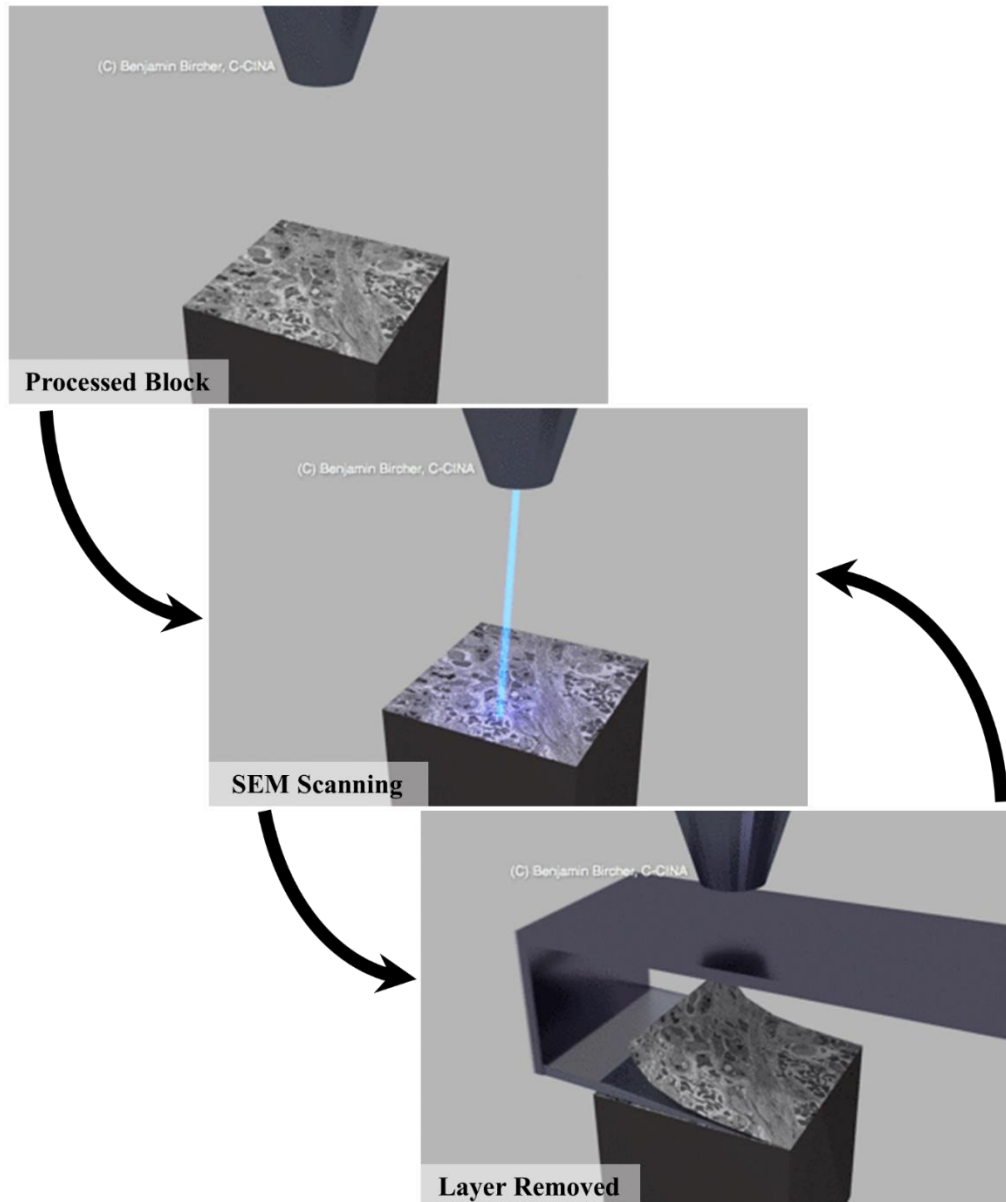


Figure 13: Process of Serial Block Face Scanning Electron Microscopy (SBFSEM) imaging.

A block of processed placenta is exposed to a SEM. This is then scanned in 2D as a standard SEM image. A diamond knife, or other blade capable of nanometre precision cutting, slices a thin layer of the block away, leaving a newly exposed image for SEM imaging. Through repeated scanning and cutting, the 3D block can be imaged in entirety. Images extracted from a GIF accessed from C-Cina.org, who retain full copyright [6]. © Benjamin Bircher, C-Cina & © DCI – Dubochet Center for Imaging.

2.6 SUMMARY

Regenerative medicine is a broad and important area of modern healthcare, and is becoming more important every year as more people live for longer across the world. Tissue

engineering, a subsection of regenerative medicine, focusses on using biomaterials and related bioengineering for tissue growth. This ranges from stem cells cultured to grow organs *in vitro* to cell-instructive biomaterial implants for regrowth of bone *in situ*.

Tissue engineering, while providing many pioneering treatments in multiple clinical settings, faces challenges such as the highly complex nature of cell response, and the dynamic multi-cellular environment of a healing body. More broadly, advances in regenerative medicine may be held back by the amount of time spent collecting or processing images for subsequent analysis.

3 DEEP LEARNING

AI-based computation methods have the potential to help researchers overcome some of the previously mentioned challenges within tissue engineering. AI models can provide statistically likely approximations without requiring complete datasets nor full scientific understanding.

The computational power necessary for deep learning, a concept within the field of AI where models have multiple layers, can be found on a standard home computer for the simplest models. However, the multidimensional matrix multiplications rapidly increase in complexity as the size of the network increases. For fully-sized network models, such as U-Nets and W-Nets, the processes become difficult to reverse engineer. However, this “black box” method of data transformation can provide, in theory, images that are both statistically likely and realistic-appearing, such that scientific understanding can still be gained from the resulting data.

Applying deep learning to medical imaging could be used to make time-intensive tasks more efficient. Examples on where AI is being applied to tissue engineering, and where it may advance the field in the future, can be found in the review “The future of bone regeneration: integrating AI into tissue engineering” Mackay *et al* [40].

The basics of deep learning are mathematical, as the transformation of data via a non-linear mathematical process within a deep learning model can result in the model acting as an approximator. A simplified mathematical introduction to deep feedforward networks is covered in section 3.1, while a more conceptual approach to deep learning is explored in section 3.2 for readers less comfortable with the standard mathematical processes covered in the former section. The architecture used throughout this thesis is briefly explained in section 3.3, and a training example is provided in section 3.4.

3.1 INTRODUCTORY MATHEMATICS OF DEEP NEURAL NETWORKS

Taking a more standard mathematical approach, the net input to any node not in the input layer, z , can be calculated as:

$$z = \sum_i w_i * x_i + b \tag{3.1}$$

Where w_i is the weight given to the input x_i . There is a total of i inputs.

The activation function φ is applied to compute a , which is the output of the same node.

$$a = \varphi(z) \tag{3.2}$$

Activation functions allow a neuron in the network to propagate learning data only when sufficiently excited. If the activation functions are all linear, the neural network is equivalent to a simple linear regression model. Therefore, non-linear activation functions are typical, such as a leaky ReLU $\varphi(x) = x1_{x \geq 0} + \alpha x 1_{x < 0}$. Other common functions include ReLU, sigmoid, and tanh.

Learning for a deep neural network (DNN) is the step of calculating the weights to produce the best approximation \hat{y}_i of the real value y_i starting from the input x_i . For this calculation, a loss function J is defined that quantifies the difference between the real and predicted values through the neural network \hat{y}_i^θ for model parameters θ on the overall training set with size m .

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \mathcal{L}(\hat{y}_i^\theta, y_i) \tag{3.3}$$

Minimising J requires two steps. First, forward propagation is used to propagate the training data through the network, and the loss function is the sum of errors in the approximations. Second, backpropagation results in calculation of the gradients of the cost function \mathcal{L} with respect to the different network parameters, to which a gradient descent algorithm can be applied for updated weight values:

$$\theta = G(\theta) \tag{3.4}$$

This process is repeated for a defined number of epochs, where a single epoch has been completed when all of the training data has been propagated through the network.

A typical DNN consists of multiple neurons with multiple layers, l . Visualising that nodes form rows in each layer, the j^{th} node of the i^{th} layer would give the following equations:

$$z_j^{[i]} = \sum_{l=1}^{n_{i-1}} w_{j,l}^{[i]} a_l^{[i-1]} + b_j^{[i]} \rightarrow a_j^{[i]} = \varphi^i(z_j^{[i]}) \tag{3.5}$$

With n_{i-1} being the number of neurons in the layer $(i - 1)$.

Across the entire network where $\mathbf{W}^{[i]}$ is all the weights w in the layer i , $\mathbf{Z}^{[i]}$ is all the node inputs z in the layer i transposed ($\mathbf{Z}^{[i]} = {}^T [z_1^{[i]}, z_2^{[i]}, \dots, z_{n_i}^{[i]}]$), $\mathbf{A}^{[i]}$ is all the outputs a in the

layer i transposed, $\boldsymbol{\psi}^{[i]}$ is all the activation functions φ in the layer i , and $\mathbf{b}^{[i]}$ is all the bias weights b in the layer i transposed:

$$\mathbf{A}^{[i]} = \boldsymbol{\psi}^{[i]}(\mathbf{Z}^{[i]}) = \boldsymbol{\psi}^{[i]}(\mathbf{W}^{[i]T} \mathbf{A}^{[i-1]} + \mathbf{b}^{[i]}) \quad (3.6)$$

The optimisation of J is typically through a descent method, and many rely on gradient descent, which requires calculating the gradient $\nabla_{\theta} J(\theta)$. This can be carried out using a computation graph which decomposes the function J into several intermediate variables. Mathematically, the gradients of J with respect to the architecture's parameters $\mathbf{W}^{[i]}$ and $\mathbf{b}^{[i]}$ are calculated using the following:

$$d\alpha = \frac{\partial J}{\partial \alpha^i} \quad (3.7)$$

Where α is any given parameter. At the i^{th} layer:

$$d\mathbf{Z}^{[i]} = d\mathbf{A}^{[i]} \circ \boldsymbol{\psi}^{[i]}(\mathbf{Z}^{[i]}) \quad (3.8)$$

$$d\mathbf{A}^{[i-1]} = \mathbf{W}^{[i]T} d\mathbf{Z}^{[i]} \quad (3.9)$$

$$d\mathbf{W}^{[i]} = d\mathbf{Z}^{[i]} \mathbf{A}^{[i-1]} \quad (3.10)$$

$$d\mathbf{b}^{[i]} = d\mathbf{Z}^{[i]} \quad (3.11)$$

Where \circ is the Hadamard product operator. These equations are then recursively applied for $i = L, L - 1, L - 2, \dots, 1$.

Before optimising the loss function, the inputs are typically normalised, the inputs can be normalised to speed up learning. J becomes more symmetric and smaller in scope, helping gradient descent to find the minimum faster and therefore requiring fewer iterations. For a convex and differentiable J , the global minimum is equivalent to solving $\nabla_{\theta} J(\theta) = 0$, with the solution of θ^* . A standard method is to start with an initial guess θ_0 and form an iterative approach:

$$\theta_{k+1} = \theta_k + \alpha_k d_k \quad (3.12)$$

$$\rightarrow J(\theta_{k+1}) = J(\theta_k) + \alpha_k \nabla J(\theta_k) d_k + o(\theta_k) \quad (3.13)$$

Where α_k is the step size and d_k is the descent direction. Since the aim is to minimise J , $\nabla J(\theta_k) d_k$ needs to be negative and therefore $d_k = -\nabla J(\theta_k)$. α_k can be fixed or variable following a specified decay law. Learning rate decay results in a reduced learning rate over time/iteration, so that bigger steps can be taken at the beginning but smaller steps are taken as the global minimum is approached [176-178].

Hypothetically, the neural network could be trained until $\nabla_{\theta}J(\theta) = 0$ and $J = 0$. However, in practice, neither of these stopping criteria are likely to be met or desirable. When J_{train} , the training loss, and J_{test} , the testing loss, are both high, this can be a sign of underfitting, where the network has failed to find the path in the data and $\nabla_{\theta}J(\theta) \neq 0$. When $J_{train} \rightarrow 0$ but J_{test} is relatively high, this can be a sign of overfitting, where the model fits the training data perfectly ($\nabla_{\theta}J(\theta) = 0$) but fails to generalise on unseen data.

3.2 CONCEPT OF DEEP LEARNING

This section covers a generalised overview of deep learning with minimal mathematical concepts. As data progresses through multiple layers of a deep learning model, the non-linear transformations can result in statistically likely approximations [179, 180]. With an infinite number of transformations, such a model would hypothetically become a universal approximator [181]. When limited to several dozen transformations, it can still be highly versatile. The architecture capable of predicting laser-machined surface profiles in 3D [13] was also used to predict the skeletal stem cell response to topography in 2D [11], with both utilising a subsection of artificial neural networks called deep learning. While deep learning was used throughout the projects covered in this thesis, architectures were adapted for novel implications and not designed from scratch for each specific task. Therefore, the mathematics involved is not covered in the greater detail required for the comparison and design of multiple neural network architectures.

The multiple transformations of data as it progresses through a deep learning model are not easily tracked or reverse engineered, which is partially responsible for deep learning methods, and sometimes AI in general, being referred to as a “black box”. However, it is the inclusion of multiple transformations that is behind the frequent success of deep learning in modern applications, specifically artificial neural networks (ANNs) for complex tasks. ANNs, computing algorithms for transformation of data that mimic the way nerve cells work in the human brain, require no pre-programmed information, such as scientific concepts, for successful application. Without any understanding of the underlying properties behind a transformation, ANNs produce statistically likely solutions by forming an approximation [181, 182]. For comparison, a young child is able to correctly identify a cat from a different (previously unencountered) animal, through exposure to several cats beforehand, without requiring the detailed knowledge behind scientific classification of mammals. The child has successfully approximated the transformation of data, such as the shape of the ears and the size of the animal into the species of animal.

Conceptually, ANNs contain multiple similarities to biological neural networks (BNNs), including the human brain, as they are directly analogous to an animal visual cortex (AVCs) (Figure 14). ANNs and AVCs both comprise of an assembly of nodes (neurons) with interconnecting synapses. For a simple biological neuron in an AVC, dendrites propagate electrochemical stimulation received from other neurons, and these impulses are carried towards the cell body and the nucleus, which then leads to a new impulse travelling from the nucleus to the axon terminals, where an electrochemical signal is released. This signal can then be detected by a different biological neuron, along with the signals of other surrounding neurons. For a simple node in an ANN, data is input from one or more sources, given an individual weighting, and then the collective weighted data is summed and passed through an activation function. This transformed data is then output from the node, possibly to be input to another node in the system.

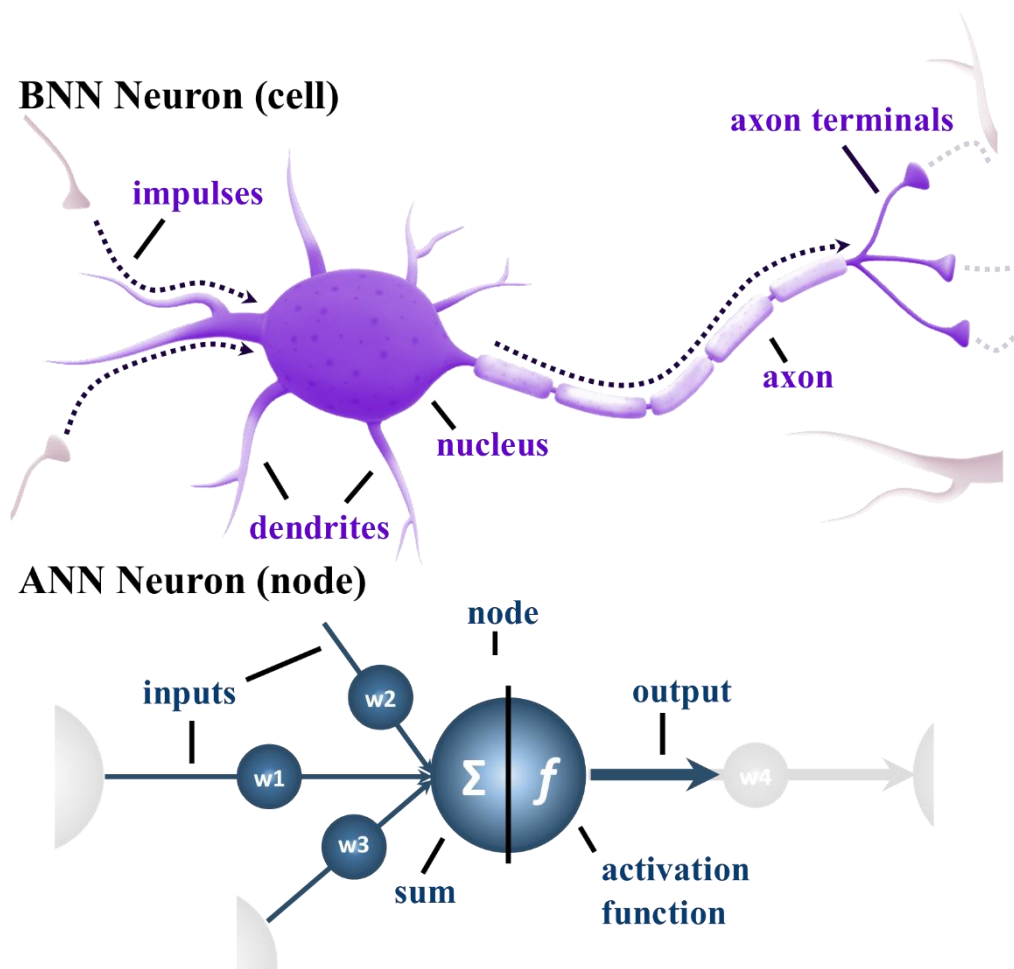


Figure 14: BNN (biological neural network) cells and ANN (artificial neural network) nodes have many similarities. A BNN cell comprises of dendrites, which transfer impulses to the nucleus. The nucleus processes impulses from multiple sources, resulting in further impulses being transferred from the nucleus along the axon to the axon terminals. Here, the impulse signals are released to be detected by other cells. An ANN node comprises of a node, which sums inputs and outputs this sum after processing form an activation function.

The weights of the synapses can change as both a BNN and ANN “learns”, similar to how pathways between neurons strengthen or weaken as we learn. A successful method of altering the weights within an ANN trained via supervised learning is through the process of backpropagation, so that the transformed data output by the system becomes closer to the target data each time (and therefore error is minimised) [183]. As with animals, connections are facilitated through larger weightings if they produce the correct outputs, simultaneously weakening those that do not, analogous to brain plasticity [40]. The mathematics behind the data transformation can be varied, from choice of summation operator and activation function in a node impacting forward propagation of the data; choice of cost function for determining error; and choice of formulas for backpropagation and learning (gradient calculation and weight update formula). Some of the mathematical processes behind these choices, and examples of the impact they can have on training an ANN, are covered in the following section 3.2.1.

When data passes through multiple nodes, data transformation can happen at each stage within the ANN. Each stage of one or more stacked nodes, a layer, can result in a different levels of data transformation. DNNs are ANNs comprising of an input layer, output layer, and “hidden” layers in-between (Figure 15). The first hidden layer processes data directly from the input layer, but the second may only process data which has been transformed by the first. These types of feedforward DNNs, such as convolutional neural network (CNNs), are commonly used for image recognition and modification and will be the focus of this section. However, feedback networks (recurrent neural networks, RNNs), where signals move forward and backward through the network, and DNNs with more complex structures, such as where all inputs are not restricted to a single layer, can be more successful in areas such clinical outcome predictions with multiple input data sources [184]. For all of these DNNs, an increase in the number of hidden layers can result in a comparatively more abstract operation at each layer. Therefore, provided there are enough layers (and enough computing power to run those layers), DNNs require no input from human computer engineers or representation design to extract relevant features. Using backpropagation, the DNN learns what features are, and which of those features are relevant, with no prior knowledge.

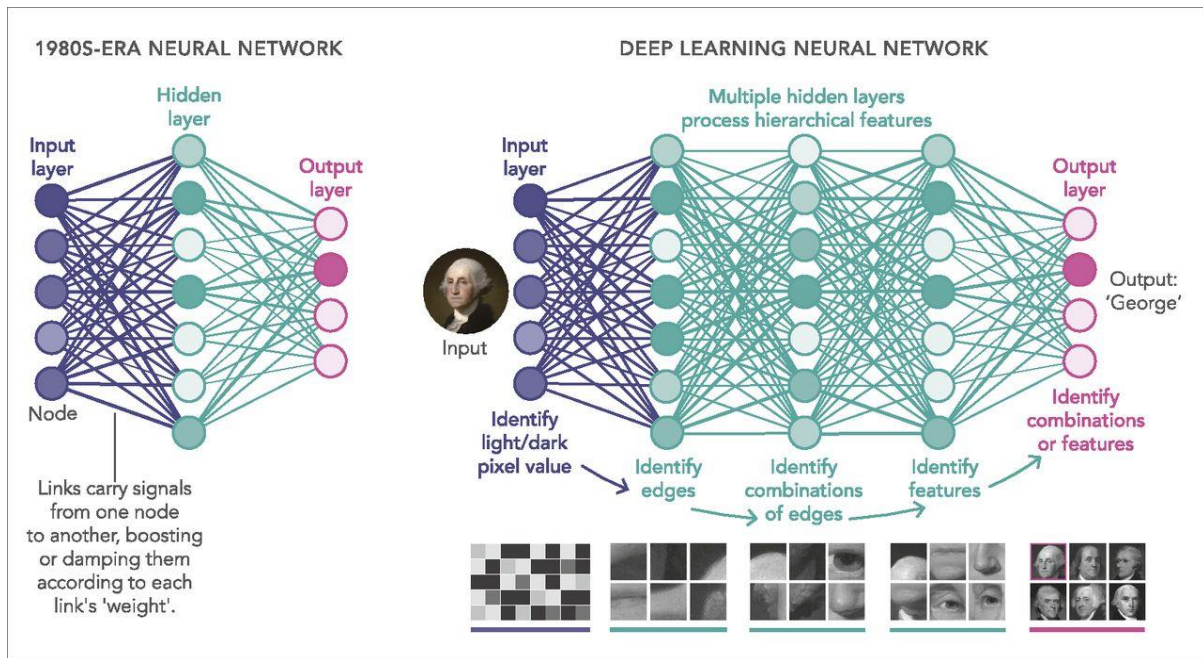


Figure 15: Comparison of a simple neural network and a deep neural network.

A DNN consists of an input layer, at least one hidden layer, and an output layer [7]. Here, the hidden layers consist of convolutional layers for spatial feature detection (identify edges). Different layers identify different features or feature combinations. Data is input as an image and transformed into a word output in this example, but any data transformation is theoretically possible with a large enough DNN. The right architecture is a fully connected network, depicting all the possible connections between nodes, but shares the concept of multiple hidden layers with a ConvNet.

The DNN in Figure 15 has only a few hidden layers, but a larger number is required for estimating the correct identification of historical individuals from varied input sources (application in a range of settings) in practice. Increasing the number by a magnitude, so that there are approximately 20 hidden layers, produces a DNN that can correctly extract key features in image identification, such as the ear shape and positioning, even with changes to background, lighting and magnification between images, and irrelevant features, such as furniture [179]. CNN are particularly successful DNNs for imaging tasks such as these, and have dominated the field of feature extraction from images for several decades [176, 185-193], including in medical imaging tasks [194-200]. At a conceptual level, a ConvNet contains similarities to how mammalian BNNs perceive visual information [201, 202], as key aspects of the architecture (the convolutional and pooling layers) are equivalent to the visual neuroscience concepts of simple and complex cells [179, 203-205].

Feature extraction within images can differ from other forms of feature detection, and can be more complex, due to the importance of spatial features, which are formed through the arrangement of pixels in multiple positions. In cell morphology identification from fluorescent images, the colour of each cell could be identical depending on the cell dyes used. Zoomed in so that only a single pixel of each cell type is compared, there would be no way to

distinguish between them. However, the shape of cells would vary, and therefore pixel values relevant to other surrounding pixel values must be detectable for successful identification. This is not limited to fluorescent images, where pixel values are limited by the number of fluorescent dyes used. Returning to identification of historical figures, the name of a person cannot be accurately determined by the colour paint (singular pixel values) of the portrait face alone – the facial structure (spatial arrangement of pixel values) must also be analysed. A convolutional layer extracts features without spatial information being lost, due to the mathematics behind the data transformation, which is covered in section 3.2.1. A pooling layer reduces the amount of data, discarding surplus information and non-features, by merging similar features, and therefore creating invariance to irrelevant shifts and distortions in the data. In other words, the network could interpret that if a cell is imaged in multiple different positions, at multiple different angles, it is still the same cell type.

The DNNs used throughout this thesis, U-Nets and W-Nets (explained more in section 3.3), transform image data into a new form of image data. Half of this transformation involves the convolutional process previously explained to extract important features from the input. The other half involves generating new features that are important for the output, which can be a more complex process [40]. During the feature extraction section, small shifts in data, such as (for the previous example of a cell morphology identifier) the position of a cell or the brightness/contrast of the image, are insignificant, as this cell retains the same morphology no matter the position and set-up of the camera. However, the feature generation section involves an element of choice, such as where to position the cell and, more importantly, where to position important identifying features on the cell. If a mononuclear cell had an equal probability of containing a nucleus towards the left of the cell body or towards the right, the DNN could learn over time, in order to match the target data where it could be left or right, to generate an image of a cell with two nuclei or none at all. This would be a failure by the DNN to generate an image of a mononuclear cell, and increased time training and increased training data (common fixes to improve DNN performance) would not necessarily lead to better results.

A conditional generative adversarial network (cGAN), a specialist type of image-to-image DNN, utilises a U-Net or W-Net feature-generating DNN (a generator) alongside an additional ConvNet (a discriminator). The discriminator only has the task of feature extraction, and can be input either a target image (i.e. an image of a cell taken from a camera in a laboratory), or the output image from the generator (i.e. a “mononuclear” cell with an

additional nucleus). The better the discriminator is at identifying whether an input is either a target or output image, the worse the generator is performing, and vice versa. By including both discriminator error and generator error into the overall loss function, the cGAN is trained to produce images that not only contain relevant features, such as a cell with a nucleus, but are also realistic (there is only one nucleus instead of two). This two-network approach overcomes potential probability-based issues in using a single DNN for data generation (Figure 16).

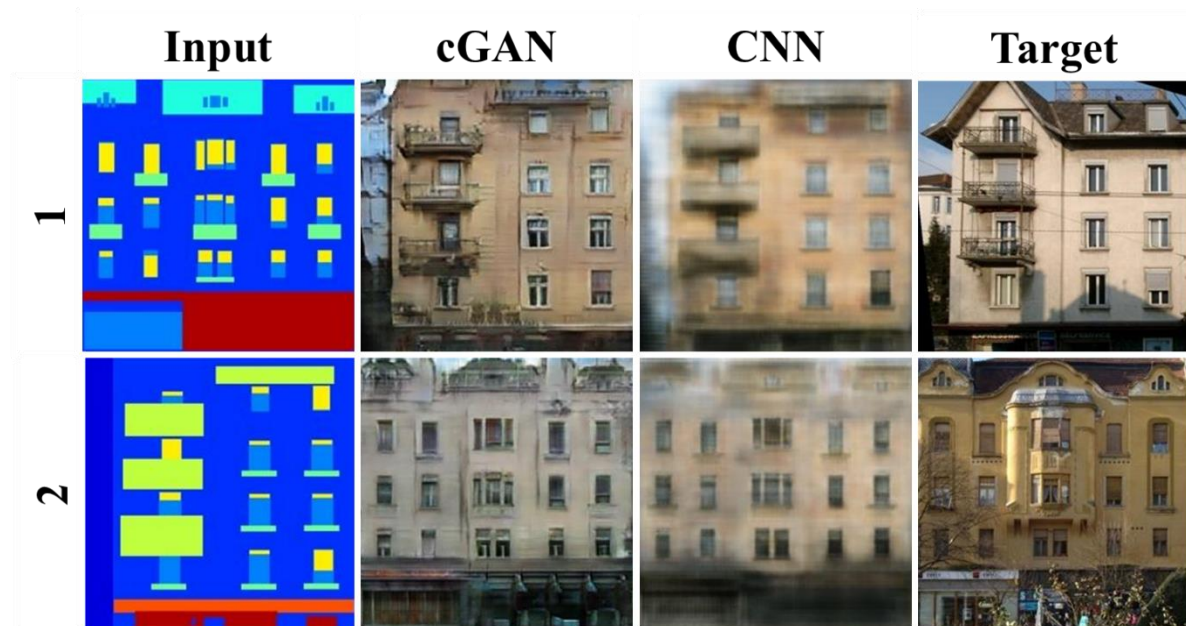


Figure 16: A U-Net is able to produce more realistic-appearing outputs than a CNN alone. Two different images are input into a cGAN utilising a U-Net architecture and a CNN network without a discriminator [5]. The target image is on the right and, while different from both cGAN and CNN outputs, the CNN produces images that are not as sharp/more blurred than both.

Choice of loss function can have major impact on training an ANN. In the previous scenario, the error was potentially greater for a cGAN output compared to a generator alone if the error is defined as the difference between output and target images, as the nucleus and cell position could be different from the target image in the cGAN (and not in multiple positions). L1 error (least absolute deviations) is typically used to quantify error by using the sum of the all the absolute differences between the outputs and target images. However, it was more important to have a clear, realistic-appearing, image than a minimised L1 error, so the total error also included the discriminator error (D error). A new loss function using both L1 and D may lead to successful training where either error independently would not. A badly chosen loss function can lead to unexpected consequences, such as when AI (not a DNN specifically) was tasked with winning multiple Nintendo™ Entertainment System (NES) games © 2018 Nintendo [206]. There are several ways “winning” can be interpreted, such as whether the AI

should be coded to achieve a high score, reach a high level, stay in play for as long as possible, or to simply avoid game-over. While being tasked to win resulted in successful, and bug exploiting, play for NES games such as Super Mario Bros., the same algorithm was unable to play Tetris. While both tasks involve winning at NES games using the same algorithm, the weighting to different elements of the loss function have to be finely tuned for each game. Similarly, the weighting between L1 and D loss in a cGAN loss function must be altered for each image-to-image task, even while the overall function remains the same. Backpropagation does not result in changes to these weightings, and they are instead included among the hyperparameters that stay constant while a DNN is training. Other hyperparameters include the number of epochs (how many times the network is input the complete training dataset), the number of hidden layers and how big those layers are, and the learning rate (how much the network updates as it learns). Too few epochs and a network may over-generalise but too many epochs can lead to overfitting, where the network learns the training dataset and cannot interpret additional input. Too few layers and the network may not be able to successfully approximate the data transformation, but too many layers that are too large may require excessive computational power and time to train. A learning rate that is too small will also lead to excessive time spent training, but a larger learning rate could never converge, where the network would never learn the correct approximation for data transformation. In practice, a decaying learning rate can help overcome the trade-offs in choosing learning rate, but fine-tuning is still necessary for successful training, as with other hyperparameters.

3.2.1 SIMPLIFIED CONCEPTUAL MATHEMATICS OF DEEP LEARNING

Each of the concepts covered so far, from activation functions to loss functions to hyperparameters, have a mathematical basis, which can include a range of higher level mathematics such as linear algebra and probability theory [180]. This section will cover only the basic mathematics behind the cGANs used for cell labelling and prediction without covering the mathematical proofs, reasonings and comparisons necessary for novel network architecture design, which is beyond the scope of this thesis. This section will use a hypothetical feedforward network comprising of one hidden layer and only five nodes and two bias nodes (Figure 17). A bias node is added to each layer besides the input layer, and connects to every node within that layer and no other layer. With a value of ± 1 for most networks, it impacts the activation function of the connecting node, with weights altered throughout training (the same as weights between nodes). Bias values are another

hyperparameter and do not change throughout training. All multidimensional variables in this section are converted to scalars for ease of conceptual understanding.

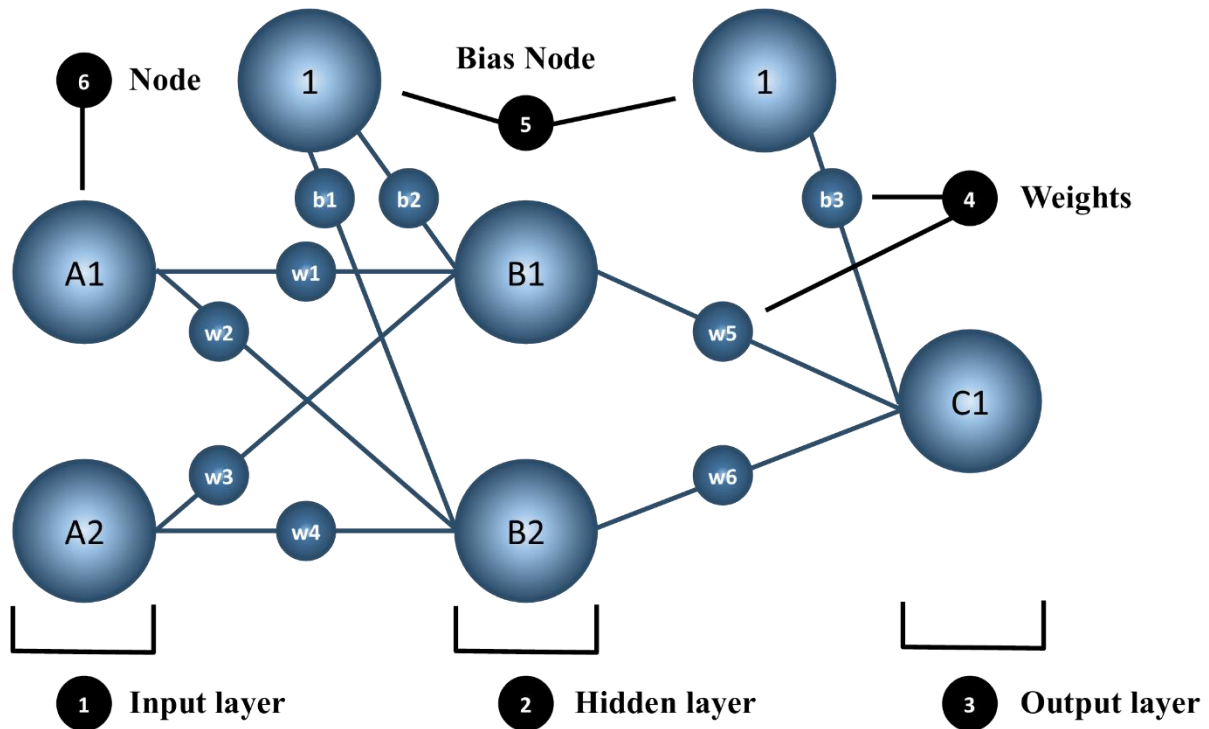


Figure 17: Simple NN architecture with a single hidden layer. A simple NN can have a single input, hidden and output node. Here, there are two input, two hidden and one output nodes, with two bias nodes. The hidden later is a fully connected layer.

Data “moves” from the input to the output through the process of forward propagation. As covered previously, every hidden and output node (Figure 17 B1, B2, and C1) contains a summation operator and an activation function. The summation operator sums all of the inputs for that node, forming a net input. The generalised net input would be calculated as follows:

$$net\ input = b + \sum_{i=1}^n x_i w_i \tag{3.14}$$

Where b represents the bias node input, n is the number of input nodes, x_i and w_i are the value and weighting from an individual node. For the node C1, the net input would therefore be:

$$net_{C1} = b_3 + (out_{B1} \cdot w_5 + out_{B2} \cdot w_6) \tag{3.15}$$

With values taken from Figure 17, where out_{B1} is the output of B1.

The activation function transforms the output value from the summation operator into a value representing how much the node should contribute (how much a node should fire). A

commonly used activation function is the sigmoid function, which would generate a value between 0 and 1.

$$S(x) = \frac{1}{1 + e^{-x}} \quad (3.16)$$

The sigmoid function, $S(x)$, also known as the logistic function, is a popular choice as the output between 0 and 1 can be beneficial to probabilistic applications (probability is always between 0 and 1). However, use of a sigmoid function does not always produce the best results, as different activation functions have their own strengths and weaknesses [207, 208]. A particularly successful activation function for improved accuracy over the sigmoid is the leaky ReLU. A ReLU is a linear function from 0 to infinity but has an output of zero for all negative input. The leaky ReLU has the following conditions:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases} \quad (3.17)$$

A leaky ReLU has been used for complex medical tasks [209], but the reasons for the advantage of a leaky ReLU over similar activation functions is still not fully understood [210]. Some networks use one type of activation function between hidden layers and another for the output layer, combining the high accuracy of i.e. a leaky ReLU, with a final output between 0 and 1 from, for example, a sigmoid function. Vector computation is also an important consideration, as is the differential of the function for backpropagation, which will be covered in more detail in the coming paragraphs. Although there are many available activation functions to choose from, they all introduce key characteristics to a network, including non-linearity (important for approximating complex data transformations) and limited output (large adjustments to weights can result in small adjustment to outputs, or vice versa, improving fine-tuning of the network during training).

Once data has travelled through the network, this transformed data is output from the network. For the network to train, the difference between the network output and the target data must be calculated, so that this difference, the total error, can be minimised, and the network can improve. The choice of loss function, or cost function/error function, will depend on the network architecture and the output data type, but also the task the network is undertaking. An example of a loss function mismatch would be if the training loss function is squared error in predicting total cell numbers, but the bioengineer only needs the colony to be above a specific density [40]. Squared error functions, and variations such as mean squared error, root mean square, and the sum of square errors (least square errors), are a common

choice for DNNs, and the least square errors, L2, is frequently applied to image processing in CNN [5, 191, 211, 212], but L2, along with L1 (the sum of the all the absolute differences between the true value and the predicted value) may lose their positions as de facto standard error metrics as newer functions reliably improve performance [213]. These loss functions can be written as:

$$L2 = \sum_{i=1}^m (t_i - z_i)^2 \quad (3.18)$$

$$L1 = \sum_{i=1}^m |t_i - z_i| \quad (3.19)$$

Where t_i and z_i are the target and network output respectively, and m is the total number of output nodes. The loss function for the whole network may utilise multiple error functions, with different modifiers attached to each depending on which should provide the largest impact.

Once the error has been calculated, the weights must be adjusted to minimise this error, which requires calculating the error of every weight in the network, and is the first stage of backpropagation. Through use of partial derivatives, changes in weights and changes in the network error can be analysed in relation to each other. This partial derivative is a gradient of the local error for a specific weight, and the minimum total error for the network, the global minimum, can be found through gradient descent, an optimisation method. The partial derivative of the total error, E , with respect to a specific weight, w_5 (Figure 17), can be calculated through the chain rule, resulting in:

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial Z} \frac{\partial Z}{\partial \text{net}_{C1}} \frac{\partial \text{net}_{C1}}{\partial w_5} \quad (3.20)$$

Where the total error changes as network output, Z , changes, which changes depending on the net input to the node C1, net_{C1} , which changes when w_5 changes. This long chain can be visualised from Figure 17.

The first derivative is easily calculated, as the relationship between the total error and network output has already been established depending on the choice of loss function. The second derivative is the relationship between the output and the net input to the node C1, which were connected through the chosen through the activation function. Similarly, the relationship between net_{C1} and w_5 has already been shown in eqn. 3.15. Assuming that the

network in Figure 17 had utilised a sigmoid activation function and a mean squared error loss function, the partial derivative would become:

$$\frac{\partial E}{\partial w_5} = (z - t) * z(1 - z) * \text{out}_{B1} = \delta_z \text{out}_{B1} \quad (3.21)$$

Eqn. 3.21 can be used to find the partial derivative of any output layer weight, with the variables w_5 and out_{B1} easily swapped with w_6 and out_{B2} . δ_z is the node delta, which is the derivative of C1's activation function multiplied by the difference between the network output and target output. Conveniently, the partial derivative of the total error with respect to a hidden node bias weight, b_3 , is simply δ_z , as bias nodes are not connected to any previous layer (so there is no out_{node} variable). The partial derivatives for hidden layer node weights are slightly more complex, as the chain rule once again works back from the network output.

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial \text{net}_{C1}} \frac{\partial \text{net}_{C1}}{\partial \text{out}_{B1}} \frac{\partial \text{out}_{B1}}{\partial \text{net}_{B1}} \frac{\partial \text{net}_{B1}}{\partial w_1} \quad (3.22)$$

Eqn. 3.20 is similar to eqn. 3.22, where the relationship between the total error and the weight into a node is linked to the input for that node, which in turn is linked to the resulting output for the node, which is then linked to the input to the following node (all easily traced through Figure 17). For a single output node, this becomes:

$$\frac{\partial E}{\partial w_1} = \delta_z * w_5 * \text{out}_{B1}(1 - \text{out}_{B1}) * \text{out}_{A1} = \delta_{B1} \text{out}_{A1} \quad (3.23)$$

With the above, the gradients of w_1 -6 can be calculated, and the next step is to alter them to reduce error. The equation for updating weights alters slightly depending on the gradient decent optimisation method used, so a generalised form will be used, here. new updated weight for w_5 , $w_{5_{\text{new}}}$, is calculated.

$$w_{5_{\text{new}}} = w_5 - \eta * \frac{\partial E}{\partial w_5} \quad (3.24)$$

The learning rate, η , impacts how much the weights change with each update. This equation shows how the contribution to error by an individual weight is subtracted from that weight, therefore reducing the total error in the system. If too much is subtracted, the global minimum (the point of lowest total error for the whole system) will be continually missed. If too little is subtracted, the network may become stuck at a local minimum, or take too long reaching the global minimum (Figure 18). Any of the weights in the system can be calculated in the same manner as w_5 .

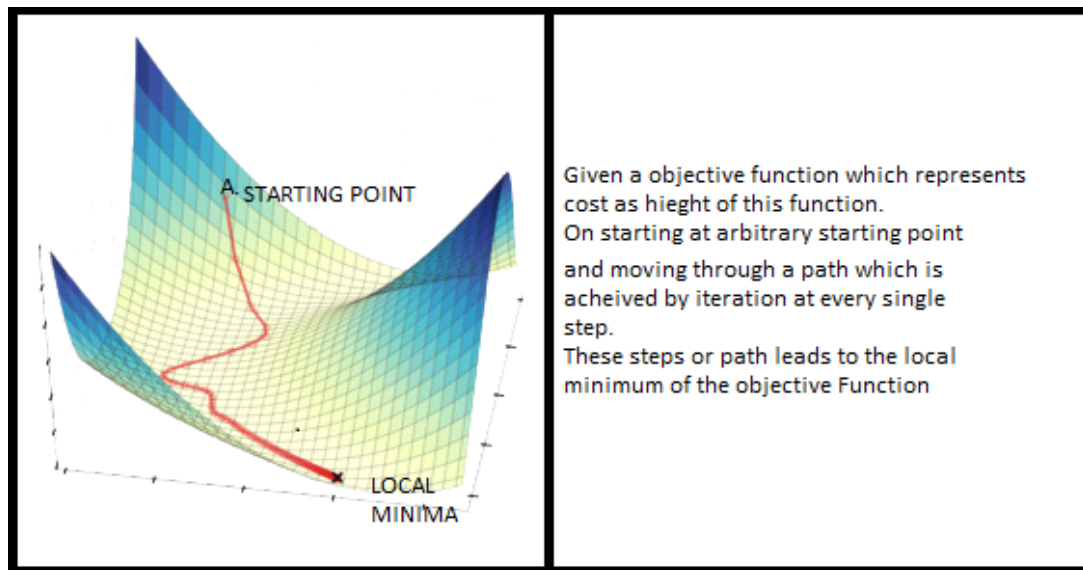


Figure 18: Gradient descent visualised with a 3D error landscape illustration.

Gradient descent can be visualised as a function for travelling down a multidimensional slope [4]. Here, the areas of large error gradient are in blue and low are in pale yellow. The local minima is the flattest part of the illustrated slope, but a true minima may be somewhere else (potentially not accessible).

Once the gradients are updated, the process begins again, with the total error reduced with each update. This was a very simple network, with a single hidden layer, no batch sizes (where multiple iterations happen together and weights are changed after a set number), no complex function choices, and no convolutional layers.

Convolutional layers reduce the image size without losing features that might otherwise be lost through simple averaging. The 2D convolution layer, conv2D, the commonly used within CNN, and is used for the U-Nets and W-Nets explained in section 3.3. A filter, also known as a kernel, in a conv2D layer acts on a small section of the 2D data (such as an image), stepping from left to right and top to bottom across the image, performing matrix multiplication. This results in a single output pixel for each step, due to the dot product. The step size, how many pixels the kernel moves across the image, is determined by the stride length, which is measured in pixels/data points. The kernel will perform the same operation for every location it steps through, transforming a 2D array of features into a smaller 2D matrix of features (Figure 19).

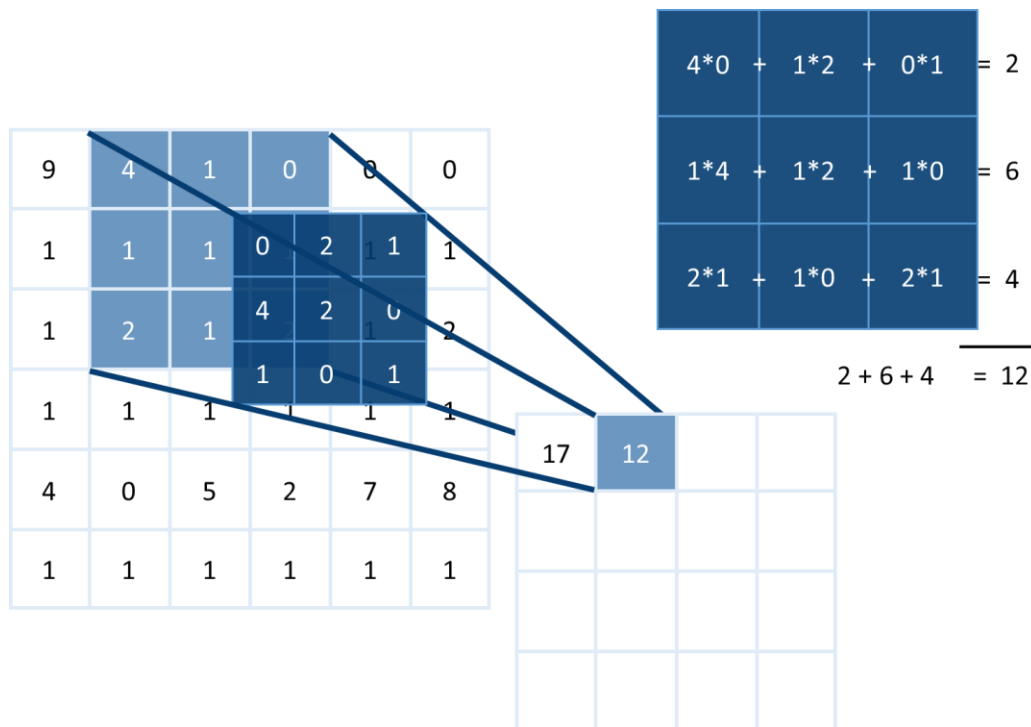


Figure 19: Visualisation of a matrix multiplication where a larger 2D array is transformed into a smaller 2D array via a filter.

While a kernel can be any size, it is typically a 3x3 matrix. This would in effect reduce a 6x6 matrix into a 4x4 matrix. The first step, for the output location [0][0] (the notation for the top left corner), would involve the calculation seen in the equation below. The next step is shown in Figure 19, where the dot product is expanded.

$$\begin{pmatrix} 9 & 4 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 & 1 \\ 4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = 17 \tag{3.25}$$

An example of an image going through a convolution layer, a ReLU later and a maximum pooling, MaxPool, layer is shown in Figure 20 [10]. MaxPool is a pooling operation that selects the maximum element from the specified region of the feature matrix, so the output results in a feature map of the most prominent features. Average pooling generates a mean average of the pixels within the specified area, and is also widely used. A MaxPool operation with a 2x2 filter and a stride of 2x2 is applied to a 2D matrix.

$$\begin{pmatrix} 14 & 5 & 2 & 7 \\ 9 & 23 & 5 & 0 \\ 12 & 13 & 0 & 1 \\ 14 & 14 & 0 & 0 \end{pmatrix} \xrightarrow{\text{MaxPool}} \begin{pmatrix} 23 & 7 \\ 14 & 1 \end{pmatrix} \tag{3.26}$$

Within a DNN, tensors are used instead of matrices, as the data can contain many dimensions, but these same operators can be applied.

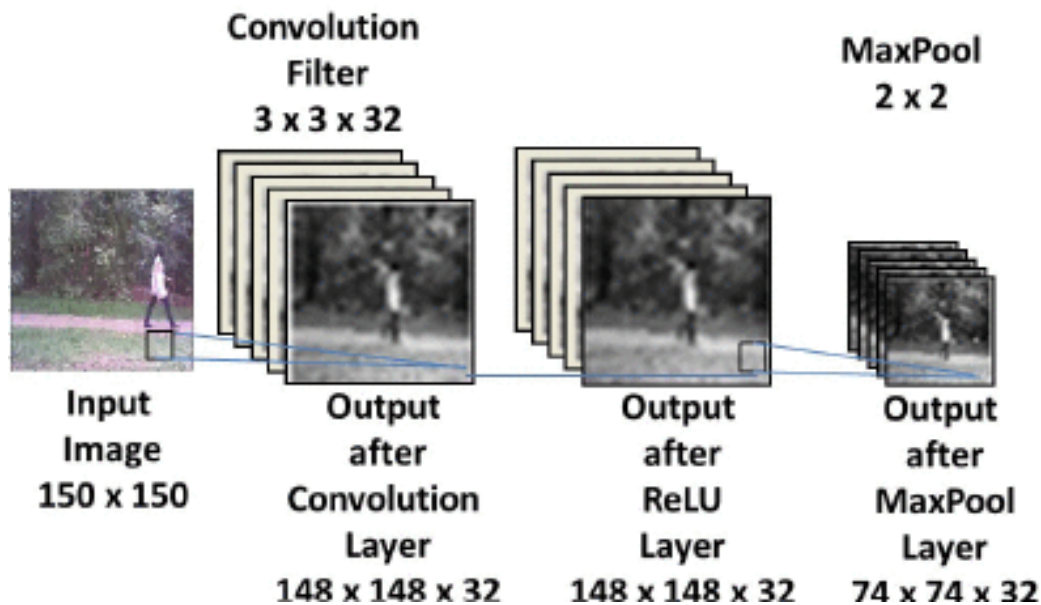


Figure 20: ConvNets utilise both a convolution and pooling layer combination, with image data transformed at each layer. An input image is transformed by a convolution layer, which visually alters the data during the feature extraction process. The “unnecessary” data is then discarded during the pooling layer. Here, the man is still visibly walking on a path, but the position and colouring has changed. Choubisa et al. [10] © 2017 IEEE

3.3 U-NETS AND W-NETS

U-Net architecture was originally an advancement of simpler CNN, first used for Biomedical Image Segmentation [214-217], and improved on the Fully Convolutional Networks of the time [218]. The name stems from a visual interpretation of the architecture, and they are capable of producing accurate predictions with relatively few training images (a few hundred as opposed to tens of thousands). The large number of feature channels make it ideal for complex and high resolution situations, avoiding many data- and calculation-heavy fully connected layers. The U-Net consists of two paths, a contracting path of convolutions and max pooling (where the spatial information is reduced as feature information is increased) and an expanding path of deconvolutions (also known as up-convolutions) and concatenations, joining sections together, combining spatial information with high-resolution features from the contracting path. A W-Net doubles up this architecture to allow for further potential feature extraction without increasing filter size beyond the capabilities of available computer hardware (Figure 21). The inclusion of a secondary U-Net is based on deep cascade learning theory [219], and can improve network performance without necessarily resulting in overfitting [220].

The U-Net and W-Net architecture makes them highly suitable for image-to-image data transformation, and are the generator in many cGANs, including the original [5]. The inclusion of an additional CNN as a discriminator is therefore included often enough in the architecture, alongside U-Nets, that the overall cGAN is frequently referred to as a U-Net. To avoid confusion, all references to U-Net and W-Nets in this thesis should be considered part of a larger cGAN architecture unless stated otherwise. (This is because the term cGAN alone could refer to either architecture.) The U-Net architecture has been used for a range of tasks, including laser machining and biomedical imaging, which makes it an ideal candidate for tissue engineering applications, which can combine both medical image data and engineering processes in cutting-edge research.

3.4 TRAINING A NEURAL NETWORK: EXAMPLE

To determine how accurately a cGAN can predict cell behaviour over time, a simpler temporal predictor is designed to operate without the introduction of topography. It is designed to eliminate cell density as a variable by inputting the exact cell number and positions in an image format – brightfield transmitted light microscope images. This can now be used to demonstrate how a neural network is trained.

Predicting how cells will behave and interact even within a short time period of a couple of hours is challenging and therefore, as with any complex problem with a high level of perceived randomisation, a large dataset is required. A standard time-lapse can only run continuously for two or three days before the time outside of ideal atmospheric and temperature conditions is overly detrimental to the cells, even without the additional risk of infection over time. However, many hundred repeated time-lapse experiments would be a wasteful use of donated human bone marrow. To circumvent this issue, raw 512×624 images are converted into multiple smaller images. This can be done through randomised cropping to ensure a varied dataset and try to minimise overfitting (memorising the correct output for seen data rather than forming a broader and more accurate prediction) – a risk when augmenting data in this fashion.

The movement of cells over a 24 hour period provides too many possible future scenarios for the network to accurately predict correct cell behaviour. This is circumvented by making composite input images. Each image is translated into a single channel greyscale image, then combined with an image taken, i.e. half an hour before and half an hour after, creating a three channel image that includes both spatial and temporal data, as seen in A and D of Figure 22.

Composite images drastically reduce the probable cell behaviour, and, with just half an hour separation, an ANN can then predict MSC behaviour 24 hours after the future-most channel.

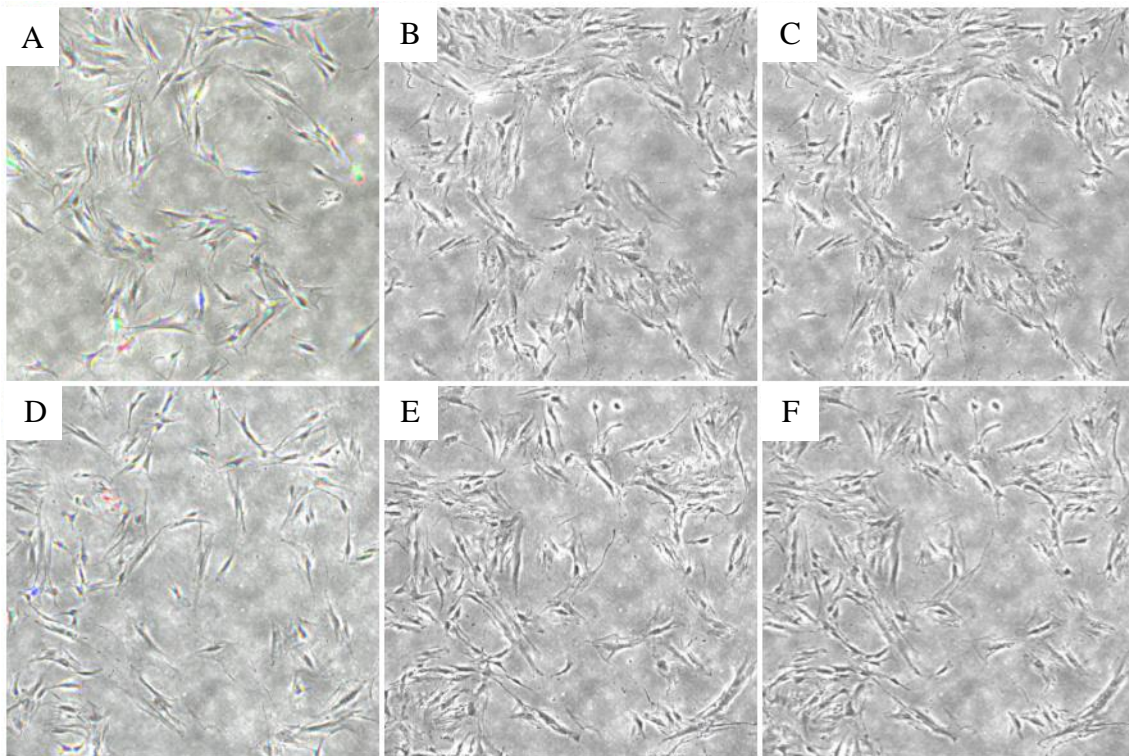


Figure 22: A combination of testing images for the 24 hour cell prediction AI. Images A-C are 256×256 images of *Stro1+* isolated SSCs and D-F are 512×512 images of the same cells. A and D are the input images, with the three channels giving additional temporal data. B and E are the predicted output of the model and C and F are the expected output.

Image A shows a composite input image of roughly 60 stem cells imaged using a brightfield microscope and cultured in standard cell conditions, consisting of three colour channels. The red channel is 25 hours, the green channel is 24.5 hours and the blue channel is 24 hours from the predicted output, which is image B. This predicted output is then compared with the expected output, image C. The AI has never seen either image A or C before, as this data is extracted from a random central point in the training data. Comparing these images by eye demonstrates that movement and proliferation (increased density of cells and their subsequent position) are robustly predicted. In addition, the AI disregards/learns to ignore air bubbles and scratches/imperfections in the images unrelated to cells (correctly unchanged across both images). Thus, the AI produces a prediction where only 9% of pixels at each position are different when directly compared with the expected output for 256×256 images.

A limitation to this model is that the input images have to be scaled to 256×256 and the output is limited to the same resolution. Therefore, work has been done on scaling this model to utilise 512×512 images and the additional data they contain. This higher resolution will be

Deep Learning

needed for accurate prediction on varied topographies. Seen on D-F, initial results are promising, the cell density and positions appear correct at a glance, but further work is needed on perfecting the network parameters, so results are comparable in error to the 256×256 model.

Optimising a network for a specific problem involves numerous iterations: changing hyperparameters such as batch size and learning rate or the architecture - layers and filter numbers. The cell predictor went through shorter training runs of only a few days before a reasonable configuration was acquired by the sixth version (V6), at which point the training runs were lengthened to a week for full analysis. V6 had a L1 to G error weight of 100. This meant that the pixel-to-pixel comparison of predicted output and expected output was multiplied by 100 in respect to the generator error in the overall error function of the AI, which the network works to minimise as it trains. There was a minimum of 128 filters per layer and the input images were 512×512 . Unlike previous versions, which were not successful for testing images, dropout was added to all deconvolution layers to help prevent overfitting of the model. These initial versions had a relatively large frame difference (FD) between the RGB channels of 50 minutes and small frame advance (FA), the time between the input blue channel and output image, of roughly two hours.

V7 was altered to have an L1 error weight of 200 – a great improvement on the previous version, as seen by the drop in L1 error in Figure 23. The next improvement was the most dramatic, v8, doubling the number of filters at each layer. Unfortunately, this meant reducing the image resolution to 128×128 on the memory-heavy parameters of previous models. V10 is the result of balancing filter numbers in different concentrations throughout the network, allowing for 256×256 images to run without issue. The FA was additionally altered to 8 hours. While the error is still substantially higher than for v8, part of this is due to the increase in pixel number, and subsequent increase in L1 error. V11 and v12 built on these adaptations and the FA could be increased to 24 hours with very little increase in error.

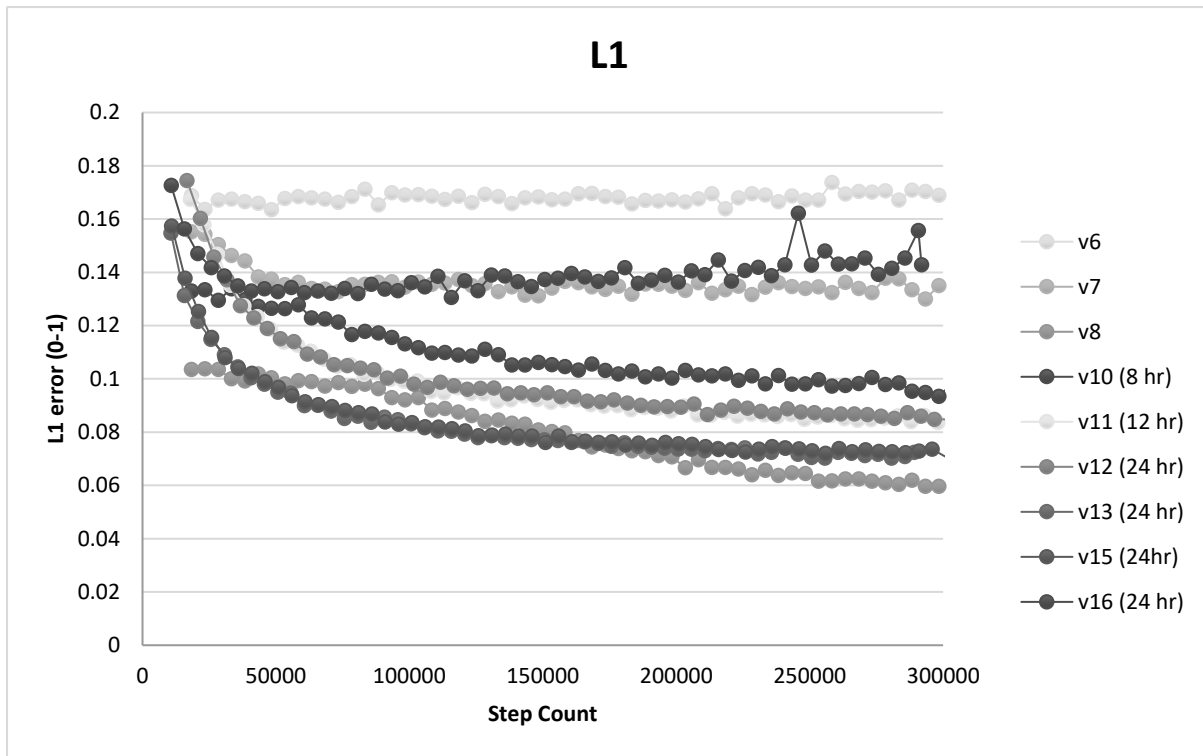


Figure 23: Graph of L1 error of versions 6-15 of the Cell Predictor cGAN as a percentage of incorrect pixels in training data vs the step count of training models.

Each version corresponds to altering hyperparameters of the network model.

Further investigation is needed as to whether there is a limit in difficulty for long term prediction, which would require training data taken over a longer period. V13 illustrates an improvement to the training data. Rare composite RGB images were a merge of different positions, harming the training process of the AI. V15 is a reduction in FD of the composite images to a much smaller 30 minutes. The L1 error remained stable considering the reduction in temporal data available, showing there is a likely optimum FD, which would balance availability of temporal data with the amount of time cells need to be imaged before accurate prediction can commence. The final version here, v16, illustrates the difficulty in scaling images to 512×512 . The necessary decrease in filter numbers combined with increase in pixels resulted in a model unable to accurately predict cell behaviour as well as previous models.

It is important to note, that a particular limitation with this result is that the AI was trained on a previous section of the time lapse images and had therefore learnt the style and behaviour of these cells. There is limited variation in that the training and testing data is from the same patient, roughly the same cell density and, as it is from the same experiment, the same temperature, humidity, lighting, and oxygen levels, which all could impact how the cells behave. Future work should explore the application of AI to a variety of patients and

Deep Learning

experimental datasets, which may be harder to predict. Importantly, the success of this provisional AI in temporal prediction shows promise for AI prediction of cell behaviour.

4 DATA: COLLECTION AND COMPOSITION

The raw data for chapters 5 and 7 was provided for by collaborators. The data for chapter 6 was independently collected for this project. All data processing, augmentation, and subsequent usage and analysis detailed in the relevant chapters, was performed by the author unless specified otherwise.

4.1 PLACENTAL SBFSEM IMAGES – DATA FOR CHAPTER 5

The SBFSEM stacks used in this project were provided by Professor Rohan Lewis at the University of Southampton (Rohan.Lewis@soton.ac.uk). Three stacks were used in total, as shown in the table below,

Stack	Used number of Z-positions	Structures Labelled	Training	Testing
1	943	Fibroblast, pericyte	Yes	Yes
2	370	Endothelial	Yes	Yes
3	140	None	No	Yes

Table 1: The SBFSEM stacks provided by Professor Rohan Lewis.

Stack 3 contained no manual labelling and was therefore used for testing only. The training and testing split, which varies for different experiments, is detailed in chapter 5.

4.2 STEM CELL RESPONSE TO SURFACE TOPOGRAPHIES – DATA FOR CHAPTER 6

The data for cell response to laser-machined topographies was created specifically for this project and was not provided by collaborators. The steps involved in generating this data, including optimising the approaches disclosed herein, took approximately a year. The steps involved in generating data for training a DNN is shown in Figure 24. The application of this data will be explained further in chapter 6.

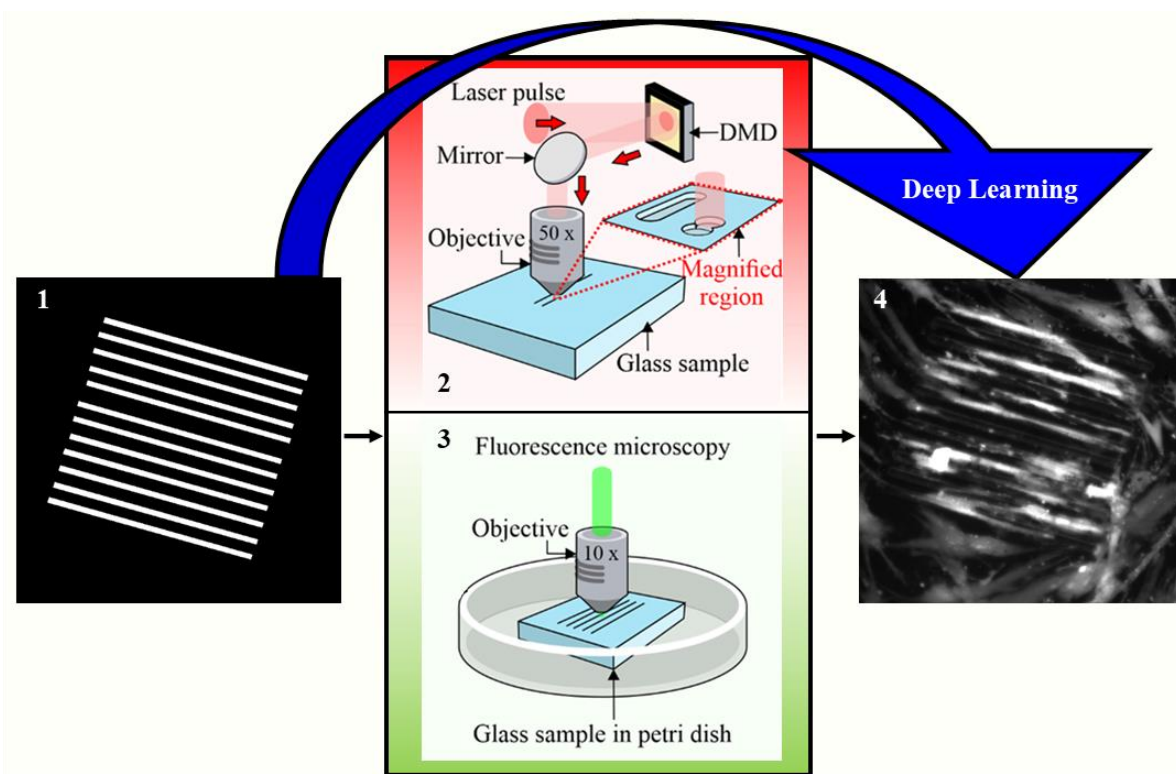


Figure 24: The path from a designed topography to associated cell response through experimental and deep learning methods. 1) A non-uniform topography is designed using a 2D computational binary array where white is areas of laser-machined grooves, and black is an area which is not machined. 2) The substrate is machined using a femtosecond pulse laser, using a beam shaper, DMD, and a 50 x objective. 3) Once cells have been cultured on the machined substrate, fluorescence microscopy is used to image the cells. 4) Fluorescent cell-tracker dye shows the cell positioning in response to the designed topography. Step 1 to 4 can be linked through deep learning.

The first stage was to design a pattern to laser-machine into the substrate sample. This pattern, (Figure 24 (1)), was shown as an array where black was areas not laser-machined and white was laser-machined areas. The next two stages involved the processes in collecting training data, including (2) femtosecond laser-machining of the desired pattern, a topography of controlled microscale features, and (3) the culture of Stro-1 positive isolated HBMSCs on the topography, which were then imaged using fluorescence microscopy. The final stage, (4), was to process the fluorescent images and then use them to train a DNN. Once the network was trained, deep learning bypassed the need for lengthy steps (2) and (3), shown by the blue arrow linking (1) to (4).

4.2.1 SURFACE TOPOGRAPHIES ON GLASS

As covered in section 2.3, laser-machining was achieved using an ultrafast pulse laser on glass coverslips, with the addition of a DMD to allow for a wider range of laser ablation patterning and future advancement of topography design (Figure 25). Overall, optimisation of laser-parameters can be time intensive as it requires extensive experimentation. However, once these parameters are obtained, laser-machining was an efficient and adaptable method

for topography generation. Therefore, to limit time spent on set-up, a single sample material was used for all topographies, which therefore only required a single set of parameters for all laser-machining within the project. As a known biocompatible substance, which is also transparent, aiding in necessary cell imaging, glass was a good option for a sample material. Specifically, glass coverslips (borosilicate glass) are designed for optimal imaging via both brightfield and fluorescence. Additionally, coverslips have very low coefficients of thermal expansion (increasing resistance to thermal shock, and use of femtosecond laser-machining further reduced likelihood of cracking due to limited thermal effects in ultrafast pulse laser ablation), so sterilisation of topographies from pressurised steam did not lead to fracturing, even after surface alteration. Importantly, cells were known to adhere to glass and femtosecond machining of glass, taking advantage of non-linear absorption, allows fabrication of features with high fidelity. Specifically, soda-lime glass coverslips were used throughout this chapter as these coverslips were less susceptible to fracture than pure silica glass post-machining.

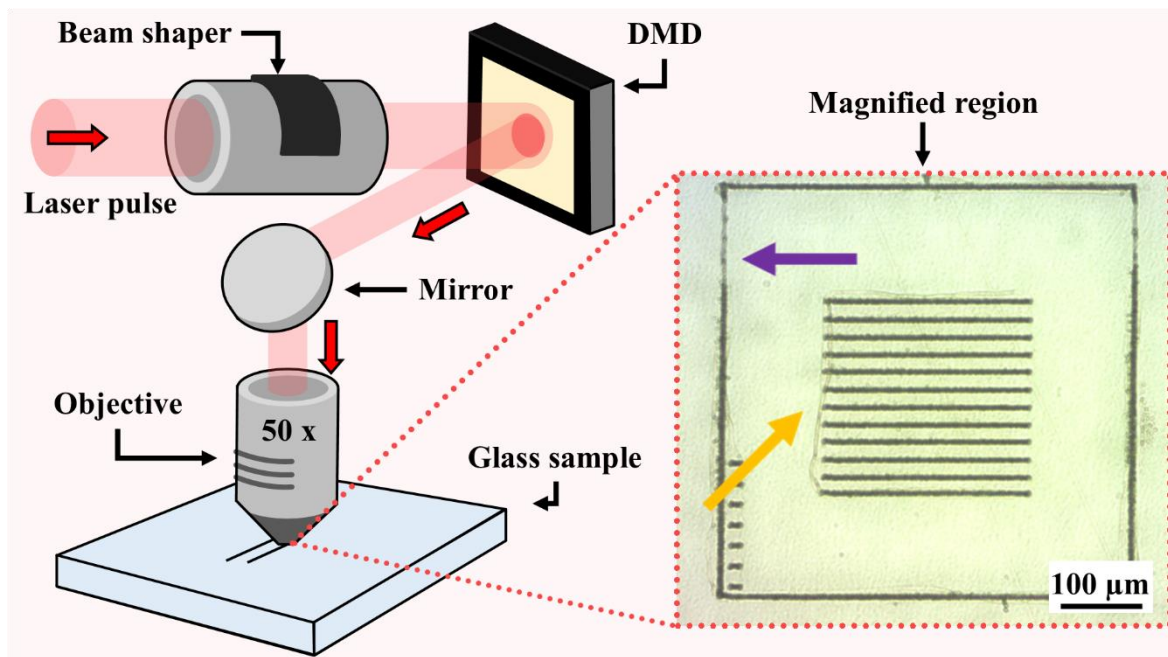


Figure 25: The laser-machining process for creating microtopographical glass substrates. A laser pulse from a femtosecond 1 W Ti:sapphire ultrafast pulse laser is shaped using a top-hat beam shaper, spatial intensity is then manipulated using a DMD, before a 50 x objective focusses the pulse onto a glass sample. This sample is moved using a motorized x-y-z stage, allowing for detailed microscale patterning to be machined.

The ultrafast pulse laser system was a 1 W Ti:sapphire, with a laser wavelength centred at 800 nm, pulse duration of 150 fs and operating at a 1 kHz repetition rate. To ensure homogenous illumination of all mirrors on the DMD, the spatial intensity profile was modified by a top hat beam shaper. Use of a DMD to control the spatial intensity of the laser pulses on the sample further controlled the topography of the ablated structures [221].

Data: Collection and Composition

Although any 2D shape could have been used, a circle was chosen to produce circularly shaped laser pulses on the surface of the substrate, so ablated areas had a uniform depth. When combined with substrate movement (via a 3-axis translation stage), the result was a continuous ablated line, with line thickness corresponding to the diameter of the projected circle shape and depth depending on the speed of the stages, and therefore the number of pulses per area.

The size of the circle pattern was optimised throughout machining to balance the creation of thin line widths (less than 10 μm) with necessary fluence for ablation, and line widths could be rapidly changed for topographical variation via the DMD. Made up of a 604 \times 684 array of square $\sim 7.6\mu\text{m}$ mirrors (DMD pixels), the spatial control over the beam shape would also allow for consistent machining over curved samples, such as bone samples and titanium implants, for future expansion and increased versatility of the network model and easy transference from *in vitro* topographical response to topographical biocompatibility and *in vivo* response [222]. With a circle pattern set on the DMD of radius between 75 and 100 pixels, and a reduced laser power of 400 mW via a variable density filter, resultant laser-machined lines had a maximum thickness range between 7.5 and 12 μm on the glass sample. The use of a variable density filter was to enable uniform machining, without gaps in lines caused by fluctuations in laser power resulting in an incidence fluence below the ablation threshold of the sample, without creating cracks and other sample deformations from high power machining. The purple arrow in Figure 25 shows an area fluctuating below the ablation threshold and the orange arrow shows a thin crack on a single sample with unchanged laser parameters. However, these discrepancies in laser-machining introduce a wider range of features in the training data, including the range of line-widths (7.5-12 μm), which can be beneficial for training neural networks. Therefore, these fluctuations were not adjusted for via, for example, CNN-monitoring or repeated machining. The laser, 3-axis translation stage, the DMD, and the variable density filter were all automated and run via an experiment-specific automation code.

When laser-machined glass was imaged using a scanning electron microscope (SEM), there was substantial variation between resultant nanotopographies, depending on the separation between laser-machined lines (Figure 26). When lines separated by less than a line-width are laser-machined, the surface became rough and nanotopographically non-uniform (nanorough) and contained no deeply ablated areas. After separation was increased to a level with no overlapping of laser-machining, deeper ablated lines became visible – the darker areas

correspond to relatively greater depth. At a separation large enough that there was no disturbance from other laser-machined areas, the lines were ablated with a rough edge, with no deformities to the otherwise smooth surface topography, and very limited build-up of debris and material between ablated depths. This was very promising, as laser-machining provided a way to precisely control microtopography for optimised design, while also producing varied nanotopographical features beneficial for cell adhesion due to differences in surrounding debris and material build-up, depending on proximity to other laser-machined areas [223].

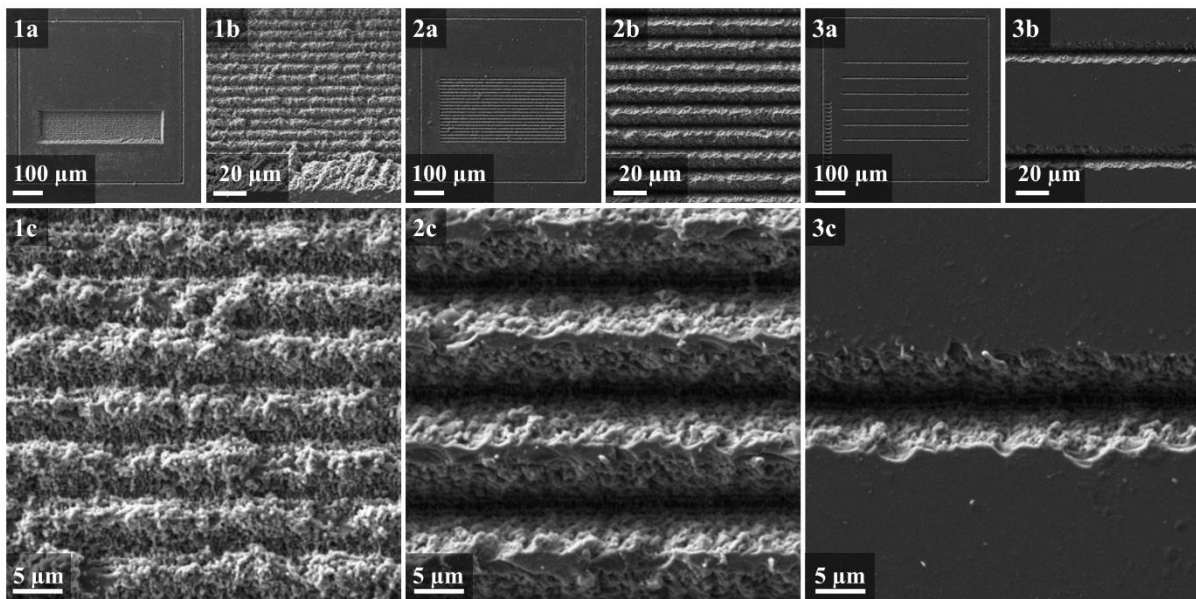


Figure 26: SEM images of laser-machined parallel lines in glass with varying line separation. (1a-c) lines machined with a separation of 5 μm . Lines are laser-machined close enough together that the deeper ablated lines, seen as darker grey and black in (2a-c) and (3a-c), are not produced. Instead, disordered nanoscale features are present in uniform microscale lines. A larger separation of 10 μm (2a-c) or 50 μm (3a-c) creates larger variation in depth, with clear ablated lines surrounded by disordered nanoscale variation. Areas not laser machined are smooth, with virtually no microscale features. Images reproduced from Mackay et al. under Creative Commons Attribution License open access [11] (© 2020 Mackay et al., published by Elsevier Ltd.). SEM images taken by Dr Matthew Praeger.

As both nanoscale and micron-scale topographies alter cell behaviour, it was important to laser-machine a dataset that included 5-50 μm line separation for accurate modelling. Due to the versatility of femtosecond laser-machining (Figure 8 and Figure 26 generated using the same laser system), this method of generating topographies could be adapted to the nanoscale for future experimentation if required. However, micron-scale laser-machining was focussed on, as nanoscale was more time-intensive than the micron-scale, and variation to nanotopography was still introduced through varied line separation. Through control over microtopography, there was an ability to adjust the nanotopography of the glass, from an entirely nanorough surface with shallow grooves (Figure 26 1c), to a mixture of smooth

substrate with deeply grooved areas surrounded by less than 5 μm of nanorough walls (Figure 26 3c).

Squares in the plane of the sample were machined with 500 μm length and width to create a boundary between individual topographical patterns so that, during imaging, there was one square per image at 10x magnification. This ensured cellular responses could not be attributed to neighbouring topographies and simplified the image capture and alignment processes. Additionally, 500 μm squares were large enough to be seen by the human eye at a meter without the need for microscopy², which aided accurate cell seeding upon topographies and microscope alignment.

Training data consisted of straight lines ranging from 10 μm to 25 μm in width, due to areas of overlapping laser-machined areas to create thicker lines, in both uniform parallel arrangements and in varied configurations where the angle, length, and number of lines was randomised using a the Mersenne Twister pseudorandom number generator (PRNG) via NumPy programming. This randomisation increased the information available to the network and ensured unique topographical configurations for testing data extraction.

4.2.2 CELL CULTURE AND IMAGING

The next step after generating topographies was to wash the laser-machined coverslips in 70% ethanol and then place the substrates in an autoclave (a system which subjected the substrate to pressurised saturated steam at 126 °C for 20 minutes), in order to sterilize the coverslips. Cells were then cultured on the substrates and the response collected using fluorescence imaging (Figure 27).

Multipotent skeletal progenitors enriched populations from human bone marrow stromal cells, isolated using a Stro-1 positive antibody, were cultured at 37 °C and 5% CO₂ atmosphere in Alpha minimum essential medium (αMEM) supplemented with 10% fetal calf serum and 1% penicillin/streptomycin, which was changed every three days. Human bone marrow aspirates were collected from haematologically normal patients undergoing routine elective hip replacement surgery, which would have been discarded if not collected, following informed consent from the patients in accordance with approval from North West - Greater Manchester East Research Ethics Committee (Ref-18/NW/0231). The Stro-1 positive

² A 500 μm square is roughly half the size of a poppy seed. A small seed held on the tip of a finger at arm's length away can still be seen, but the texture cannot. Similarly, the topographies were difficult to spot by eye, but the surrounding squares made easy visual location possible.

isolated cell population enrichment was completed using standard protocols within the Bone and Joint Research Group [224, 225]. An example Stro-1 isolation protocol can be viewed in Stro-1 isolation method and protocol.

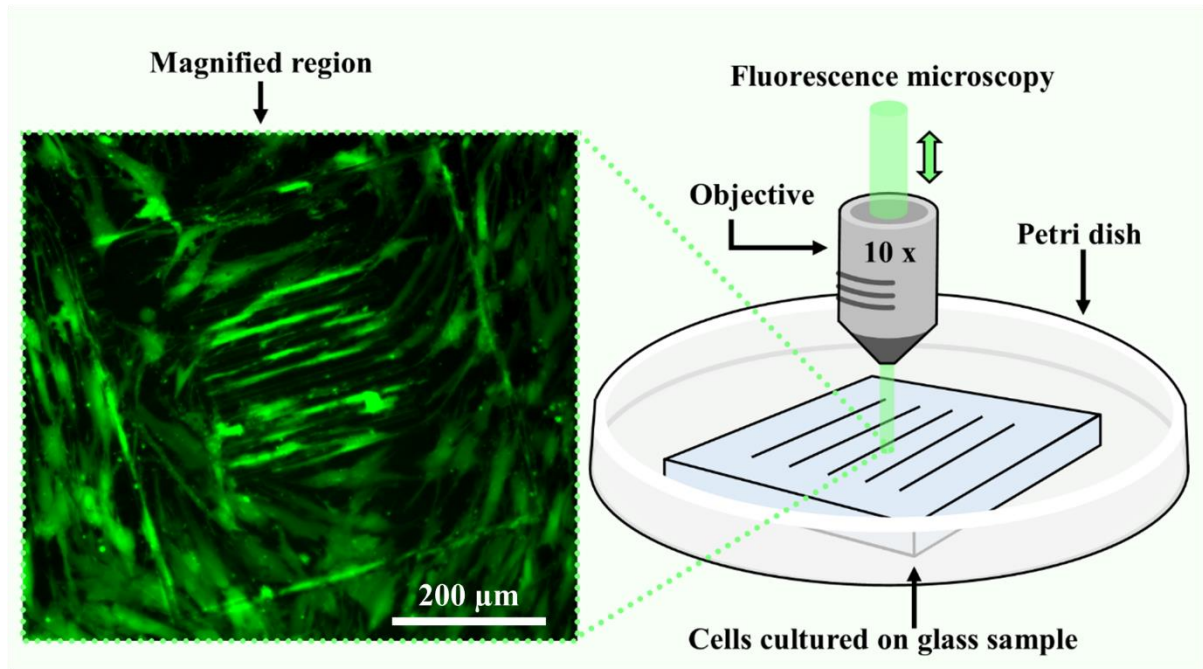


Figure 27: The fluorescent image capture of cells cultured on a laser-machined glass sample. Cells are cultured in standard conditions before being labelled with fluorescent cell-tracker dye, allowing for repeated imaging as cell fixation is not required. Cell positioning and alignment is collected using 10 x magnification, which balances resolution with capturing the whole topography in a simple image.

Fluorescence images were collected using Vybrant cell tracker dye via a Nikon Eclipse microscope at various time points from day 0 until cells were fully confluent (covered the coverslips completely). Full confluency occurred at different time points due to deliberately altered seeding densities to allow for greater variety of training data, with the longest time point set at day 30. This randomisation was to minimise possible network constriction to a limited specific cell density, and cell response could then be determined for a range of densities. Post-imaging editing to images was kept to a minimum, with alterations in brightness and contrast performed through normalisation manipulation to compensate for any alterations in brightness: While all images were taken with an equal exposure time, some fluctuation in fluorescence was unavoidable. These alterations ensured that values for cell density and cell positions, which were derived using image recognition techniques, were consistent between measurements and cell samples. In total, 203 unique combinations of topography and time-point (from day 0 to day 28) were used to train the network.

Data: Collection and Composition

After laser-machining topographies, culturing cells on the topographies and then imaging the cells, three images were formed for each of the 203 combinations. This dataset included:

- i. The design of the topography, such as examples Figure 28 a and d, was stored as a binary array of laser-machined lines, white, on a smooth unaltered surface, black. This was the input, alongside the time-point and cell density, to the network.
- ii. A brightfield image of the cells after a recorded number of days on the surface topography, such as examples Figure 28 b and e, was used to check the cell positioning relative to the clearly visible laser-machined lines, and to check for fluorescent debris (organic waste from dead or damaged cells).
- iii. The fluorescent image of cells, with the same magnification and positioning as the brightfield image, such as examples Figure 28 c and f, showed live cells and their positioning. This was the output of the network.

Figure 28 a-c shows one dataset corresponding to a topography with parallel laser-machined lines and Figure 28 d-f shows a dataset for an unmachined topography.

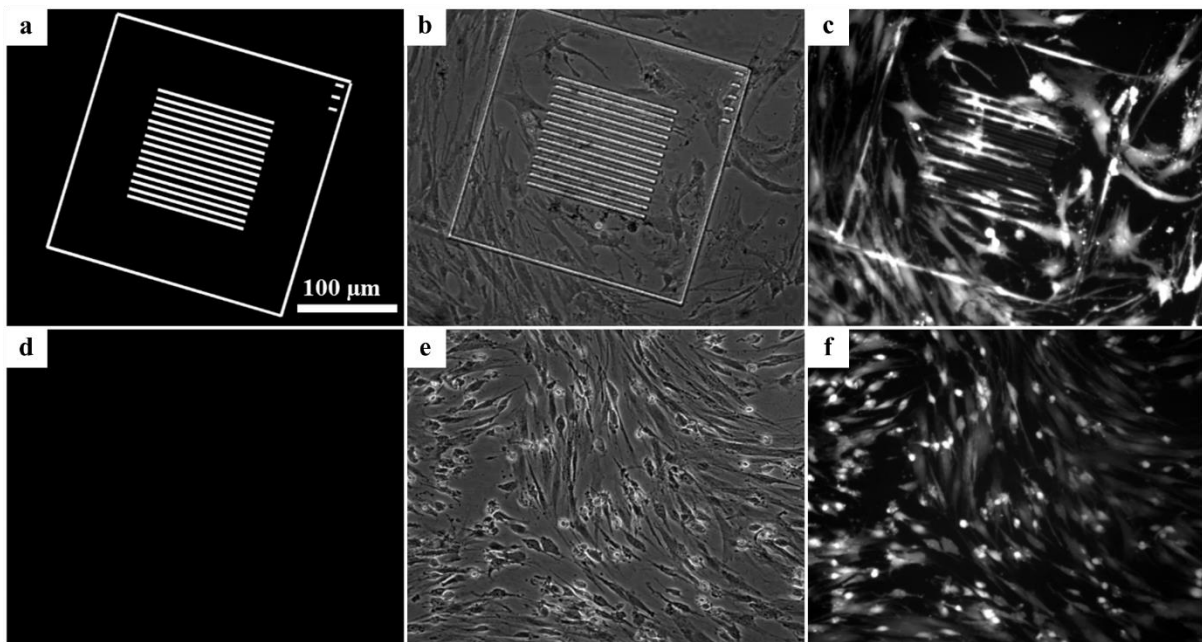


Figure 28: Skeletal Stem cell enriched population response to laser-machined topographies, with adhesion and alignment along the parallel lines.

(a, d) The surface topography, illustrated by a binary array where white is laser-machined and black is unmachined areas. (b, e) Brightfield microscopy shows both the laser-machined lines and the cell positioning. (c, f) Fluorescent microscopy provides a clearer visual of cell position, with cells aligning parallel to and adhering to laser-machined areas. In contrast, cell positioning is more randomised in unmachined areas of (a-c) and in all areas of (d-f), where there are no laser-machined areas. All images are at the same magnification. Image modified from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

While the brightfield image was not used by the network, and instead was used to check the quality of both the laser-machined topography and cell culture, the topography array and

fluorescent image were used as input and target output data for the network. The brightfield image was also useful for alignment guidance, as it contained partially visible cell and topography data. Additionally, autofocus software was more effective with brightfield than fluorescent microscopy during the imaging process, thus collection of brightfield images at identical positions to fluorescent images added less than a second per position to the process time. Once a binary array was aligned to the corresponding brightfield image, the array was resized so that both images had the same resolution, while keeping the microscope data unchanged.

4.2.3 COLLECTED DATA COMPOSITION

The following table, Table 2, shows the number of raw image sets collected via the above process. Each image set consists of a binary image for topography, a brightfield image for cell positioning relative to topography, and a fluorescent image of cell positioning and response.

Tag	Gender	Age (Years)	Days Imaged	Number (raw)	Training
F63	Female	63	0, 1, 8	21	No
F77	Female	77	0, 1, 8, 15	149	No
F83	Female	83	0, 1, 8, 15	28	No
F62	Female	62	30	22	Yes
F84	Female	84	1, 8	61	Yes
M55	Male	55	0, 1	27	Yes
M59	Male	59	1, 8, 30	78	Yes
B0	-	-	0, 1, 8	15	Yes

Table 2: A complete list of the collected stem cell images used in training and testing. Combinations in bold ("No" in the testing column) were extracted for testing.

In total, before augmentation, 203 images were collected for training and 198 were collected for testing. The comparative increase in the number of images within the F77 set was due to the timing of data collection. Lockdown measures had legally come into force in the UK and therefore an increase in the number of images was taken to guard against the event that more data would not be collected at a later date. The B0 set included topographies without stem cells present.

A limited cross-validation was performed on the dataset to determine whether extracting different patient image sets for testing resulted in a change in the relative success of the

Data: Collection and Composition

network output. The images corresponding to stem cells extracted from an individual patient were extracted from the overall dataset for testing, while the remaining images were used to train a U-Net for 24 hours. B0 was not included in this comparison. This is not sufficient for providing accurate outputs, the error was above 50% for all network outputs, but a comparison is provided in Figure 29. As this is not a true cross-validation, the results were not used to quantify uncertainty.

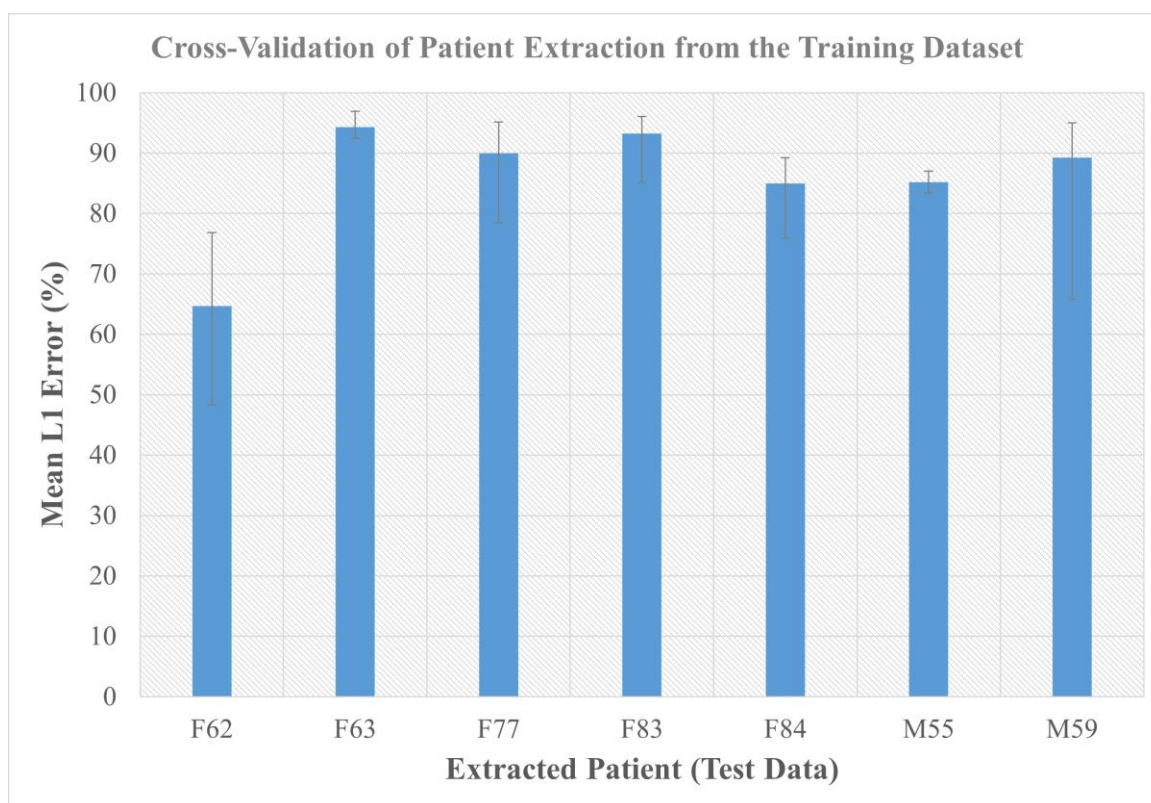


Figure 29: A graph for comparing the mean L1 error of network output when the images of stem cells from a single patient are extracted from the training dataset and are instead used for testing.

The neural network was trained for 24 hours on seven different training datasets. The datasets contained images from each patient set except a single extracted patient image set that was used for testing. Error bars show the maximum and minimum L1 error for each testing set.

When the exclusion of F62 images is not considered, extracting any specific patient images does not affect the output of the neural network by more than 10%. However, excluding F62 images provided an error roughly 20% lower. This is because the number of cells in the F62 were higher than for the other patient image sets, so the fluorescent images were brighter (more cells result in more areas of fluorescence). As the outputs from the neural network after 24 hours were relatively binary, with block areas of black and white, the high levels of white in the outputs images were closer in pixel value to the F62 test images.

While this limited cross-validation cannot be used to provide a quantitative uncertainty for the outputs in Chapter 6, it showed that F62 should be included into the training data as any

calculated error based on F62 test images may be lower than for testing based on other patient image sets.

4.3 STEM CELL RESPONSE TO SCAFFOLDS – DATA FOR CHAPTER 7

The following table, Table 3, is a list of the materials and coating combinations provided by Dr Karen Marshall of the University of Southampton (K.M.Marshall@soton.ac.uk).

All egg CAM procedures were carried out in accordance with the guidelines and regulations stipulated in the Animals (Scientific Procedures) Act, UK 1986. The chick embryo CAM model was under Home Office Project license P3E01C456.

4.3.1 DATA FOR *IN VITRO* BIOCOMPATIBILITY

This dataset consisted of fluorescent images of seeded cells after 1 and 14 days of culture on corresponding scaffolds, which could then be analysed or used by a neural network (Figure 30). The collection of experimental data was performed by collaborators, principally Karen Marshall at the University of Southampton, which involved creating, coating and seeding cells on the scaffolds, and then imaging using fluorescent dyes. Processing, application and analysis of this data was then undertaken by the author for network training and testing.

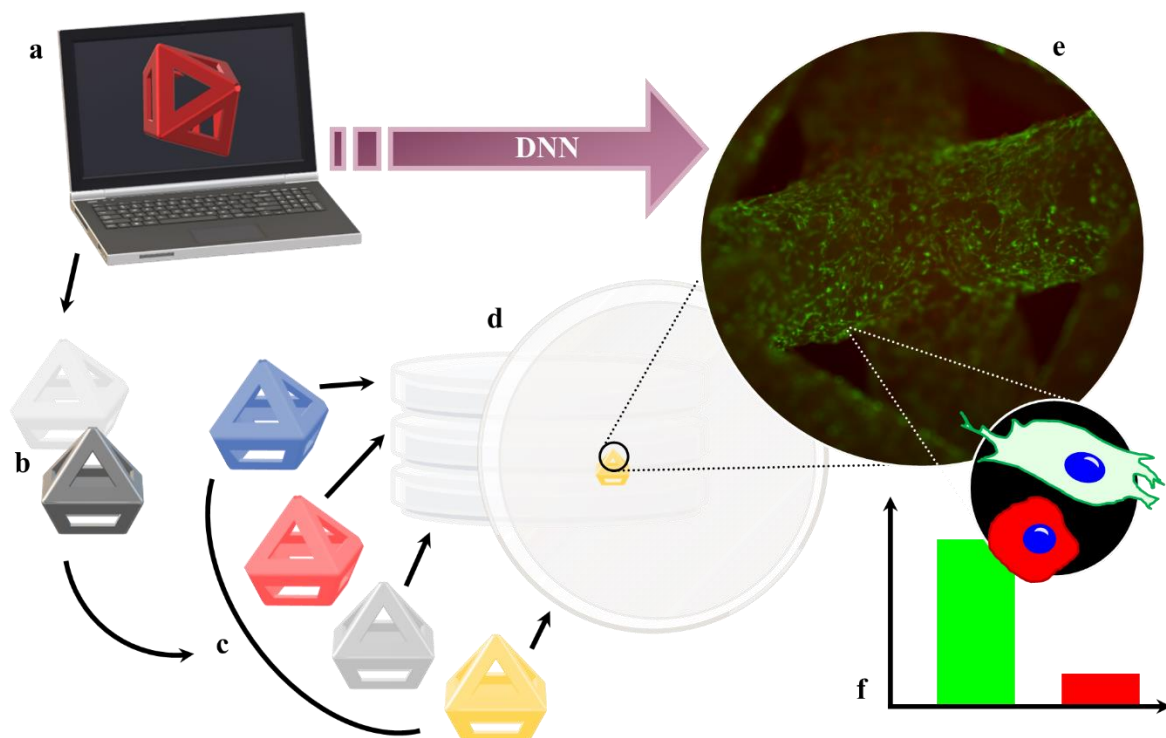


Figure 30: Material biocompatibility prediction method for the *in vitro* data transformation. Once a scaffold has been designed (a), the two different material scaffolds chosen (b) must be coated with a range of coatings (c) designed to promote angiogenesis (formation of new blood vessels) or biocompatibility. Cells are then seeded and cultured on the scaffolds (d) and imaged using fluorescent cell staining (e) for future analysis (f).

Data: Collection and Composition

HBMSCs were grown on either nylon or titanium scaffolds (Figure 30 b), with a variety of coatings (Figure 30 c), generating varying cell responses, including a mineral coating, a coating generated by Glasgow University for osteogenesis promotion (referred to as ‘Glasgow’ coating), and scaffolds left uncoated (plain) [226, 227]³. Cultured for 14 days using conditions used in previous work (chapter 6) (Figure 30 d), these coated and non-coated scaffolds were imaged at day 1 and day 14 using calcein acetoxymethyl (AM) dye and a Zeiss stereomicroscope and digital camera (Figure 30 e) [17]. Enzymatic conversion of virtually non-fluorescent calcein AM resulted in fluorescent calcein, producing green wavelengths in live cells. Additionally, ethidium homodimer-1 (EthD-1) was used to determine cell health, as it is a nucleic acid stain that enters cells with damaged membranes, therefore staining dead and dying cells red (EthD-1 is excluded by the intact plasma membrane of viable live cells).

The microscope images were subsequently analysed computationally to compare the number and health of cell culture for each scaffold-coating combination, which were approximated through the amounts of calcein AM (green pixels) and EthD-1 (red pixels) fluorescence (Figure 30 f). Collecting training data of scaffold material-coating combinations (Figure 30 a) and fluorescent imaging of cell-seeded scaffolds after 14 days of culture (Figure 30 e) takes several weeks, due to cell culture and processing times. As in previous work (chapter 6), a neural network was used to speed up the process of testing cell response to environmental cues by transforming information on cues, which contained the choice of scaffold material and coating, to the resulting cell response via the fluorescent image. As 3D scaffolds were imaged through 2D microscopy, only portions of the fluorescent images were in focus, and this was included in the input data. The full list of scaffold material-coating combinations is seen in Table 3.

Tag	Material	Coating	Day Imaged	Training
NB01	Nylon	BMP 2 peptide	1	Yes
ND01	Nylon	CD3 peptide	1	Yes
NE01	Nylon	CE3 peptide	1	Yes
NG01	Nylon	Glasgow	1	No

³ Further details on these scaffolds and coatings are awaiting publication and outside the scope of this project. Research funded by the UK Regenerative Medicine Platform (MR/R01565/1) and is a collaboration between the University of Southampton, the University of Glasgow, and Imperial College London.

NM01	Nylon	Mineral	1	Yes
NP01	Nylon	Plain	1	Yes
TB01	Titanium	BMP 2 peptide	1	Yes
TD01	Titanium	CD3 peptide	1	Yes
TE01	Titanium	CE3 peptide	1	Yes
TG01	Titanium	Glasgow	1	Yes
TM01	Titanium	Mineral	1	Yes
TP01	Titanium	Plain	1	Yes
NB14	Nylon	BMP 2 peptide	14	Yes
ND14	Nylon	CD3 peptide	14	Yes
NE14	Nylon	CE3 peptide	14	Yes
NG14	Nylon	Glasgow	14	No
NM14	Nylon	Mineral	14	Yes
NP14	Nylon	Plain	14	Yes
TB14	Titanium	BMP 2 peptide	14	Yes
TD14	Titanium	CD3 peptide	14	Yes
TE14	Titanium	CE3 peptide	14	Yes
TG14	Titanium	Glasgow	14	Yes
TM14	Titanium	Mineral	14	Yes
TP14	Titanium	Plain	14	No

Table 3: A complete list of the material-coating combinations used in training. Combinations in bold (“No” in the testing column) were extracted for testing.

Out of the 24 scaffold-coating combinations used for the first data transformation, two were extracted for testing. One, NG01 and NG14, was a nylon scaffold with the ‘Glasgow’ mineral coating after 1 and 14 days of cell culture, and the other combination, TP14, was a titanium scaffold without a coating (left plain) after 14 days of cell culture. The inclusion of both nylon and titanium scaffolds gave possibility for comparison between materials, and the inclusion of a coating and a scaffold without a coating provided further potential comparisons. Additionally, the error in predictions when the network had training data at day 1, but not day 14, and with no training data at either day 1 or 14 could be investigated.

4.3.2 PREDICTING *IN VIVO* BIOCOMPATIBILITY

Collecting data for *in vivo* biocompatibility was collected by a collaborator for separate biocompatibility experiments [17], and complied with NC3Rs guidelines. The CAM assay

Data: Collection and Composition

images were then repurposed for use in this study through cropping and pairing with corresponding fluorescence imaged scaffolds.

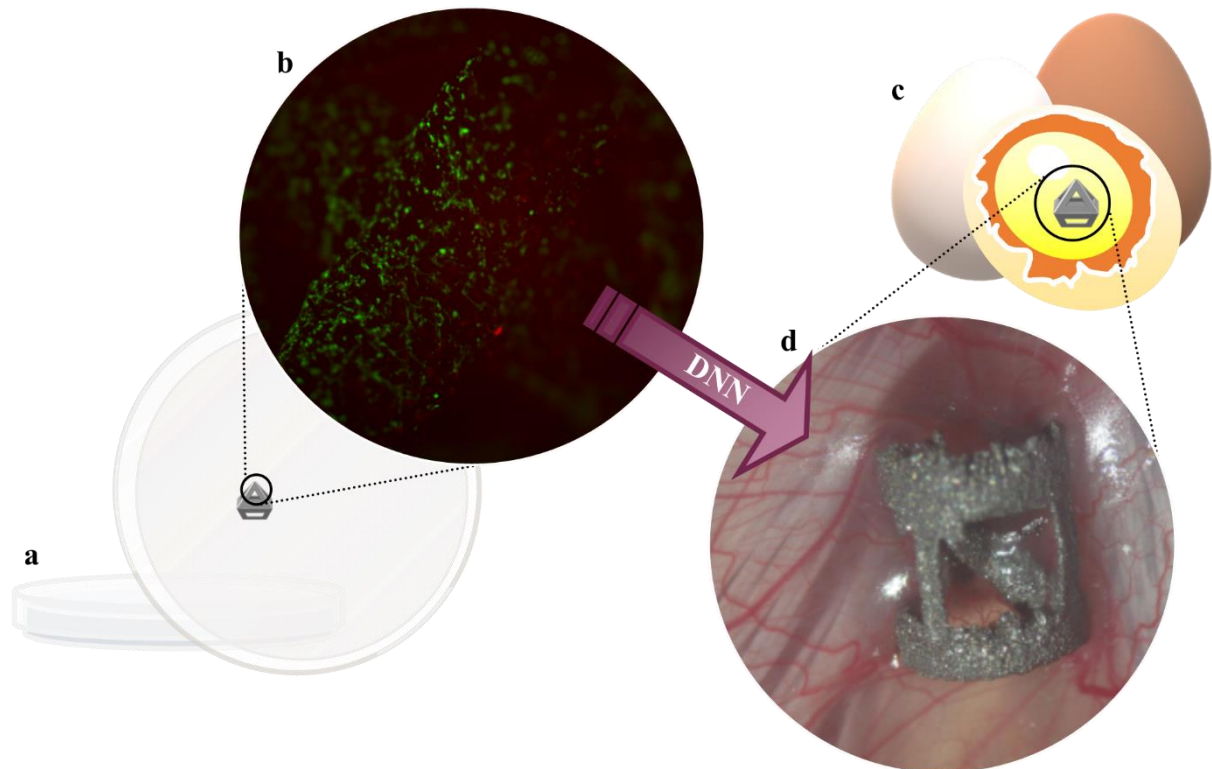


Figure 31: Generating data input-output pairs for training a DNN to transform fluorescent images to CAM assay images. (a) A scaffold is seeded with HBMSCs cells and cultured for 14 days. (b) Fluorescent images are taken to determine cell health on the scaffold. (c) A scaffold of the same material and coating is implanted into a chick egg and left for 14 days. (d) CAM assay images are taken of the scaffold to determine health and integration. Fluorescent images are used as the input to the DNN, which transforms the data into CAM assay images. These network generated images can be compared to laboratory obtained images for network training and testing.

The scaffolds were seeded in parallel experiments, where each scaffold material and coating combination was used for *in vitro* biocompatibility testing (Figure 31 a-b) and also used for *in vivo* testing (Figure 31 c-d). For *in vitro* testing, this involved seeding cells on a scaffold that was then cultured in a petri dish for 14 days, and then fluorescent imaging of the cells. For *in vivo*, a different scaffold with the same material and coating combination was placed on a CAM and then imaged after 8 days. While the same material and coating combination was used for both scaffolds, the experiments were independent.

There were a greater number of augmented fluorescent images to CAM assay image, so the fluorescent images were randomly paired with corresponding CAM assay images with the same material and coating. This was possible due to the independence of *in vitro* and *in vivo* experiments and datasets, but an individual augmented fluorescent image was used no more than once per dataset to try and limit overfitting. While this could result in limited outputs from the network, this increased the size of the training dataset with limited *in vivo* data. The

scaffolds were once again either nylon (N) or titanium (T), with coatings such as mineral (M), the coating developed by the University of Glasgow, 'Glasgow' (G), and left plain (P) (uncoated). These tags were combined to indicate the material-coating combination of each scaffold, such as NG for 'Glasgow' coated nylon. Uncoated titanium (TP) and uncoated nylon (NP) scaffolds were included in both the training and testing data, yet the images in the testing data were from a different experiment performed on a different day under the same conditions. The mineral coated titanium (TM) and nylon (NM) scaffolds were not included in the training data, and instead were extracted for testing. Additionally, two NG and two TP scaffolds were removed from both training and testing data for a network-series test, where the two networks were tested together. The material/coating information (NG and TP) was input into the first network, and the output fluorescent image was then used as the input to the second network.

5 LABELLING 3-DIMENSIONAL PLACENTAL TISSUE

The research covered in this chapter was published and short-listed for best presentation at BIOSTEC2020 in Valetta, Malta [220]. A book chapter published by Springer was also based on this research [228].

5.1 INTRODUCTION

While there has been a lot of work over the last decade in non-invasive labelling of cells via machine learning, most of this work has involved simple 2D cultures, where cells are cultured on cell-culture plastic and are not confluent at the time of imaging [196, 229]. There are significant benefits to such imaging, such as real-time monitoring of cell density and morphology, but similar labelling for 3D structures have not progressed as quickly, even though it is crucial for understanding cellular behaviour, the biology of the tissue, and how cells work together. For example, the interaction between fibroblasts, pericytes, and endothelial cells in placenta can improve understanding of how these features affect placental nutrient and waste transfer [2]. To model and analyse these complex interactions, 3D visualisation is required, and 3D-labelling is therefore a necessity to the medical imaging field (Figure 32).

Labelling 3-Dimensional Placental Tissue

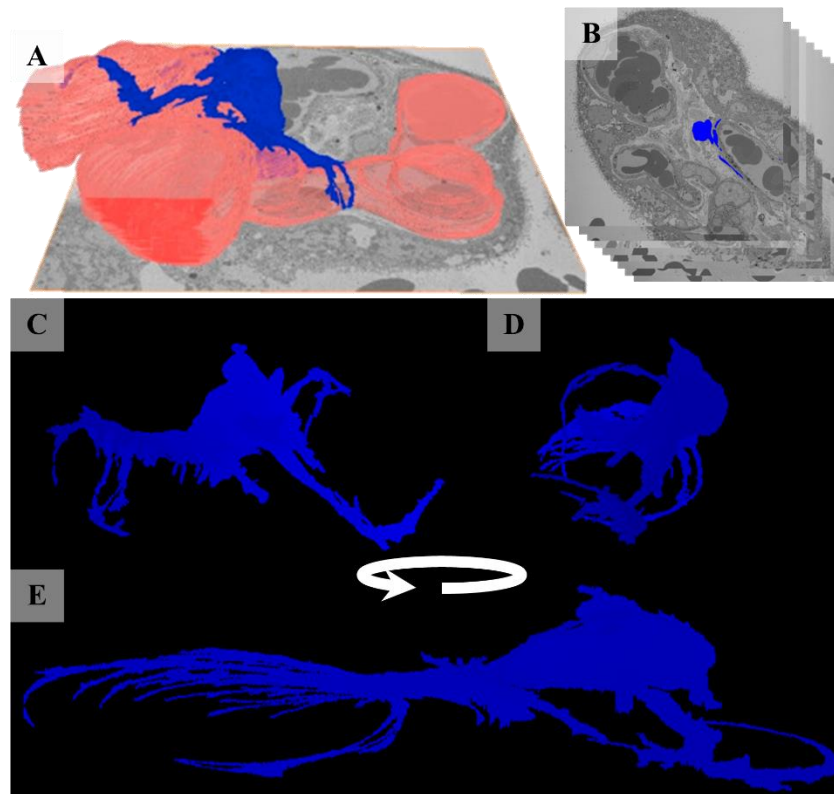


Figure 32: Labelling SBFSEM-imaged placenta to gain 3D projections of pericytes.

To obtain an understanding of how pericytes (blue) interact with placental capillaries (pink), 3D projections are required (A). SBFSEM images are taken and labelled (B), which consist of thousands of SEM images in a stack. By combining the labels throughout the stack, large 3D models can be created using readily available software (C-E), which allows for rotation and visualisation not possible in 2D. Figure (A) modified from Harris et al. under Creative Commons Attribution license open access (© 2020 Harris et al. Published by Elsevier) [2].

3D visualisation requires multiple steps in order to form high-quality models for cell behaviour analysis. The first step is imaging the cell using methods where the 3D structure can be obtained, such as with the 3D technique for scanning electron microscopy (SEM) image, serial block face scanning electron microscopy (SBFSEM), which shows the internal cellular structures as well as surfaces. SBFSEM and other imaging methods face a lengthy process of labelling the target cell, or in some way extracting it from the surrounding cellular and acellular material [160]. While methods such as confocal microscopy could in theory bypass this stage, it would not allow for the nanoscale precision required for high-quality individual cell analysis. This step can severely limit the use of 3D-imaging methods, as labelling a single cell from a standard SBFSEM stack (a collection of SEM images collected via the serial block face technique) can take an expert as long as a month, or longer for larger and more detailed stacks. Only after a stack has been completely labelled can the extracted labels be projected into a single 3D visual, which is easily done through open-source software such as ImageJ [230].

With the use of faster AI-based approaches, 3D imaging could be more widely adopted and new understanding of cell behaviour accelerated, which is briefly explained in section 2.5. A U-Net, a form of DNN typically used for image-to-image data transformation, was trained on SBFSEM stacks labelled manually, and corresponding unlabelled stacks, which is covered in section 5.3. The features labelled included fibroblasts (5.4.1), pericytes (5.4.2) and endothelial cells (5.4.3) within placental tissue. The results are then summarised in section 5.5.

5.2 HYPOTHESIS, AIMS AND OBJECTIVES

While analysis of 3D cellular structures within placenta is necessary for advancing research into placental growth-factors, 3D image processing can be labour intensive due to the length of time spent manually labelling each section (of which there can be thousands per image). Using a network to extract and label desired features can potentially overcome time and labour constraints in research reliant on 3D imaging.

5.2.1 HYPOTHESIS

Unlabelled SBFSEM sections can be automatically labelled to identify and label specific cell types without labelling undesired features on unseen SBFSEM sections.

5.2.2 AIMS AND OBJECTIVES

The project is split into several sections, with the tasks rising in complexity with each additional experimental aim.

- 1) Automatically label a single cell within a SBFSEM stack, with an average error of less than 5%.
 - a) Use manually labelled SBFSEM sections to train a network architecture developed for image-to-image data transformation.
 - b) Compare network labelled and manually labelled SEM images and quantify the differences to form an appropriate error.
- 2) Determine whether any single cell within a SBFSEM stack can be automatically labelled when manually labelled features are limited to less than 20% of the stack.
 - a) Train a U-Net based network architecture on multiple different structures to evaluate versatility of the network.
 - b) Limit the training data to less than 1 in 5 of the SBFSEM stack and compare the error to network output when more and less training data was used.

- 3) Automatically label multicellular structures and maintain an average error of less than 5%.
 - a) Compare network labelled and manually labelled SEM images for a multicellular structure.
 - b) Evaluate the network labelling of a SBFSEM stack that has not been used in training.

5.3 METHOD

This study used term human placental tissue, collected after delivery from uncomplicated pregnancies with written informed consent and ethical approval from the Southampton and Southwest Hampshire Local Ethics Committee (11/SC/0529). All methods and experiments were approved and performed in accordance with the relevant guidelines from the Southampton and Southwest Hampshire Local Ethics Committee and the University of Southampton. Full methodology of imaging processes of placental tissue can be found in Lewis *et al.* and Harris *et al.* [2, 160].

5.3.1 SINGLE CELL LABELLING

Automated labelling of a 3D single cell from SBFSEM data requires multiple steps (Figure 33), so that a single SBFSEM z-stack can be used for both training and testing. The first step required inputting a 2D section, a z-stack image, of the SBFSEM data to a trained neural network, capable of identifying and labelling the desired cell. The second step was collection of the output from the neural network, followed by (the third step) isolation of the label. The final step was collecting all isolated labels and using readily-available open-source software, such as ImageJ, for creation of a 3D projection. While an entire SBFSEM z-stack could be input into a (much larger) neural network and trained to output a 3D projection, therefore skipping the second and third steps, this would increase several requirements:

- i. Up to several thousand labelled SBFSEM z-stacks, which would take a dedicated team several years to produce, instead of the three used for this chapter.
- ii. A GPU larger, and therefore more expensive, than can be used in a standard high-quality PC set-up (such as those used for gaming or image editing).
- iii. Additional time spent training the network, as there would be more data parameters.

By keeping the input and output to the network limited to 2D, less training data was required, less time training the network was required, and relatively cheaper computer set-ups could be used.

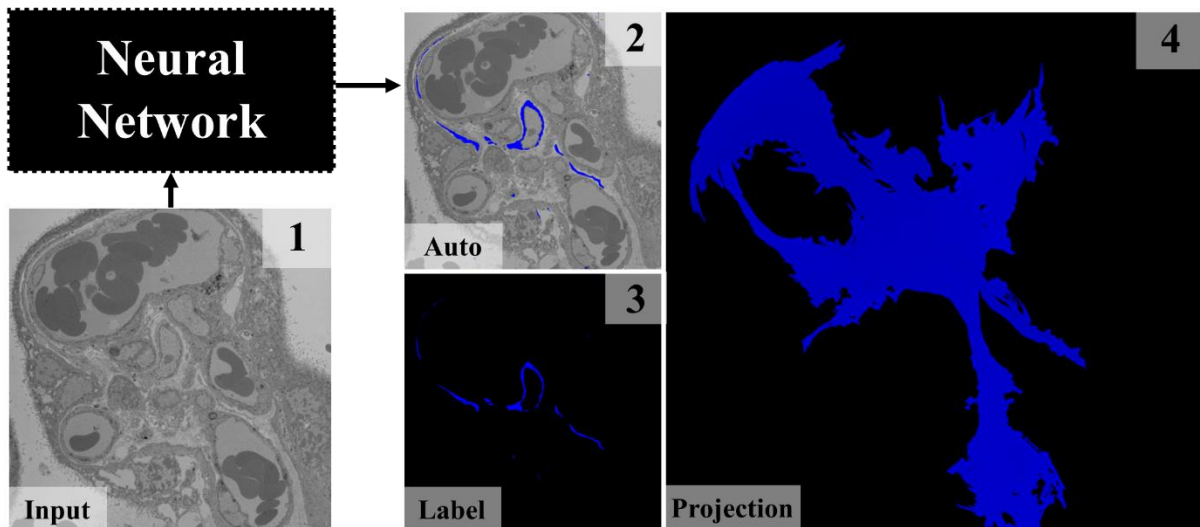


Figure 33: The method of single cell automated labelling, which contains several steps.

The first step (1) is to input a 2D segments, a z-stack image, from the 3D SBFSEM block into a trained neural network. The second step (2) is the output of this trained neural network, which produces a SEM image which the feature cell labelled. The third step (3) is extraction of the label from the surrounding SEM image. The final step (4) is 3D-projection of all extracted labels into one 3D cellular structure.

Both the network architecture and the data used provide large variations to the success of training, as badly designed or poorly chosen architectures will either fail to train or will take too long to produce a desired output, and a lack of adequate data can lead to a range of problems, from overfitting (looked at in more detail in section 5.4.2) to failed optimisation of minimising the loss function (no solution for transforming input data to output data is formed). For the task of transforming one image (the input z-position) to another image (the output labelled z-position), a U-net architecture was chosen, which was trained on sections of a labelled SBFSEM z-stack (Figure 34).

The U-Net consisted of 17 layers, each containing a convolution or deconvolution/up-convolution with a stride of 2, a 4 by 4 kernel size, and an accompanying rectified linear unit activation function. The dimensions of the input data therefore altered from 256 by 256 pixels over three channels to 1 by 1 across 1024, before being resized back up into a 256 by 256 three channel image. Skip connections between the encoder, the convolution side, and the decoder, the deconvolution side, resulted in spatially relevant information passing across the architecture, therefore resulting in more realistic output images. To further improve the realism of output images, the architecture also included a discriminator, which had a random input of either an automated or manually labelled image. The discriminator was tasked with determining whether the input image was real or network generated, and was a simple 4 layer ConvNet, utilising a sigmoid activation function to produce a real vs fake prediction.

Compared to the prediction of the discriminator, pixel-to-pixel comparison error (L1 loss)

Labelling 3-Dimensional Placental Tissue

between the U-Net generated image and the manual image was given a weighting 100 times larger within the loss-function. This combined discriminator-L1 error function was then used via standard back-propagation, which results in the network altering weights to produce different data transformations. Through optimising the transfer of data propagating through the network, and minimising the loss-function, the network is trained.

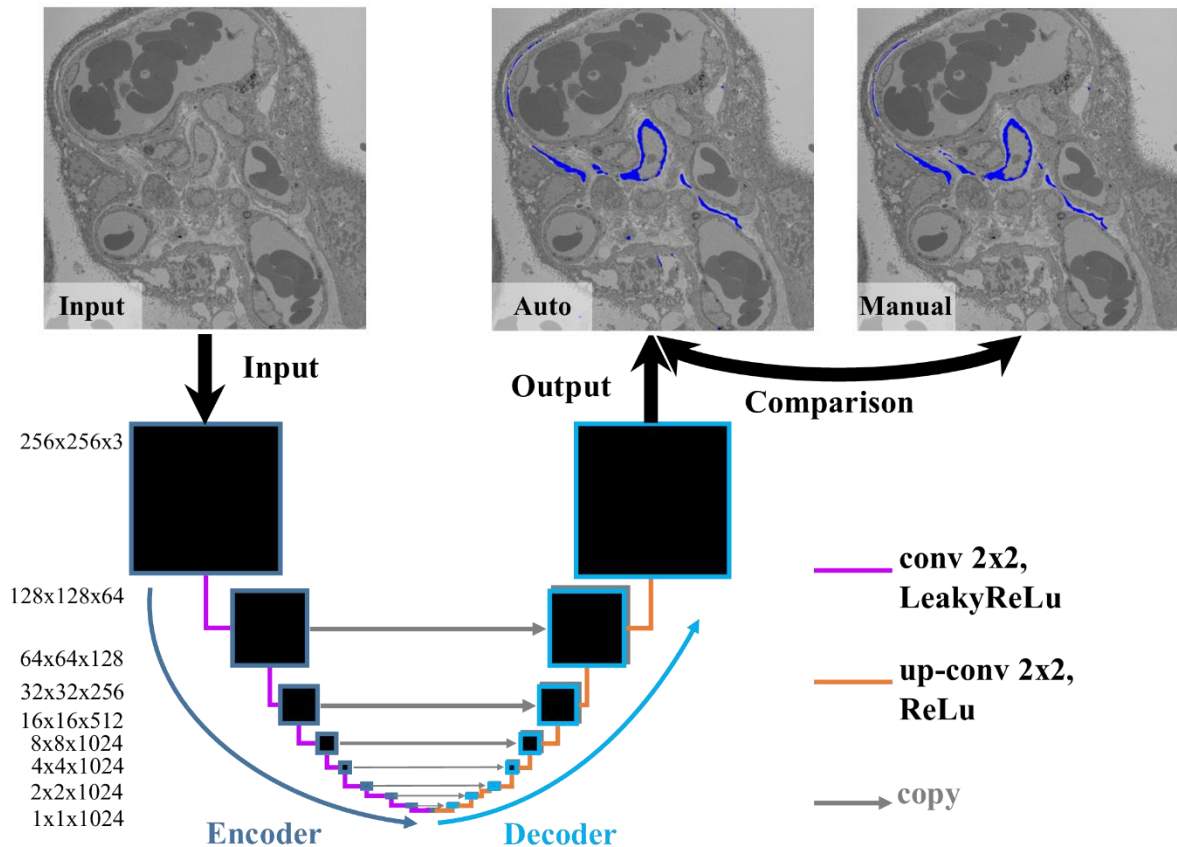


Figure 34: The training of a neural network for the automated labelling of feature structures.

An unlabelled image is input into a neural network consisting of many convolutional layers, deconvolution or up-convolution layers and skip-layers between the encoder-decoder layers. Output images are compared to manually labelled images, and differences lead to changes and optimisation of the network layers through backpropagation. (Additional error correction is performed from a separate convolutional neural network to determine how realistic the output images are in comparison to manually labelled target images.)

Although three z-stacks were used in this chapter, only the first used for labelling fibroblasts and pericytes, which consisted of over 1000 z-positions, of which 943 were used. The second and third were used for the labelling of endothelial cells and consisted of 370 and 140 z-positions respectively. For labelling fibroblasts, the network was trained on 50% of the stack and then tested on the remaining 50%. The 50% used for training, 471 z-positions, were expanded through use of standard data augmentation techniques, including resizing the images from 4096 by 4096 pixels into 2000 by 2000, cropping the resized images into a minimum of eight 512 by 512 images, and then resizing again into 256 by 256. The reduction

in resolution resulted in the network taking only six hours on an NVIDIA TITAN Xp GPU to perform 100 epochs (the network had every training image input 100 times).

The only data input to the network was the unlabelled-labelled pairs, as is standard for training deep learning models without any pre-training. Therefore, this technique is not limited to SBFSEM images and can be applied to other 3D imaging techniques.

5.3.2 MULTI-CELLULAR STRUCTURES

For larger cellular structures, such as a multi-cellular ring of endothelial cells, a W-Net is required to allow for the increase in complexity [11, 220, 228]. A series of U-Nets, the W-Net consists of two encoder-decoder pairs, with 33 layers in total, but only a single discriminator. Increasing the number of layers increased the complexity of the network, so training took approximately 24 hours for 50 epochs, compared to only ~ 6 hours for the U-Net, even while all other hyperparameters and parameters remained unchanged. Unlike increases in filter number, increasing the number of layers did not require an increase in the necessary GPU size requirements for the network to train and run, as the dimensions of the data throughout the network are no bigger. The effective image size was reduced from 256 by 256 to 1 by 1 and increased back up to 256 by 256, as for the U-Net, before being repeated.

Labelling of endothelial cells focussed on the adaptability of the network, and therefore general usability, instead of on the overall accuracy and performance. As well as extracting 50% of the images for testing, a further 10% section was extracted from the centre of the stack to produce a “blind spot” for the network. An additional section at end of the stack was deliberately left unlabelled to determine the extent of network adaptability while avoiding overfitting.

5.3.3 TESTING THE NETWORK

When using the network for testing, there is no use of backpropagation or changes to the network, as there are no comparisons made by the network to manually labelled images. All analysis is conducted separate from the network, using a range of statistical markers rather than a loss-function, of every output. Once analysis of each individual output is complete, all output labels generated from the network (from both the unseen testing data and seen training data inputs) are extracted and 3D projected for a final 3D visual analysis to determine the success of the network.

5.4 RESULTS AND DISCUSSION

Error analysis was completed on every unseen z-position using pixel-to-pixel comparisons between automated and manually labelled images. A labelled pixel had the red, green, and blue (RGB) value [0,0,255], visualised as blue on a standard three channel image, and any deviation from this value was classified as an unlabelled pixel. To elaborate, even a single value difference from the label values, a pixel difference of ± 1 or greater in any of the three colour channels, such as [0,1,255] or [0,0,254], would not be identified as a labelled pixel, even though it is virtually indistinguishable to the human eye. The aim of the network was to obtain 100% total matching pixels, where all labelled and unlabelled areas on the network output matched the manually labelled image identically.

5.4.1 ANALYSIS OF FIBROBLAST LABELLING

An area where both an automated and a manually labelled z-position image had a labelled pixel was considered a true positive (blue in Figure 35). Areas where the same area in both were unlabelled was considered a true negative (black in Figure 35). If the area in the network output did not match the same area in the manual image, this was considered a false positive (green in Figure 35) or false negative (red in Figure 35), depending on whether the network had labelled or not labelled respectively. It would be imprecise to categorise areas of true positives and negatives as correct, and areas of false positives and negatives as incorrect, as there is the possibility of human error, especially in areas of highly detailed, thin, structures.

Three positions in the central region of the z-stack were chosen for further visual analysis (Figure 35), at the z-position 300, 400 and 500 (out of roughly 1000 positions), as the central region had some of the lowest total matching pixels. This was not surprising, as this was the area where the cellular structure has the largest amount of surface area (the location where imperfect labelling was most likely). For each position, the input z-position image, automated labelled image, manually labelled image and a comparison image was inspected. The input image provided a comparison for both the labelled images, to ascertain which areas should have been labelled or left unlabelled. As the automated and manually labelled images were very similar, the comparison image provided a clearer visual for areas of differences between the two. The vast majority of pixels in the comparison images were blue and black, with over 99.5% true positives or true negatives, shown when the positions are analysed numerically (Figure 36).

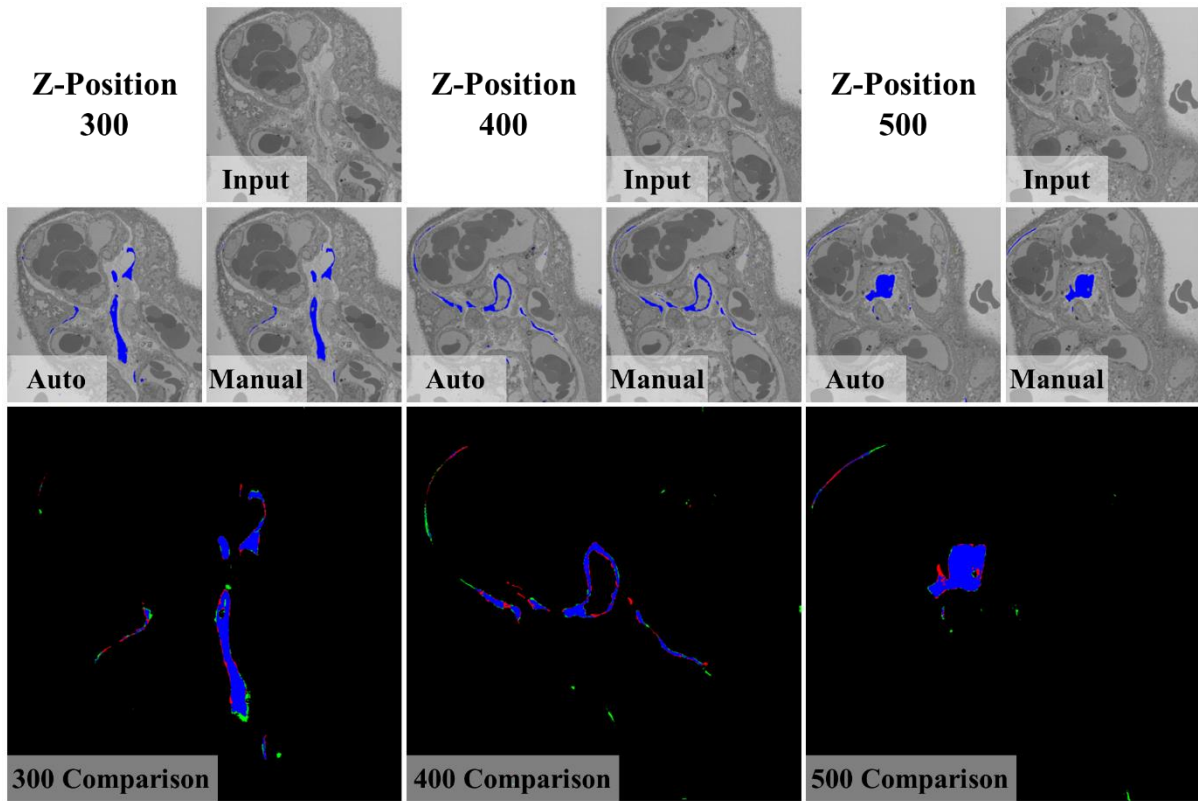


Figure 35: Comparison of automated and manually labelled SEM imaged placenta for three different z-stack positions. Each position, 300, 400 and 500, shows the input to the network (an unlabelled SEM image), automatic network labelled output and manual labelled image, and the comparison image. The comparison images contain four different colours. Black pixels are a position of a true negative, where the automated and manual images show no labelling. Blue pixels are positions of true positive, where both show labelling. Red pixels are positions of false negatives, where there is manual labelling but no automated labelling. Green pixels are positions of false positives, where there is automated labelling but no manual labelling.

There were some areas of discrepancy in position 300, with areas around the labelled sections speckled in both red and green. This mixture of both false negatives and false positives in close proximity was a sign of minute differences between manual and automated labelling of boundary regions, which were inherently subjective due to both computational and human limitations. Computers are limited by data resolution, while humans can also be limited by screen resolution, hand-eye coordination and cursor control, etc. The differences in areas on the left were therefore not concerning. However, the small green circle on the bottom right was not correct, as it was labelling a position within a blood vessel (not a fibroblast section). This speckling of green, small areas of less than 0.5% of the image in total, was also seen in position 400 and 500, which could be a network bias towards labelling small areas of uncertainty. As there was only one manually labelled fibroblast, but more than one fibroblast in the z-stack, this bias was a promising sign that the network could identify more than one fibroblast if not trained on a limited dataset. This inherent uncertainty likely lowered the

Labelling 3-Dimensional Placental Tissue

overall ability of the network, and future work should focus on a more optimal training dataset, where all desired features are labelled.

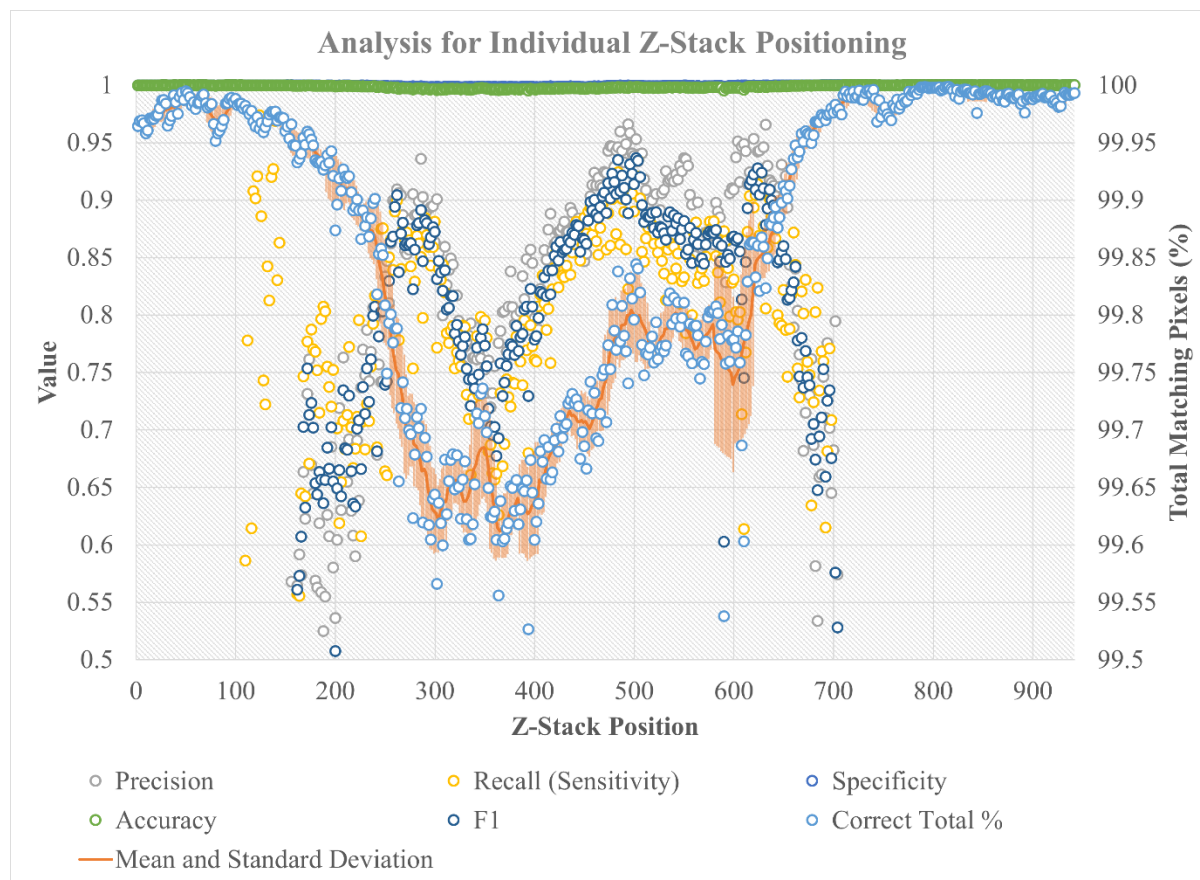


Figure 36: Graph of multiple statistical analysis factors for comparing automated to manually labelled SBFSEM placenta. Left axis of 0.5-1.0 is for precision (grey circles), accuracy (green circles), recall (yellow circles), F1 (navy circles), and specificity (blue circles, almost hidden behind accuracy). Right axis of 99.5-100.0% is for Correct Total % (light blue circles) and the accompanying rolling mean and standard deviation of these points (orange). A pixel is considered correct, a true positive or true negative, when automated labelling matches manual labelling.

Even with a suboptimal dataset, where there were potentially multiple fibroblasts but only one was labelled, the network was very successful in labelling the fibroblast from unseen z-positions. When the entire stack was numerically analysed, the total matching pixels for each position never dropped below 99.5% (Figure 36), shown as light blue circle markers. The orange line that intersected the light blue markers was the mean average pixel match over ten images, and the larger orange region showed the standard deviation, also across ten images, throughout the stack. The grey markers showed the precision of each z-stack position (the ratio of correctly automated labelled pixels to the total labelled pixels); the yellow markers showed the recall, also known as sensitivity, for each position (the ratio of correctly automated labelled pixels to the manually labelled pixels); the blue markers showed the specificity (correct automated unlabelled pixels as a ratio to manually unlabelled pixels), which were mostly hidden by the green markers, the accuracy (the ratio of correctly

automated labelled and unlabelled pixels to every pixel within the image). The F1 value, the harmonic mean average of precision and recall, was shown in navy.

Regions where the total matching pixels were highest, the first 100 z-positions and last 200 in the z-stack, averaged over 99.95%. However, the F1 value was lower than 0.5 for this region, and was therefore not visible on the graph. Due to the green speckling effect, the automated labelling of small areas which may or may not be unlabelled fibroblasts in the stack, these regions had proportionally high false positives, as there was little or no manually labelled fibroblast areas in these regions. The accuracy and specificity were close to 1 (perfect value) across the entire dataset and false positives occurred in less than 0.2% of all unlabelled pixels. Specificity and recall were closer in value towards the centre of the z-stack because the singularly labelled fibroblast dominated the SBFSEM stack and the likelihood of other unlabelled fibroblasts being present was reduced.

The recall (the measure of how well the network has automatically labelled the pixels that were labelled manually) had two regions of large variation. The first region was where the fibroblast started to appear within the stack and there were very few pixels to label, and the second was where the fibroblast stopped appearing and, once again, the labelled pixel regions were a lot smaller. The precision also deviated around these areas for the same reason. Some variation in automated and manually labelled regions was due to subjective boundaries in which areas of the z-position images can be defined as fibroblast or not. These boundary regions were up to roughly ten pixels in width around labelled areas and consequently, in areas of low quantity of labelled pixel, the percentage of corresponding pixels fluctuates and precision lowers. The smaller areas of variation, such as the dip between positions 300 and 400, were areas where the fibroblast labelling took up the highest proportion of image space compared to other regions in the z-stack. Therefore, there was a higher number of boundary regions and greater variability in the results. The standard deviation was noticeably greater in the region around position 600. A few z-positions contained fewer total matching pixels than average, including 590, which was subsequently analysed in more detail (Figure 37).

With one of the lowest amount of matching pixels between automated and manually labelled images, 99.5378%, position 590 had visibly different automated and manually labelled images, which was especially evident in the comparison image. While there were small flecks of green, as seen in the previous comparison images, the main discrepancy was a large green arc in the top right of the image. This represented an area where the network labelled an area

Labelling 3-Dimensional Placental Tissue

of tissue which was not manually labelled in the testing data. Further inspection on both the input (non-labelled) and manually labelled images showed that this was an area of fibroblast, and instead of being fully labelled was instead simply marked with a thin line. This was likely a marker for the labeller to label that section, which was then forgotten or otherwise left incomplete.

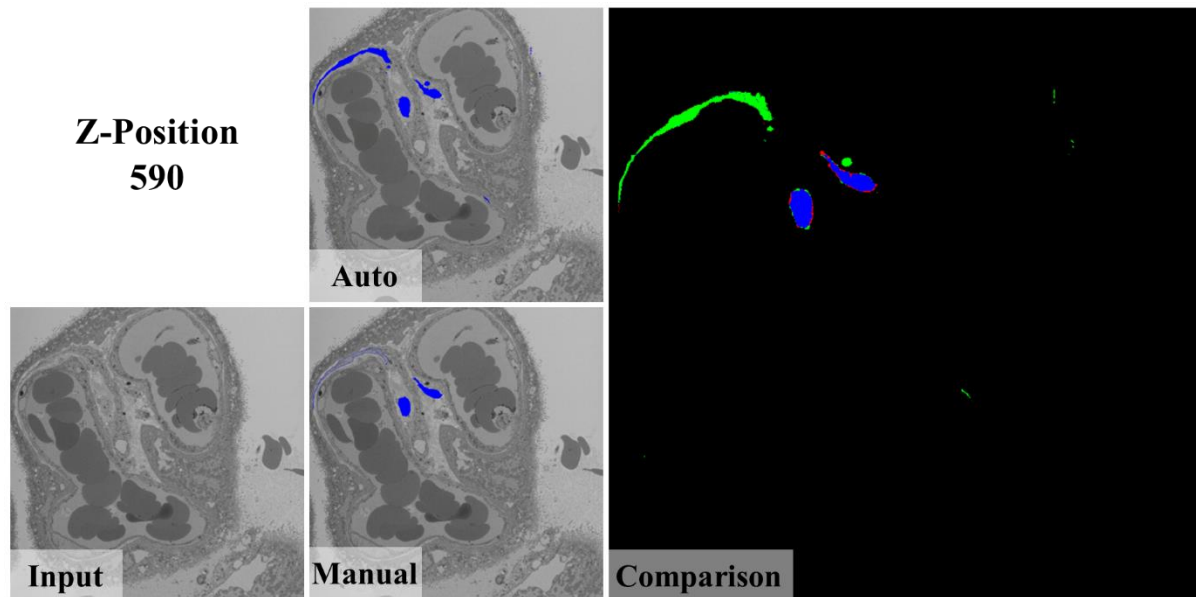


Figure 37: The extracted data from Z-position 590, with a visual comparison.

The unlabelled input SBFSEM image produces a network output of an auto-labelled SBFSEM, image for the corresponding z-position. This is then compared to the manually labelled position to create a visual comparison image, where there is a large section of green. This green is an area of “false positive” where the network labelled a region not labelled manually. Upon close inspection, this is a region which should have been labelled, and instead only a quick line was drawn (human error). However, smaller patches of green were areas of incorrect labelling at regions which contrasted with surrounding tissue.

Not only can the network be used for automated labelling of features, such as a fibroblast, within a SBFSEM z-stack, it could also be used for quality control purposes. Areas with the lowest correlation, or lower than a specified threshold in precision and F1 statistical markers (relevant to “false positives”), would be flagged for further visual inspection. This quality control could be used for time-efficient checking of either automated or manually labelled datasets, with multiple networks working in tandem but trained on varied SBFSEM data. Combining manual labelling with network quality control could also be applied in reverse, with manual quality control of network-labelled data. For example, when the automated labels of the previous z-stack were projected, the overall structure of the fibroblast was clear, yet there were small areas of labelling around the fibroblast that were not needed (likely due to the presence of unlabelled fibroblasts in the training data, as previously discussed). With some simple editing techniques on readily available software, such as outlier removal on ImageJ, the structures can be more clearly seen. However, detail can be lost during this

process, so a training dataset should be fully labelled to prevent the over-labelling seen in Figure 38.

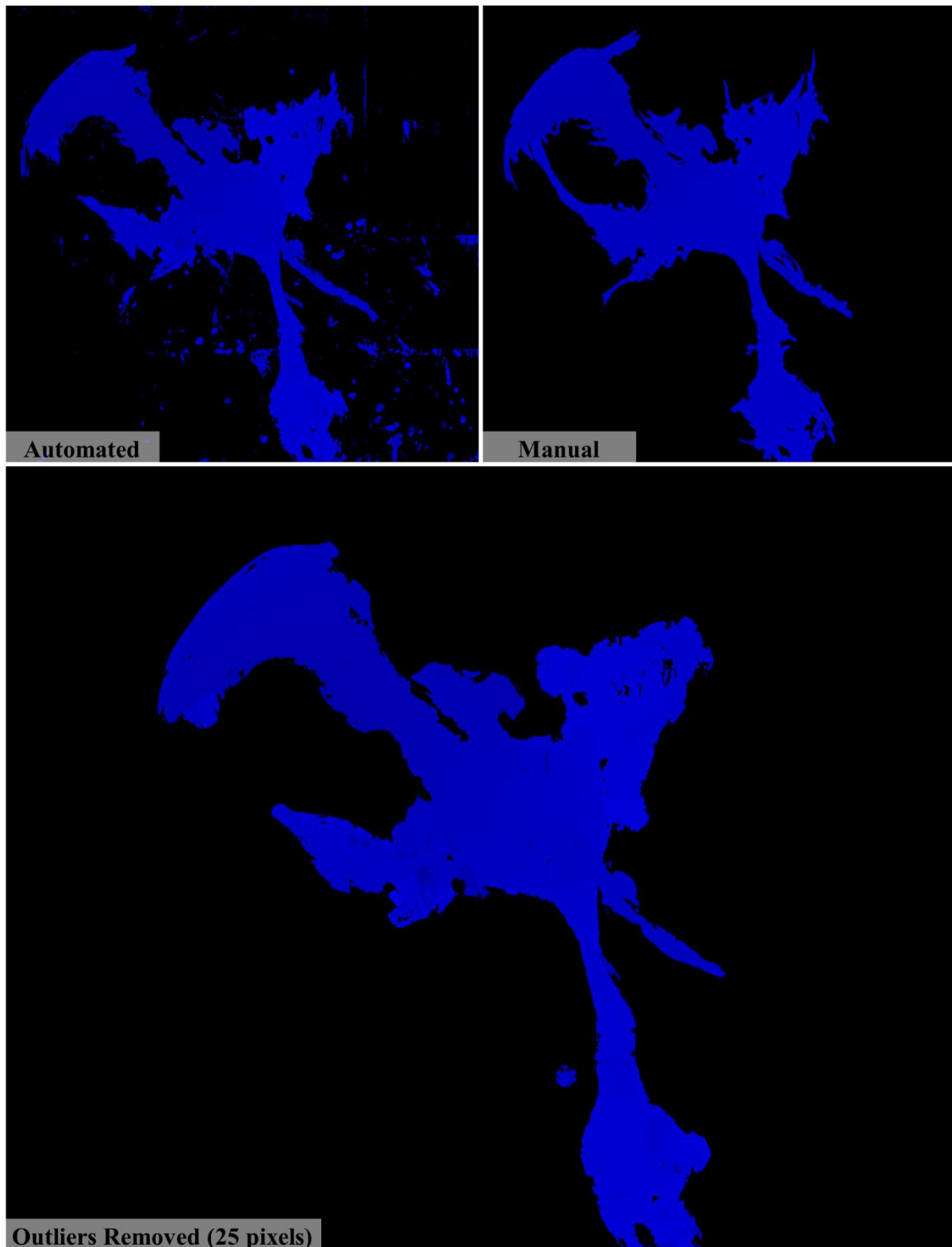


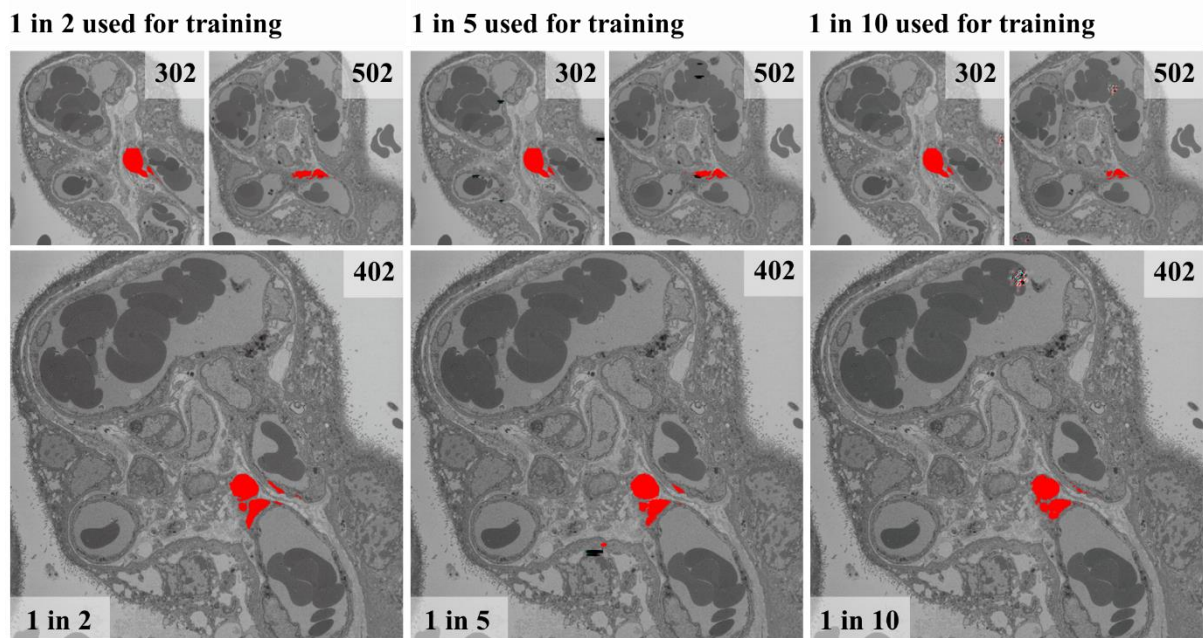
Figure 38: Comparison of a projected 3D structure of a fibroblast in SBFSEM imaged-placenta from both automated and manual labelling.

The automated structure, without any computerised data-cleaning, contains areas of poor detail and areas of over-labelling, giving a “messy” appearance. However, the overall structure is clearly visible and maintains a similar shape as the manually-produced 3D projection, as shown when outliers are removed using ImageJ.

5.4.2 REDUCING THE NETWORK REQUIREMENTS FOR PERICYTE LABELLING

As the same SBFSEM z-stack was used to train the network on pericytes as for fibroblasts, the label colour was altered to red, [255,0,0], to avoid any possible cross-contamination of the data. Rather than repeating the previous work with a fibroblast substituted for a pericyte, the capability of the network was tested, with respect to the amount of training data required for automated labelling to match manual labelling. Comparing the automated labelling output for each network (Figure 39), an optimal training amount for a single stack may be determined, balancing time needed for manual labelling (generating training data) with reliable network output.

Labelling Comparison for:



Input and Manual Labelling

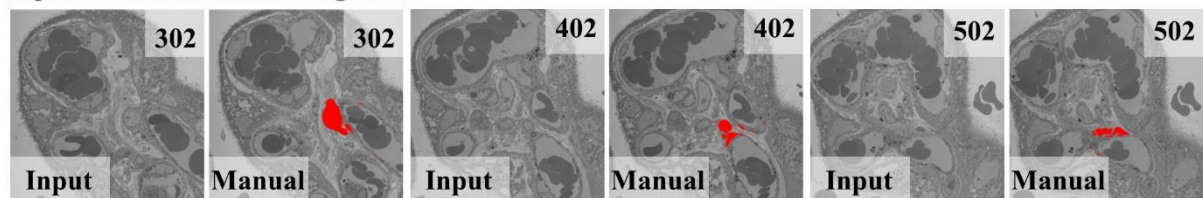


Figure 39: Comparison of automated and manually labelled SEM imaged placenta for three different z-stack positions. Each position, 302, 402 and 502, shows the input to the network (an unlabelled SEM image), a manually labelled image, and the automated labelling comparison of 1 in 2 images used for training data (the remaining used for testing), 1 in 5 used for training and 1 in 10 used for training. The position 402 is enlarged for closer comparison between the automated output for decreases in available training data. When 1 in 5 is used for training, there is more false-labelling and less detail, showing a lower confidence in the network output compared to 1 in 2. When 1 in 10 is used for training, there is less detail again, with added pixilation issues in the reproduction of the SEM image, seen in the top right of the image and repeated for multiple outputs (302 and 502 in different areas of the image), showing a breakdown in network performance.

As before, 1 in 2 images (50%) of the z-stack were used for training and the remaining data used for testing. Additionally, a separate network was trained on only 1 in 5 images and a final network was trained on 1 in 10 image, with remaining images used for testing.

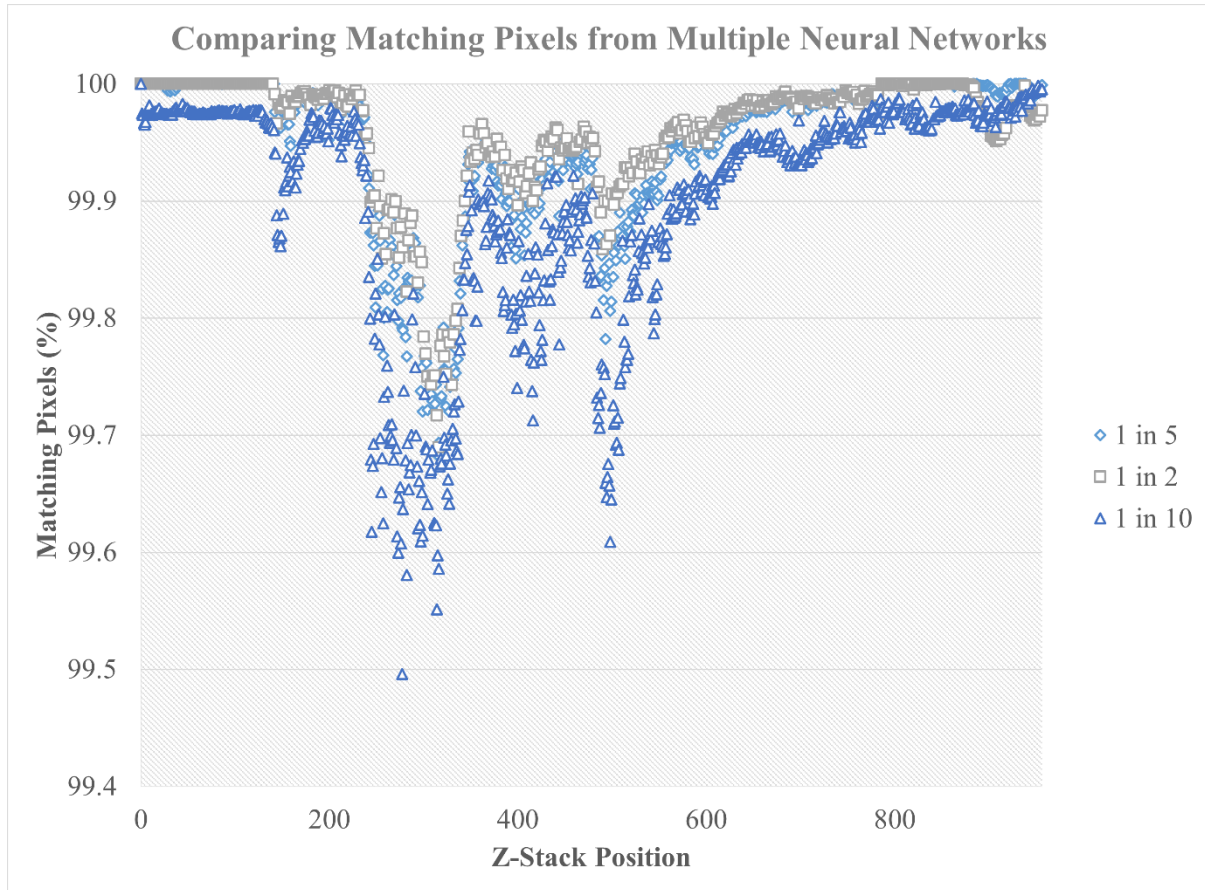


Figure 40: A graph comparing the total corresponding pixels for network output when 1 in 2, 1 in 5 and 1 in 10 images are used for training.

A corresponding pixel refers to a pixel with the same RGB values in both manual and network output images. Network output for 1 in 2 used for training is shown with grey square markers, 1 in 5 is shown with light blue diamond markers, very similar in value to 1 in 2, and 1 in 10 is shown with navy diamond markers, which are visibly lower in value.

The 1 in 2 network produced an output 99.96% identical to the manually labelled image (Figure 40). As there are relatively small differences between adjacent slices in the stack, the network was trained on data very similar to the testing data and anything above 99% identical was expected. However, the 1 in 5 network produced small black regions on the image which were not present on the input image. As these were not [255, 0, 0], the impact on labelling was negligible, although some loss of detail was evident. Overall, the output was still 99.95% identical. Differences between training data and testing data were more pronounced, here, so this was a very promising result. The 1 in 10 network has a greater drop in matching pixels, at 99.90%, which was due to both a small loss of detail and also the presence of small blocks of apparently randomised cyan, red and grey pixels, typically ranging in 10 to 20 pixels across.

Labelling 3-Dimensional Placental Tissue

Like the black regions in the 1 in 5 outputs, these were a likely sign of overfitting, where the networks were less confident producing outputs in testing data that had greater variation from training data (a consequence of less training data being used). By adding multiple stacks into the training data, overfitting in the 1 in 5 and 1 in 10 networks could be reduced.

Labelled output from the 1 in 5 network, shown as light blue diamonds, was almost indistinguishable from the 1 in 2 network, grey square markers, throughout the whole z-stack, averaging only 0.01% difference. As there was only one dominant pericyte, the uncertainty present in the fibroblast labelling network was not present, and there were areas in both the 1 in 2 and 1 in 5 network of 100% identical output. The 1 in 10 network output, navy triangle markers, was different for each position and drops below 99.5% at an area of greatest labelling surface area. It was also unable to manage 100% in areas with little labelling, due to the previously mentioned areas of random pixel generation, which included label-red [255, 0, 0].

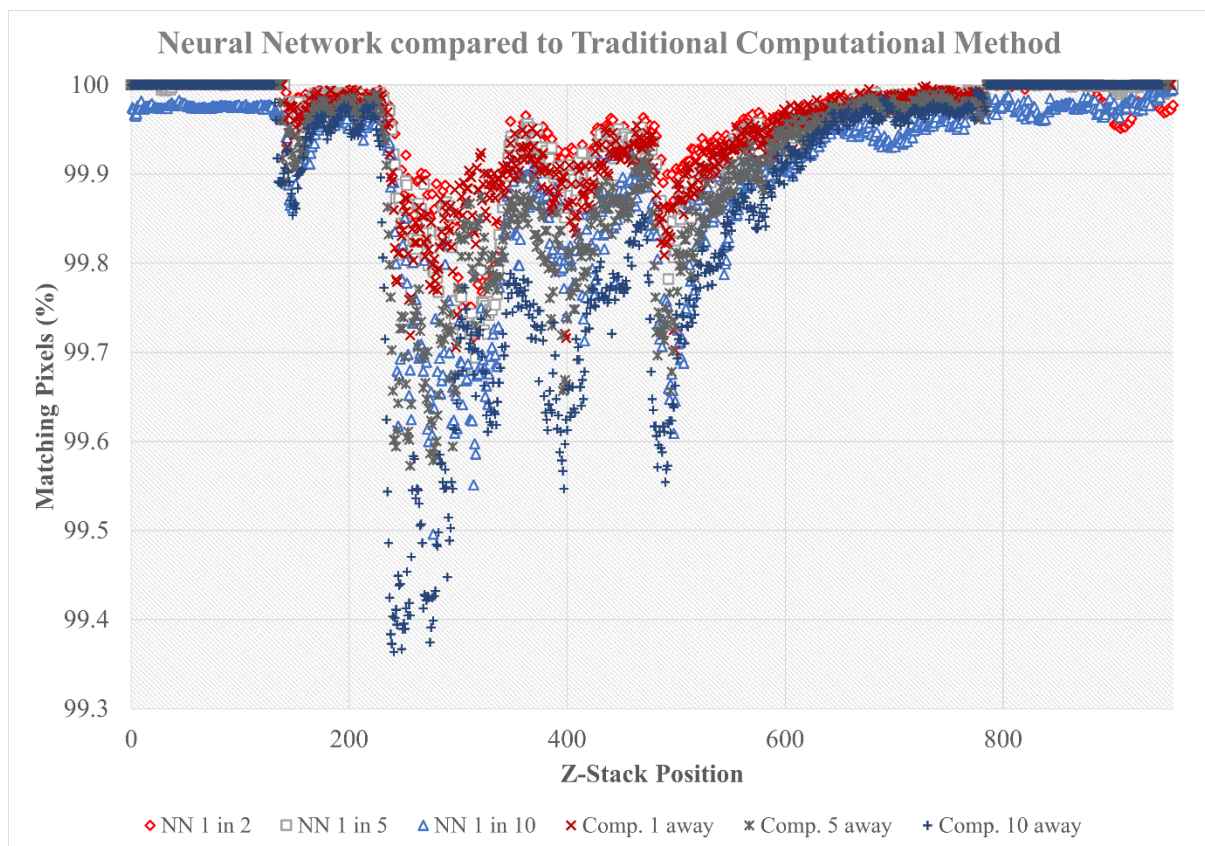


Figure 41: A graph comparing the matching pixels from network output when 1 in 2, 1 in 5 and 1 in 10 images are used for training with a traditional computational method when an image 1 position away, 5 positions away, and 10 positions away is used.

A matching pixel refers to a pixel with the same RGB values in both manual and network output images. Network output for 1 in 2 used for training is shown with red diamond markers, 1 in 5 is shown with light grey square marker and 1 in 10 is shown with blue triangle marker. The traditional computational output for 1 image away is shown with a dark red x-cross, 5 images away is shown with a dark grey star-cross, and 10 images away shown with a dark blue cross.

These results were then compared to a traditional computational method of labelling that did not utilise a neural network (Figure 41). This method relied on image slices varying little from adjacent image slices to form a naïve baseline. When the matching pixels for each image was averaged over the entire stack, as stated previously, the 1 in 2 network produced 99.96% matching pixels, 1 in 5 network produced 99.95% matching pixels and 1 in 10 produced 99.90% matching pixels. The traditional method 1 image away produced 99.95% matching pixels, 5 images away produced 99.92% matching pixels, and 10 images away produced 99.87% matching pixels.

When the general positioning of the traditional computational approach (the crosses) is graphically compared to the neural network approach to labelling (the markers), the crosses are visually lower than their respective marker counterparts. The only area where this is not the case is for areas with no pericyte to label, at both the beginning and the end of the stack. For areas where a pericyte is present, the neural network provided more pixels that matched manual labelling than the traditional computational approach.

Looking at individual z-positions, the lowest 1 in 10 network labelled position was z-position 282, with 99.58% matching pixels. The traditional computational method only produced 99.50% matching pixels for this z-position. The lowest traditional computational method produced 99.36% matching pixels while the network output was at least 99.89%.

Across the entire stack, but especially for areas with pericytes, and for individual z-positions that resulted in less matching pixels across both approaches, the neural network approach provided a higher percentage of matching pixels. This improvement became more evident when less manually labelled images were provided for either training or referencing.

5.4.3 MULTI-CELLULAR LABELLING WITH ENDOTHELIAL CELLS

The next step after labelling a single cell was to label multi-cellular structures, to determine whether a network was capable of recognising more complicated cellular systems. The method of training a network for labelling multiple cells, such as endothelial cells, was similar to the method for labelling single cells. However, labelling larger structures required the use of a W-Net, a series of U-Nets, with 33 layers rather than 17 (Figure 42). Unlike increases to filter number or other data-dimension changing parameter modifications, increases to the number of layers did not require an increase in GPU size requirements. The increase in complexity did increase the time needed for training, with 50 epochs taking approximately 24 hours, compared to roughly 6 hours for the previous U-Nets.

Labelling 3-Dimensional Placental Tissue

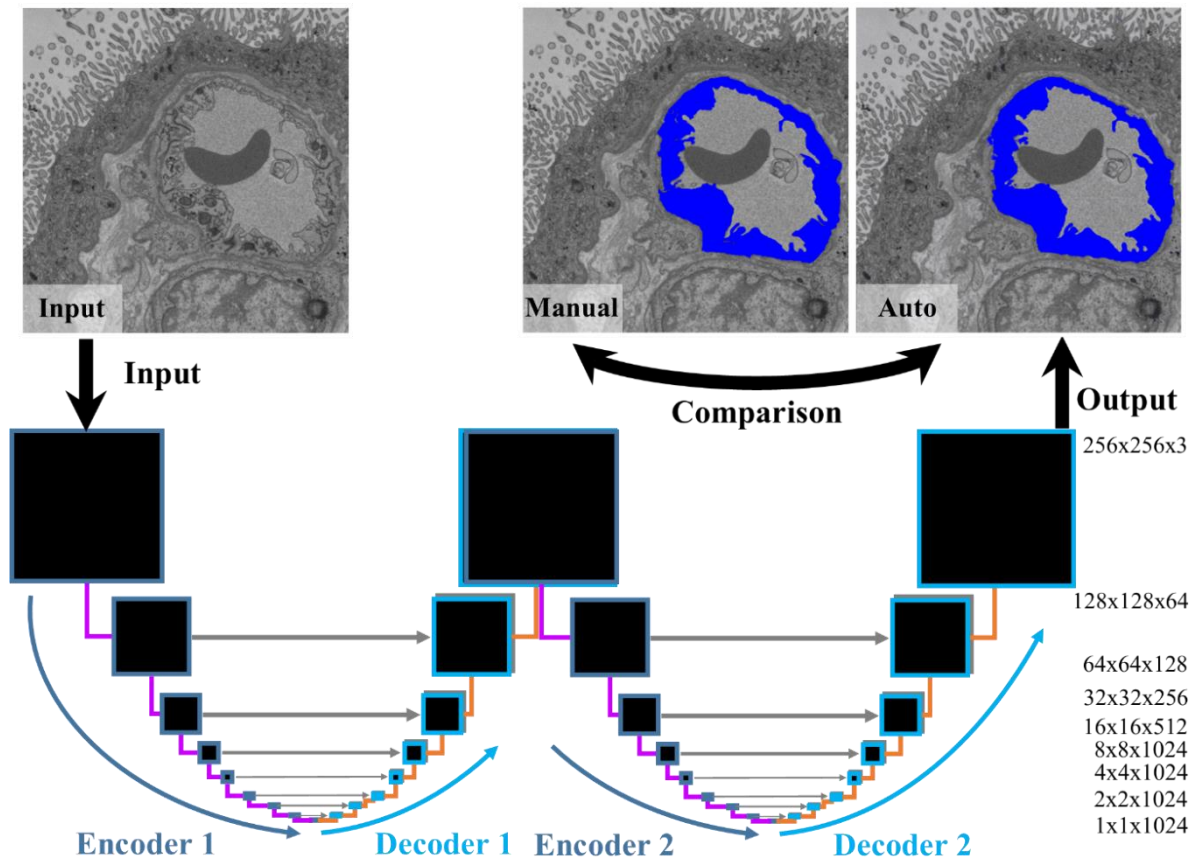


Figure 42: Illustration of the W-Net architecture.

An unlabelled image is input into a neural network consisting of many convolutional layers (pink), deconvolution or up-convolution layers (orange) and skip-layers between the encoder-decoder layers (grey). Output images are compared to manually labelled images, and differences lead to changes and optimisation of the network layers through backpropagation. (Additional error correction is performed from a separate convolutional neural network to determine how realistic the output images are in comparison to manually labelled target images.)

A different SBFSEM stack was used for endothelial cells than for previous work to confirm the adaptability of the network labelling method, and therefore the usability and applicability. Even though as little as 10% of the stack produced network output above 99% pixel match, 50% was used for training for continuity between each labelled feature. However, to further test the network, an additional 10% block of data was extracted from the centre of the stack, along with a block at the end of the stack, which was left unlabelled. These deliberate “blind-spots” were used to determine how well the network could adapt to higher variety inputs without causing the overfitting seen in previous work, as the amount of training data remained adequate for training.

Visual analysis was conducted at z-positions 10, 110 and 210, and showed similar performance to previous network models (Figure 43). Comparison images showed the network was capable of matching the vast majority of pixels to manually labelled images, shown by the blue and black pixels. Compared to the single cell structures, less detail in the

boundary region appeared lost in the automation process, likely due to both the increase in layers (allowing for greater complexity in network-generated output) and the clear boundary of a blood vessel compared to the more complex regions of placental tissue. However, the boundary regions were still subjective up to 10 pixels in width, and close examination of the comparison images showed the location of the red and green pixels, where the network and manual labelling diverge, were virtually all around endothelial sections. While there were frames with the green speckling seen in fibroblast labelling, these were the minority, and only in areas close to training data blind-spots.

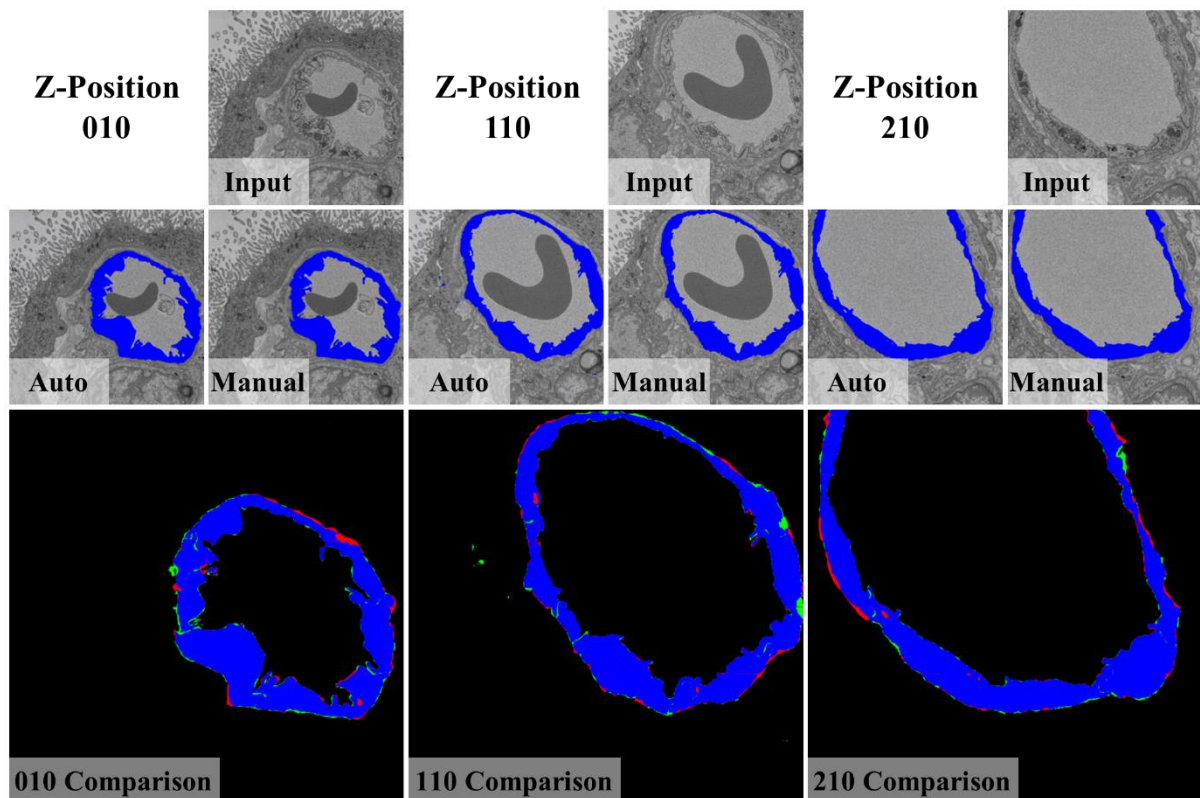


Figure 43: Comparison of automated and manually labelled SEM imaged placenta for three different z-stack positions. Each position, 010, 110 and 210, shows the input to the network (an unlabelled SEM image), automatic network labelled output and manual labelled image, and the comparison image. The comparison images contain four different colours. Black pixels are a position of a true negative, where the automated and manual images show no labelling. Blue pixels are positions of true positive, where both show labelling. Red pixels are positions of false negatives, where there is manual labelling but no automated labelling. Green pixels are positions of false positives, where there is automated labelling but no manual labelling.

When each z-position output throughout the SBFSEM stack was analysed, statistical values were plotted using the left axis and total corresponding pixels (percentage of pixels which were identical in automated and manual labelling) using the right vertical axis (Figure 44). Precision (light blue markers), a fraction of how much of the automated labelling was relevant, whether the automated labelling was also manually labelled, was greater than 0.9 on average throughout the whole stack. Recall (orange markers), a fraction of relevant labelling

Labelling 3-Dimensional Placental Tissue

to all automated labelling, was also greater than 0.9 on average. F1, the mean average of precision and recall, showed that the vast majority of areas labelled by the network were areas that had been labelled manually. In other words, endothelial cells were labelled by the network without also labelling undesired features. Specificity (grey marker), a fraction of areas not labelled manually nor by the network over areas that were only labelled by the network, was above 0.95 across the entire stack. This indicated that there were very little areas lacking automated labelling that were not also left unlabelled manually, and therefore the network did not miss areas of endothelial cells. Accuracy (yellow marker), the fraction of matching labelled and unlabelled areas over the entire output, was similar to specificity, as the area of endothelial cells was relatively small in comparison to the size of the entire image. Consequently, the network ability to correctly identify areas which did not require labelling held greater weight.

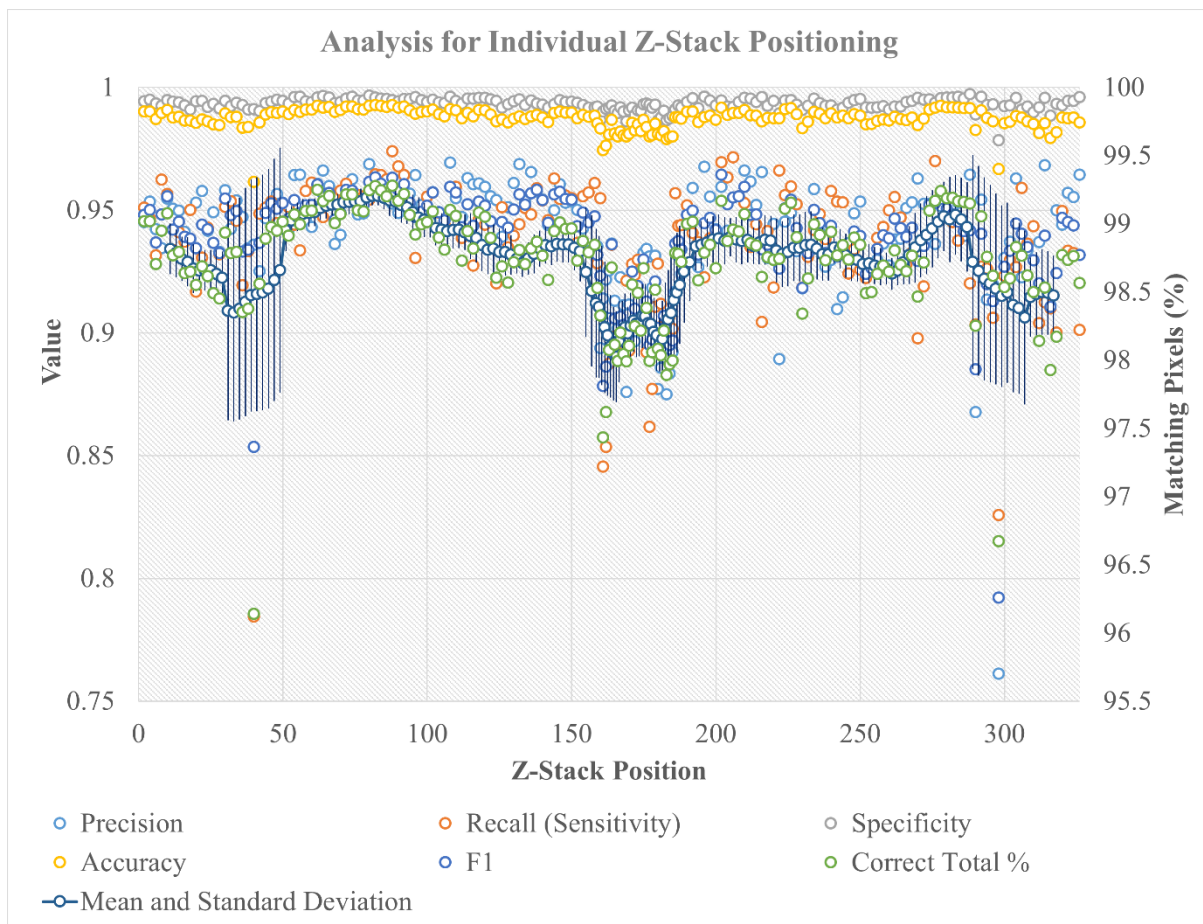


Figure 44: Graph of multiple statistical analysis factors for comparing automated to manually labelled SBFSEM placenta. Left axis of 0.75-1.0 is for precision (light blue circles), accuracy (yellow circles), recall (orange circles), F1 (blue circles), and specificity (grey circles, almost hidden behind accuracy). Right axis of 95.5-100.0% is for Correct Total % (green circles) and the accompanying rolling mean and standard deviation of these points (navy). A pixel is considered correct, a true positive or true negative, when automated labelling matches manual labelling.

The “Correct Total %” (green marker) was the percentage of pixels that were identical in both automated and manual images, here assumed to be correct, with a rolling mean and standard deviation in navy. This followed the same trend as seen with previous statistical values, with a dip below average (98.7%) between positions 150 and 200. This corresponded to the network blind-spot where no training data was provided in this region. These values were examined in more detail on a separate graph, to more easily differentiate between percentage and statistical values (Figure 45).

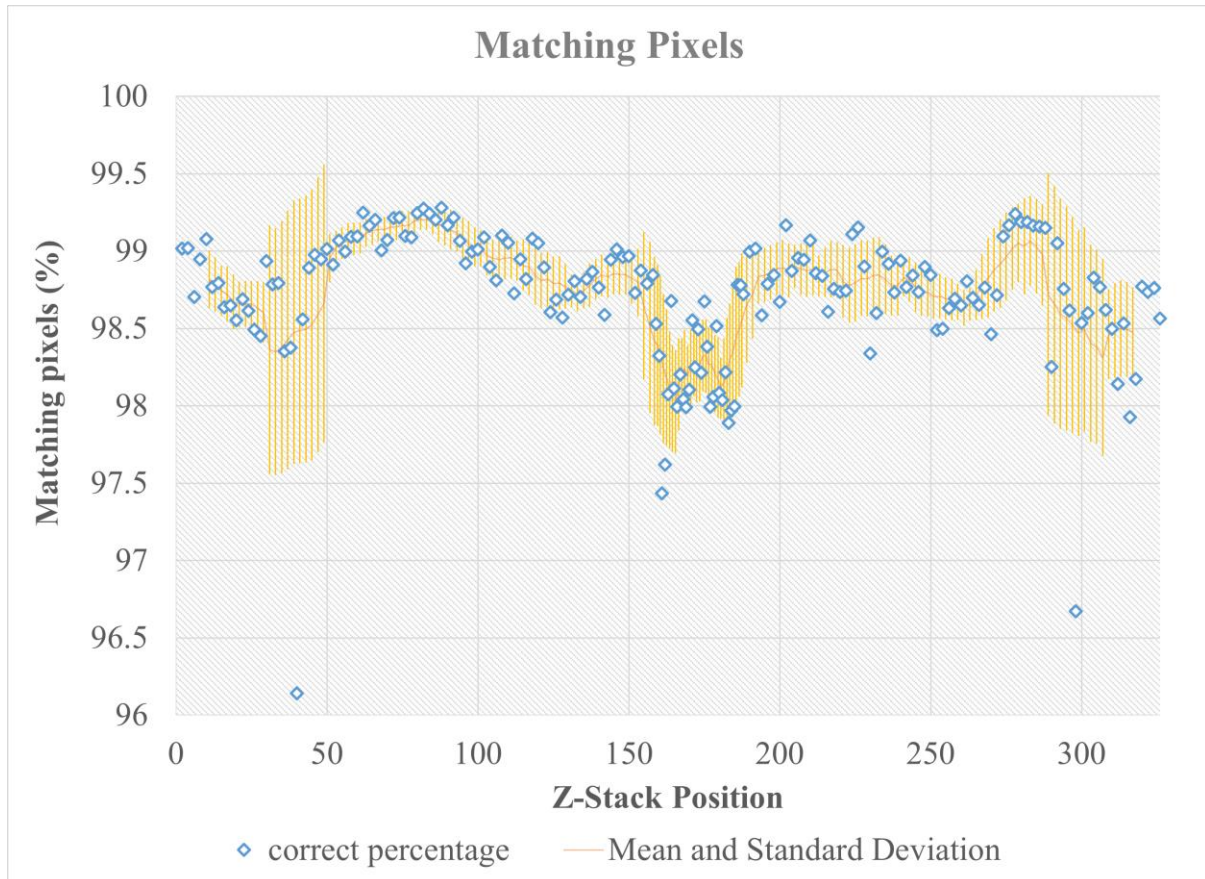


Figure 45: In-depth graph of corresponding pixels across the z-stack of SBFSEM imaged placenta. A corresponding pixel refers to a pixel with the same RGB values in both manual and network output images, and is marked with blue diamonds. The rolling mean and standard deviation of corresponding pixels are marked in yellow. If a network-generated label and manual label produce the same corresponding pixel value, the pixel is considered correct.

The percentage of pixels considered correct (blue diamond marker), where the network and manual images matched, and the rolling mean and standard deviation (yellow), fluctuated in several positions along the z-stack. While the average of 98.7% was lower than for single cell labelling, the fraction of labelled structures was higher, and endothelial cells were present throughout the whole stack. The standard deviation was only 0.4, which showed very low variation of corresponding pixels across the z-stack, but there were two areas of high standard deviation. In both cases, one at the start of the stack and one at the end, the increase in

Labelling 3-Dimensional Placental Tissue

standard deviation was due to a single z-position, 40 and 298 respectively, which were analysed in more detail (Figure 46). However, the central blind-spot did not have a significant increase in standard deviation, which showed the stability of network performance over the region. Even though the centre of the region could have been the worst performing space, as it was furthest from training data, performance did not continue to drop, and some of the central positions were the best performing in that region. This was a very promising result, it shows that larger blind spots could be created as long as there is enough variability in the training data to not adversely impact performance. Images over 15 positions away from the last seen image were labelled as well as images 5 positions away, therefore the network could be capable of successfully labelling stacks with fewer than 1 in 15 images manually labelled, provided overfitting could be avoided.

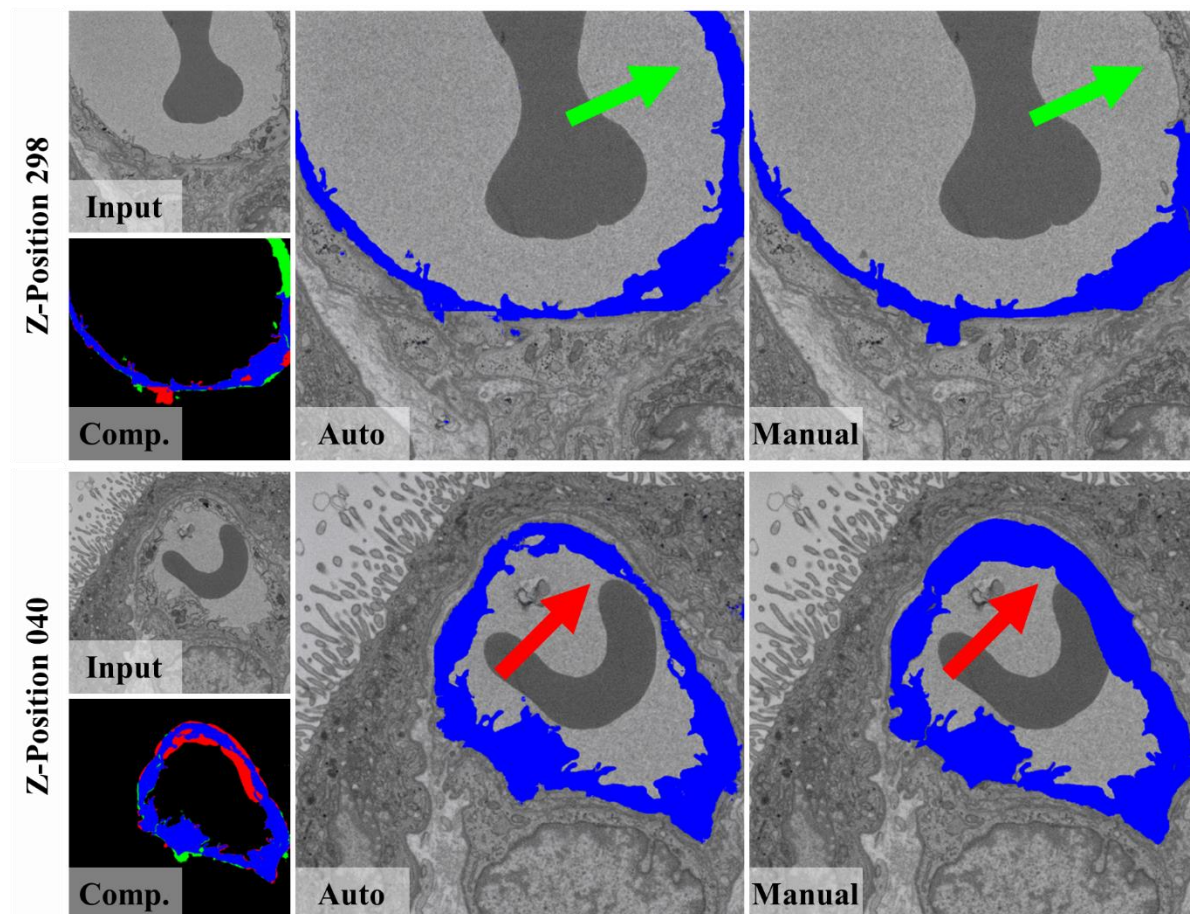


Figure 46: The extracted data from Z-positions 40 and 298.

The unlabelled input SBFSEM image produces a network output of an auto-labelled SBFSEM, image for the corresponding z-position. This is then compared to the manually labelled position to create a visual comparison image, where there is a large section of green for position 40 and red for position 298. These areas of difference are shown on the auto and manual images with green and red arrows. The green is an area of “false positive” where the network labelled a region not labelled manually. Upon close inspection, this is a region which should have been labelled, and was incorrectly left unfinished. Similarly, the area of red on position 298 also shows human error. While red shows a “false negative” where the network failed to label a region labelled manually, the manual labelling was an inaccurate over-labelled region. For both positions, the network was superior to manual labelling.

The two positions of below 97% matching pixels, 40 and 298, were examined in more detail, and each position contained a different abnormality. Position 298 contained less labelling in the network output than was used in the manual labelling, shown by large red arrows. This was not a mistake by the network, as the difference was due to poor labelling by the manual specialist. For this z-position, the detail was lost and the manual label was an oversimplification, which contained a large surface area where no endothelial cells were present. In contrast, position 40 had a section of apparent over-labelling in the top right corner, shown by green arrows. However, once again difference between network and manual labelling was due to mistakes in the manual labelling process, where an endothelial cell was incorrectly left unlabelled. The network did not miss this endothelial cell, which resulted in a more correct output but a large drop in total matching pixels. While both outputs, positions 40 and 298, had small areas of other labelling differences, the vast majority of differences were due to incorrect manual labelling, which showed once again the potential for networks to be used both for automated labelling and for quality control purposes.

5.4.4 LABELLING THE UNKNOWN

For widespread application and real-world usage, a network should be able to label unseen SBFSEM stacks (stacks where no z-positions have been manually labelled for training purposes). Additional augmentation of training data with random changes to contrast and brightness were used to increase variation in the training data and limit overfitting. The testing of this ability was split into two sections, with one input part of the same placenta as used in training, using the same imaging conditions, 40 z-positions away from the previous SBFSEM stack. The other input was from a different SBFSEM stack, with possible differences in imaging conditions, such as magnification and contrast. Output from both testing sections was then visually analysed (Figure 47).

The first tested image had an input from an unseen area of the same imaged placenta used in training, 40 images away from the last image used in training the network. A realistic appearing output image was produced, and an overlay of translucent labelling showed the endothelial cells were labelled with very little labelling of undesired features. The second tested image produced a similarly well labelled image, with the endothelial cells labelled and even less additional speckling/over-labelling than for the first tested image. This was a very successful result, as this stack was different from the training data, and yet the network did produce abnormalities shown from overfitting in previous experimentation with varied training and testing data (Figure 39).

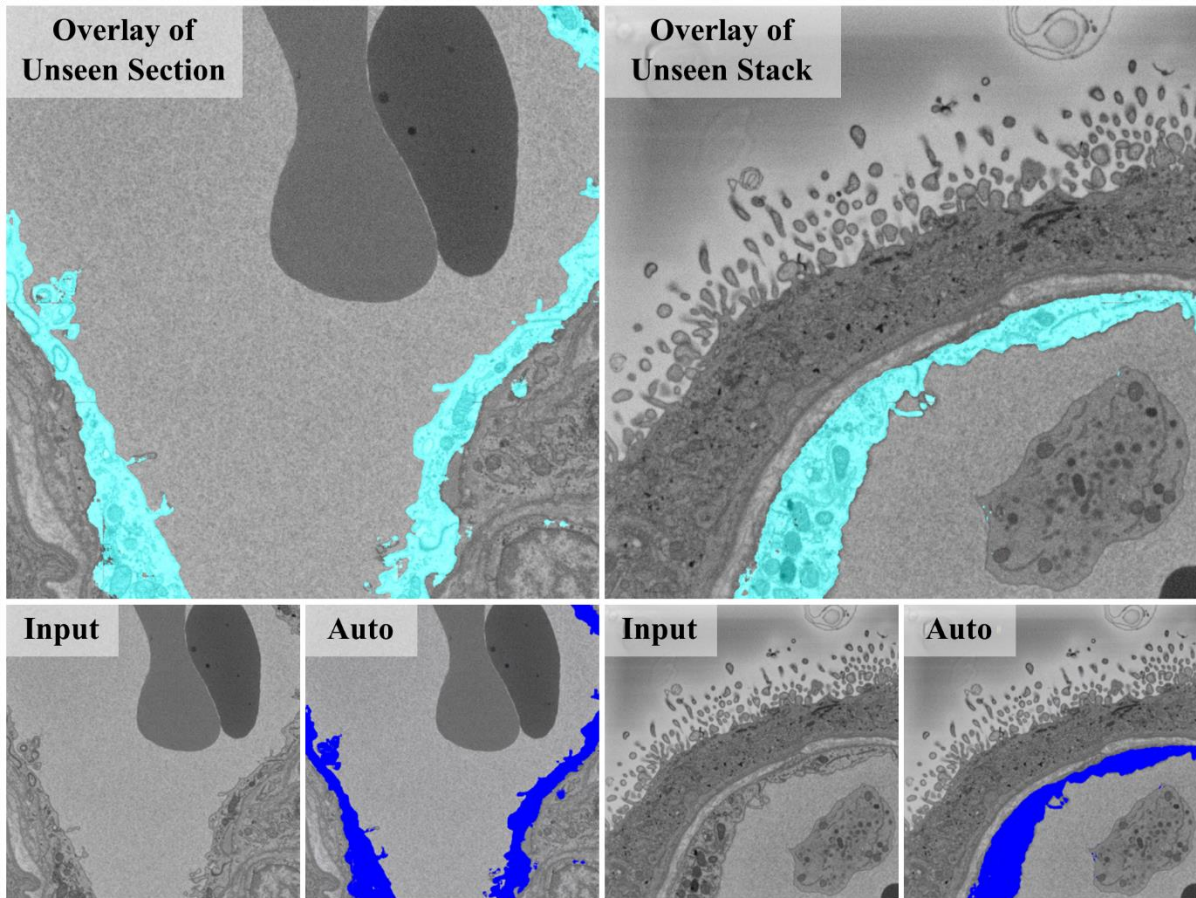


Figure 47: Visual evaluation of unlabelled regions and unlabelled stacks not used in training the network. The unseen section is over 40 positions away from the last section used for training. The unseen stack is part of a SBFSEM stack completely “unseen” by the network, with different magnification and imaging parameters. The input image is the unlabelled SBFSEM image, and the auto-labelled image is shown with the endothelial cell labelled blue. The overlay images show the labelling as a transparent cyan, so the detail of the labelling can be evaluated.

Automated labelling of both an unseen region and unseen stack without any manual labelling showed that the network method is capable of labelling new and unseen 3D image data while being trained on a separate manually labelled sample, promising for real-world usage. Future inclusions of additional image enhancement, such as super-resolution techniques, could be combined with this neural network labelling approach to both improve similarity of automated and manual labelling, and also to increase the range of detected features [231].

It is worth noting that the traditional computational method applied to pericyte labelling can not be reliably applied to an unseen stack with different imaging conditions.

5.5 SUMMARY

SBFSEM imaging can be an important tool in analysing nanostructures of cells within placenta in 3D, and visualising how 3D features interact, which is important for

understanding and modelling placentae. However, although novel structures and understanding has been revealed through this technique, SBFSEM imaging is slow, as determining i.e. the spatial relationship between two cells can take months of dedicated manual labelling of several thousand 2D SEM images. Development of a DNN approach to automated labelling removes this restriction to future computational modelling of the placenta, as network labelling significantly speeds up the labelling process (from weeks/months to seconds/minutes for a full SBFSEM stack). This in turn enables more quantitative analytical approaches as subjective labelling at boundary regions is limited.

The first two aims for this project were all achieved successfully, as the average percentage of matching pixels between manually labelled and DNN-labelled images was above 99% (see Figure 36 and Figure 40). The most time-intensive part of using a network for automated labelling is the time spent training, which took between 6 and 24 hours, yet testing an image took less than a second. Even when training time is included, the process is substantially faster than a month of dedicated time manually labelling individual stacks. Image resolution is only limited by the size of available GPU, and cropping images can overcome such restrictions, so there is potentially no data loss using networks for labelling. These networks not only identified singular cells, but also distinguished specific cell types and labelled large cellular structures with an average error less than 2%. The third aim also included evaluating labelling of a completely unseen stack, which was successfully achieved. As all aims were successfully completed, the hypothesis was found true and networks trained on manually labelled SBFSEM sections were able to identify and label specific cell types.

However, this approach is not limited to specific cells, as other cell and tissue types aside from placenta could be labelled using the same techniques applied here (such as osteoblasts within bone), where a network is trained on stacks of unlabelled and their associated labelled images with a desired feature extracted. Similarly, as the networks were given no additional information beyond the image data on SBFSEM imaging, other 2D and 3D imaging techniques could have processing and feature extracted automated using DNNs. This ability to identify and label a range of featured cells, such as fibroblasts, pericytes and endothelial cells, shows that a network is capable of extracting cellular (and therefore acellular) data from an image input, which will be necessary to determine cell response to biophysical and biochemical stimuli in further work.

6 MODELLING THE CELL RESPONSE TO SURFACE TOPOGRAPHIES

The research covered in this chapter was published in *Tissue and Cell* [11]. Application of AI in regenerative medicine was reviewed in *Biomedical Physics and Engineering Express* [40].

After networks have been shown capable of identifying single and multiple cells in a single image, as shown in 5 Labelling 3-Dimensional Placental Tissue, the challenge advances into whether a network is capable of predicting cell behaviour. While this is a substantially more complex task than the identification task, by modelling stem cells computationally, the previously mentioned problem of scope (section 2.1.2) can be overcome, and those without access to the currently necessary biomedical research facilities can perform simulations similar to experimentation, which was previously not possible outside specialist laboratories, once networks are suitably trained.

6.1 INTRODUCTION

There is no substitution for laboratory-based experimentation. However, computer-based experimentation and statistical models have been used for decades alongside real-world data successfully. For example, scientists who advise leaders in UK government were able to create multiple case-count predictions during the COVID-19 pandemic by combining data from different past outbreaks and modern experimental data to predict what would happen in the future, with updated, real-time, data used to refine the model further [232]. Another example is the compulsory use of flight-simulations for the continual training and development of commercial pilots, who spend a collective 4.6 million hours in Full Flight Simulators every year [233]. Pilots can test responses to statistically unlikely phenomena, and try new or untested manoeuvres without risking lives or billions of pounds of equipment. Just as with flight simulations, a model for stem cell behaviour would allow for repeat experimentation without the required expenditure of traditional experimentation, which would allow for a more detailed range of features, including those initially viewed as less promising, to be examined.

Through harnessing savings in money and time, a network model has the potential to increase the quantity of research being produced by the biomedical field. This is due to both increases in efficiency for academics and clinicians already carrying out experimentation, and also from the ease of access for potential researchers who are currently unable to carry out laboratory-based stem cell investigation, either as a consequence of lack of tissue samples,

equipment, funding, or geographical/physical limitations. Increasing research in the field will lead to improved scientific understanding, an improved collection of more relevant data, and therefore an improved network model. This cycle, seen in Figure 48, leads to continual improvement, even without accounting for the predicted increases in network capabilities in the upcoming decades. However, network models should not replace traditional experimentation entirely, as novel models require experimentally-obtained data for verification, training, and predicted outputs from each model need vigorous testing against real-world data to determine applicability, scope, and validity. Additionally, new training data should be added to the network frequently, so models can be continually updated, refined, and the model does not become limited in scope or become obsolete.

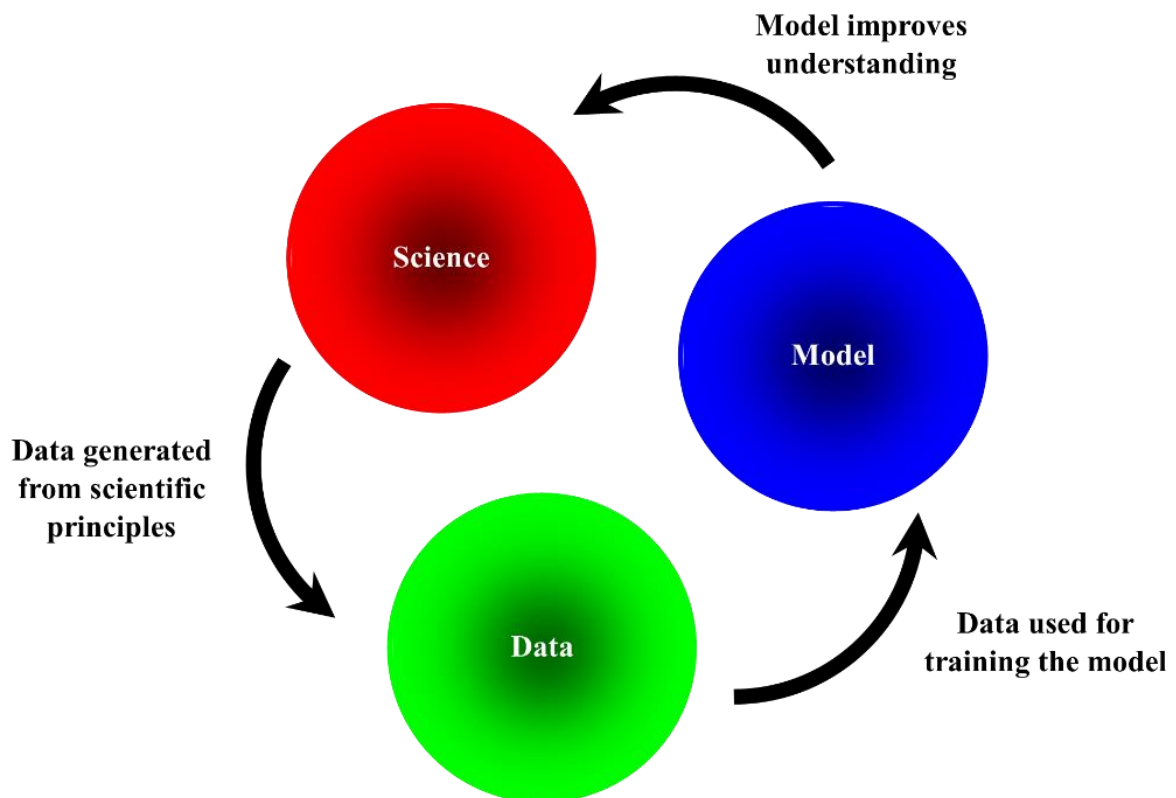


Figure 48: Circular nature of improvement in science, data, and AI models.

Science improves data acquisition through targeted areas of interest and variables based on scientific principles and understanding. Data improves AI models, such as deep neural networks (DNNs), because network design can be optimised to the data, and the outputs of the model depend on data quality during training. Models can be used to improve scientific understanding, as they are capable of generating statistically probabilistic results to scenarios much faster than is possible in traditional experimentation, including for variables which could not be investigated ethically and/or physically in real world experimentation. As one section improves, the whole system improves: Improved scientific understanding leads to improved data acquisition leads to improved AI models and, ultimately, to improved scientific understanding.

The network model created in this chapter was designed to save both time and money through the prediction of stem cell behaviour, specifically, skeletal stem cell behaviour, as a result of microtopographical cues, with the aim of accurately predicting stem cell position,

alignment, and visual morphology for unseen topographies. Topography is of growing importance to biomaterial design, as cell behaviour is influenced through the topographical engineering of surfaces and the use of surface-directed biotechnologies, allowing for improved osseointegration and tissue regeneration [9, 12, 16, 30, 49, 71, 72, 74, 75, 78, 80-82, 86, 87, 234]. Time saving is vital, as fully-investigating four topographies with on four materials (leading to 108 unique combinations) would take longer than two working weeks. An increase to ten topographies with a hundred variations on ten materials and subsequent investigation of each unique combination would take 400 days. Combining these two examples into even more unique combinations could take over 10,000 years to investigate. Instead of exploring each unique combination of features experimentally, which would not be practical, a network model can create a predicted response in a fraction of a second, without the necessity of topography creation, cell extraction, cell culture, and cell imaging. Through determining optimal topographies computationally, biomaterial design can advance at a faster pace and the current cost of associated research will reduce, due to less time and fewer material requirements.

The cell response to topographies is complex, with different cells displaying a myriad of different responses to the multiple topographical cues [30]. While research into cell-topography research can be traced back over many decades, there remains a lack of predictive capability as to how a specific cell will respond to a specific topographical environment, although some trends have been discovered in recent years [60, 85]. Some of this work will be summarised in section 2.2, and the challenges behind generation of cell-instructive topographies briefly explored in section 2.3. Through generation of laser-machined microtopographies (structures $\sim 10\ \mu\text{m}$ in scale), the response of skeletal stem cells to various microtopographical patterning can be used to train a network model, which then requires a series of tests to prove such a model is reliable and capable of real-world application. This naturally divides the project into two sections, one (section 6.3) focussing on collecting data sufficient for training a model, and training a network with a suitable architecture to utilise this data, and the other section (section 6.4) focussing on evaluating and testing the trained model. The summary (section 6.5) will focus on both the success of the network model, and how this technology can be advanced in future work.

6.2 HYPOTHESIS, AIMS AND OBJECTIVES

Through use of laser-machining, varied topographies can be created, which can result in a range of skeletal stem cell behaviour. These behaviours are difficult to predict, as there are many combined biophysical cues, and biochemical cues, available to the cells, and each individual and combined cue can lead to different cell responses. A neural network may have the ability to model cell response when input topographical data because DNNs are successful approximators, and therefore would not require detailed theories behind all cue-response relationships. Instead, a DNN could predict a statistically likely cell response, if provided with a suitable training dataset.

6.2.1 HYPOTHESIS

A model can be developed with a dataset of binary and fluorescent cell images to simulate cell behaviour.

6.2.2 AIMS AND OBJECTIVES

There are three aims in this chapter and each aim consists of multiple objectives:

- 1) Develop a model capable of producing realistic-appearing fluorescent images of cells.
 - a) Create a small but diverse dataset from experimental images. This will consist of multiple different topographies and cell densities after multiple days of cell culture.
 - i) Laser-machine topographies onto a substrate suitable for cell culture and imaging.
 - ii) Seed cells onto the laser-machined substrates and culture for up to 30 days.
 - iii) Image the cells using brightfield and fluorescent microscopy at a combination of day 0, 1, 8 and 30.
 - b) Augment the images to maximise the variability of the dataset, through use of standard augmentation practices, such as rotation of the images.
 - c) Optimise a DNN for image-to-image transformation, so output images are believable without resulting in overfitting of the network.
 - d) Determine whether the images are realistic-appearing for multiple different inputs via visual comparison to experimental images.
- 2) Generate a model capable of producing likely cell responses, based on comparing cell positioning in model-generated and target images.
 - a) Use testing data that contains experimentally-obtained outputs, allowing for mathematical comparison between generated and target images.

- b) Calculate the probability of the simulated cell positions matching experimental cell positions.
- 3) Validate the model through determining a minimum line separation for stem cell alignment using model-generated data and comparing to experimentally-obtained data.
 - a) Determine a minimum line separation for stem cell alignment to link cell behaviour and topography.
 - b) Verify through comparison to experimental data. This data will be collected before the variable is derived to protect against bias.

6.3 METHOD

The collection of data is explained in section 4.2. This method section will explain how the data is applied.

6.3.1 TRAINING THE NEURAL NETWORK

In Figure 24, deep learning (the blue arrow) bypasses the lengthier experimentation involved in laser-machining samples and cell culture and imaging. This is a simplified illustration to depict how a DNN can be used as a model to investigate cell response to microscopic topographies. It was previously stated that 203 images were used to train the network. However, deep learning requires a large quantity of data, in the magnitude of 10,000 images or greater, to deliver accurate results [235-240]. Therefore, the 203 images were multiplied into 36,540 unique training images through standard data augmentation techniques of randomised cropping, followed by rotation, an approach been used successfully in machine learning for many years [241]. Augmentation of a dataset creates a more comprehensive set of possible data points due to the increase in variability of the data, without having to collect additional experimental data, which minimises the differences between the training and testing datasets (as well as further unseen data input during future application). This not only helps prevent overfitting, a problem encountered and overcome in chapter 5, but also maximises the amount of information extracted from the dataset by the network.

The input to the network consisted of three data channels: (a) a topography channel corresponding to the corresponding laser-machined pattern; (b) a time channel corresponding to the time point at which the output image was taken (i.e. day 0, 1, 8 or 30); and (c) a cell density channel, which was a randomised array corresponding to the brightness of the original, unprocessed fluorescent image and therefore an approximation of the cell density (Figure 49). Creating the topography channel was explained in section 4.2.1, and both the

time channel and cell density channel also resulted from automated code. The maximum length of cell culture achieved before full confluence was 30 days in both training and testing datasets, which could then corresponded to the maximum pixel value of 255. As dividing 255 by 30 days would result in a value of 8.5 per day, the ration was altered to 8 per day so that whole-integer pixel values were preserved. Therefore day 0, 1, 8 and 30 would become pixel values of 0, 8, 64 and 240, over a uniform 2D array (the whole channel contained the same value).

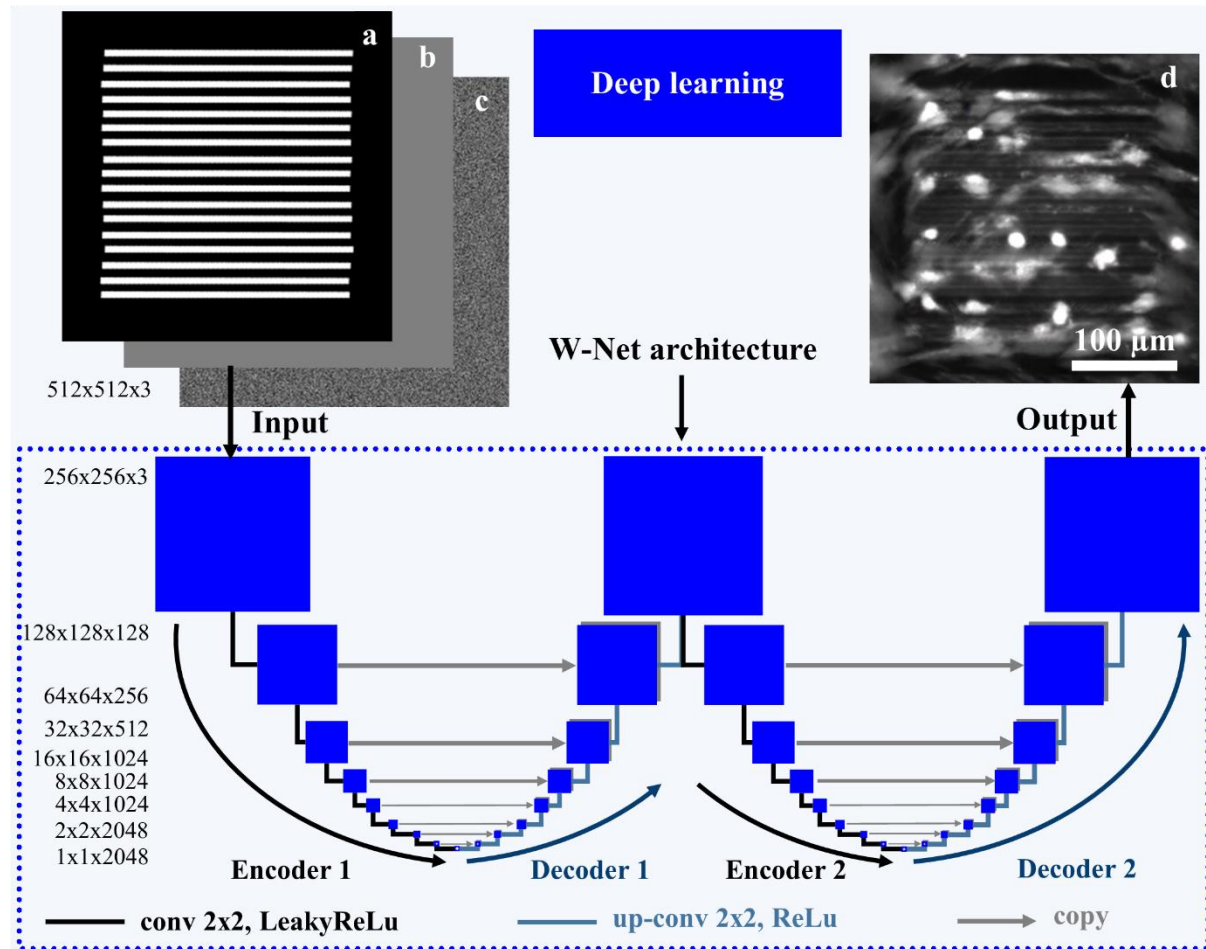


Figure 49: The deep neural network W-Net architecture consists of multiple convolutional layers and skip connections between encoder and decoder sections.

The input contains three data channels: topography (a), time (b) and density (c). The output (d) shows the neural network prediction of cell growth (as it would appear under a fluorescence microscope) for the topography, time-point and randomized cell density seed shown in (a), (b) and (c) respectively. The non-uniform input topography (a) is a result of randomisation introduced into the laser-machining process to increase parameters in the training dataset. Image modified from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

The cell density channel did not contain a uniform array, as each pixel was generated randomly using a Mersenne Twister PRNG via NumPy, which has a statistically uniform distribution. The maximum value for the PRNG was set by the mean fluorescence of an image, where the maximum pixel value 255 corresponded to the image with the highest mean fluorescence value in the dataset. As the minimum value for the PRNG was not altered, only

the maximum value for each dataset, the statistical distribution remained a Gaussian for all images, and the mean value of generated pixels increased for higher density images.

The network architecture is comparable to that used in previous chapters. Images were reduced in size from a maximum 1280 x 1024 to 512 x 512 through previously mentioned data augmentation methods of randomised cropping and rotation, before being resized using bilinear interpolation to 256 x 256, which balanced minimal data loss with time required to train the network. Training, consisting of 25 epochs (where one epoch is defined as every augmented image in the training dataset having been input to the network once) with a learning rate of 0.0005 and a batch size of 1, took two weeks on a NVIDIA Quadro P6000 GPU. The generator network, a W-Net, was an encoder-decoder-encoder-decoder architecture, with 34 layers, illustrated by the blue squares in Figure 49, a stride of 2, a 4 x 4 kernel size and used rectified linear activation functions. This can be visualised as an image decreasing in size from 256 x 256 down to 1 x 1, then increasing back up to 256 x 256, then repeatedly decreased and increased. The generator also contained skip connections between the mirrored layers, illustrated by the grey arrows in Figure 49. An additional network not included in the figure, a discriminator, was formed of 4 convolution layers with a stride of 2, resizing the image from 256 x 256 down to 32 x 32. This was then converted to a single output via a sigmoid activation function, which labelled images as realistic or unrealistic.

After each iteration (a single image is input through the network once), the output from the network was compared, pixel to pixel, to the corresponding experimental image (L1 error), leading to network improvements achieved via backpropagation. Simultaneously, the discriminator network was inputted either the generated output image or the experimental image, which the discriminator would correctly or incorrectly identify. The higher the discriminator error (D error) in identifying the authenticity of the image, the better the generated image. D error was therefore not only used to further improve the discriminator network, but also used to help train the generator by appropriately weighting and combining both the L1 and D error for backpropagation, which lead to alterations in the network throughout training to encourage realistic images and to statistically correct images at different stages. The use of a discriminator helps train the generator to produce realistic fluorescent images of cells in the most statistically likely position, rather than a blurred image that combined all likely cell positions.

6.4 NETWORK OUTPUTS AND ANALYSIS

6.4.1 INTRODUCTION TO MODEL TESTING

After the network had been trained, testing was performed using datasets consisting of experimentally obtained target images and the corresponding input images. To analyse the scope and limitations of using a DNN as a network model, only new feature combinations were used for testing, collected from entirely separate experiments than the experiments used for training data creation. While this potentially limited the variability of training data, as these experiments could have added additional variety to the training dataset, this ensured the network was tested in a scenario necessary for future application (where new and different topographies are being input to the network), preventing the network from providing apparently successful outputs that are copies of similar input-output images used in training. Unlike during training, when weights are altered within the network via backpropagation, the network parameters remain unchanged during testing. Experimentally obtained target images were individually compared to the network generated output images, rather than relying on L1 and D error.

Additional testing from artificially generated input features (with no experimentally obtained target images) investigated the ability of the network to cope with a wider range of input data to determine applicability as a model, where not all topographies consist of laser-machined squares. Testing therefore contained several different steps:

- The versatility of the network was analysed through independent and multiple alterations to input channels of topography, cell density (relative to confluence at the time of imaging, not the seeding density), and time-point (the day in which the fluorescent image is taken after cell seeding) in sections 6.4.2-6.4.4.
- The limits to network ability were found through introduction of curved, circular, and intricate line arrangements, when training data only consisted of straight laser-machined lines in section 6.4.5.
- Statistical significance of the network predictions in comparison to experimental data was obtained for input combinations unseen by the network during training (new topographies, new time-points, and cells isolated from different samples), and excluding overfitting, in section 6.4.6.

Once these steps were performed, the network was validated through discovery of a previously unknown variable, the minimum line separation required for cell alignment, which was then confirmed via comparison with experimental data in section 6.4.7.

6.4.2 TIME AND DENSITY INDEPENDENCE

During collection of experimental data, cell response to topographies changed over time, even when cell density remained relatively unchanged (Figure 50). For the low cell density, in this case a single cell, day 1 and day 8 showed highly varied responses.

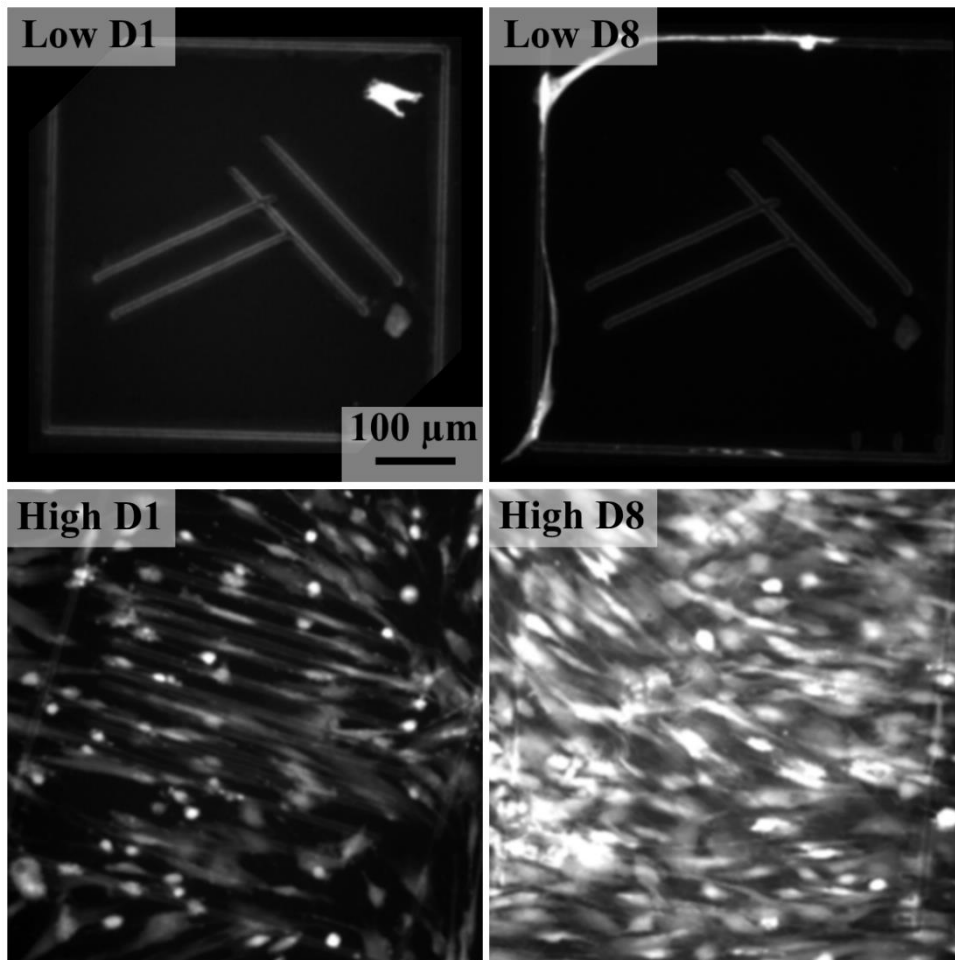


Figure 50: Changes to cell behaviour over time for different densities and surface patterning in experimental images. Low D1 shows a low cell density, in this case a single cell, at day 1 in a laser-machined square containing two parallel line pairs at random angles after one day of cell culture. Low D8 shows the same laser-machined box after eight days of cell culture at low density. High D1 shows a high density, multiple cells close to confluency, aligning to a set of laser-machined parallel lines within a laser-machined square and not aligning at other areas on the substrate around the square. High D8 shows the same cells on the same substrate after an additional week of culture, with no clear alignment/non-alignment of cells respective to the parallel lines and clear substrate areas.

At day 1, one day after cell seeding, there was no noticeable response from the cell to the surrounding topography. However, at day 8, it had adhered to the square laser-machined into

the topography. For a high cell density, where the cells were close to confluency, there was less visible change to cell response to topography. While there was clear alignment to the laser-machined parallel lines at day 1, the alignment was not clear at day 8 as coincidental alignment from the collective cell positioning could not be ruled out. Therefore, alignment and density were both independent of time.

The first test to the network was performed by independently altering density, relative to confluency, then independently altering the time-point. Experimental data showed visual differences between different time-points when density remained unchanged in (Figure 50). Network generated outputs were visually evaluated for whether high cell density, close to confluency, at a lower time-point, such as day 0, was different to a high cell density at a longer time-point, such as day 30 (Figure 51), which would then match experimental findings. As both a higher density and longer time-point can result in a large number of cells, these were input configurations that needed special attention. The input topography was blank and unchanged throughout this test (Figure 51: Maintaining a blank topography, time-points and density were varied to evaluate time-density independence. a), as alterations to the input topography would have altered the output of the network and the differences between high density and high time-point would have been difficult to evaluate.

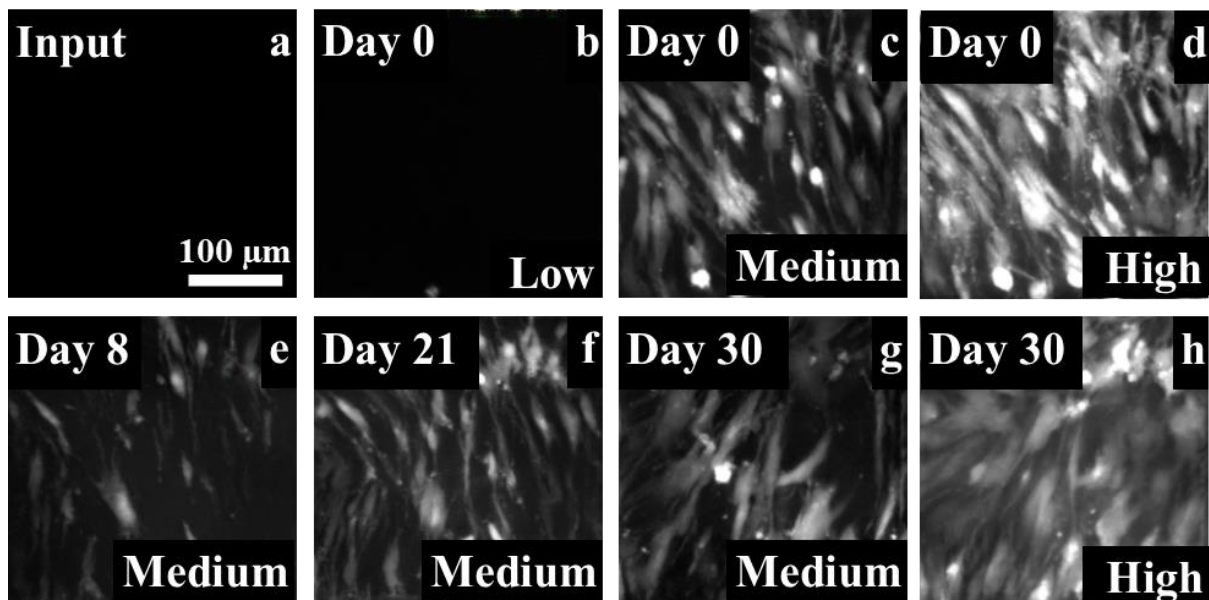


Figure 51: Maintaining a blank topography, time-points and density were varied to evaluate time-density independence. A blank topographical input with no laser machining (a) was constant while density variations (b-d and g-h) and time variations (c, e-g) were input into the network to produce predicted fluorescent image of stem cell response to the blank topography for the specified time and density input (b-h). Images (b-h) are network generated. Low, Medium and High refer to the cell density relative to confluency, where Low contains only a few cells whereas Medium and High is close to confluent. Image reproduced from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

As the time point was kept stable at day 0, the density input was varied from a lower to

higher cell density relative to confluency (Figure 51 b-d), which showed that variations did affect the network generated outputs. Generated images were realistic, with varied cell positioning across each image (differing amounts of space around individual cells), which likely resulted from the fluctuations in pixel value associated with the partially randomised input of the density channel. When the time-point was altered from day 0 and density was kept stable at a medium cell density, where there were several cells but they were not confluent, (Figure 51 e-g), cell positioning continued to change for every output, and the number of cells across most time-points (Figure 51 c, f and g) remained stable. Promisingly, there were morphological changes to the cells across these time-points, with more bright and spherical cells observed on day 0 (Figure 51 c) than after a month of adherence and spreading (Figure 51 g). The differences between a high density at a low time-point (Figure 51 d) and at a high time-point (Figure 51 h) confirm that the network had not incorrectly oversimplified increases in both density and time-point to only lead to an increase in cell number – changes to independent input channels resulted in unique network outputs.

6.4.3 TOPOGRAPHY AND DENSITY DEPENDENCE

Assuming there was no overfitting (investigated in section 6.4.6), a blank input topography would result in many available cell positions. However, a patterned input topography would result in a limited number of highly-likely cell positions, therefore impacting the positioning of cells for low, medium and high densities. Density was low if there were very few (0-5) cells, with a maximum pixel value of 50 in the density input channel. High cell density was close to confluent, and a maximum input pixel value of 255. Medium cell density contained more than 5 cells, but was not close to confluency, and had a typical maximum input pixel value of 150. The second test involved topographies with crossed-parallel and parallel lines, combined with a range of density inputs and a constant time-point of day 0 (Figure 52).

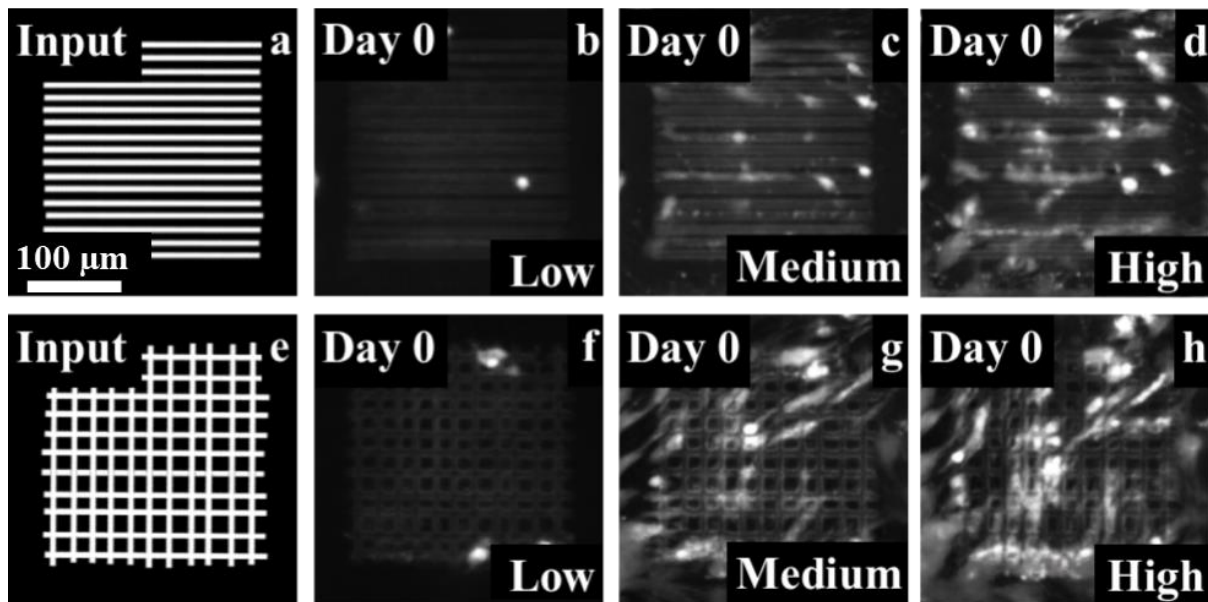


Figure 52: Two topographical inputs were used to test the network output for varying densities. One input was parallel lines (a-d), while the other was a grid of parallel lines crossing perpendicularly (e-h). Low (b, f), medium (c, g), and high densities (d, h) relative to confluency were input along with the topography of parallel lines for a constant time-point of day 0. Images (b-d, f-h) are network generated. Image reproduced from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

For each increase in input density, there was a visible change in network output but, unlike for the first test of a blank input, the cell positioning retains similarities across all densities, even while the cell number increases (Figure 52 b-d, f-h). For parallel lines (Figure 52 b-d), the single cell at low density (Figure 52 b) is visible in the network output at medium and high density (Figure 52 c and d). For cross-parallel lines (Figure 52 f-h), the network did not produce the clear cellular structures seen in previous tests, probably due to the inherent statistical variability in which direction cells align with.

When experimental data of crossed parallel lines was inspected (Figure 53), it was difficult to determine a relationship between cell alignment and the primary laser-machining direction. As the binary array also convey the directionality of the cross-parallel lines, the generated outputs were unclear. No statistical significance with the limited experimental data available on cross-parallel lines was discovered – there was no rule regarding cell alignment directionality. A higher density, with more cells to position, resulted in a large unstructured cellular mass (Figure 52 h), which was different from experimental results (Figure 53). More training data on crossed lines may provide the necessary information to prevent low-quality output, alongside depth characteristics added to the topography input channel, instead of a simple binary array. However, while the network did not successfully predict cell response to complex input topographies, it generated visible laser-machined lines that were

comparable to experimental images, seen in Figure 50 and Figure 53, from only binary input data. Modelling of laser-machining had previously been accomplished using DNNs and dedicated datasets [13, 182], but had been achieved by this network as a necessary step in generating fluorescent images of cell positioning on laser-machined topographies.

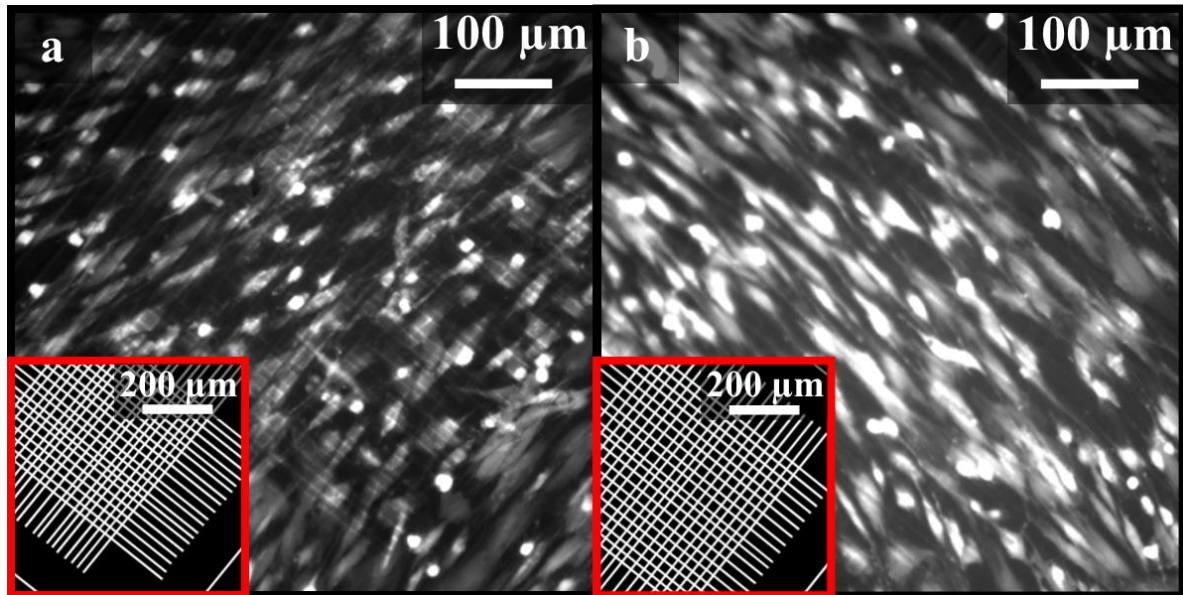


Figure 53: Experimental images of skeletal cells grown on two cross-parallel topographies. The alignment of skeletal cells in (a) is opposite to the alignment of cells in (b) even though the topographies are both cross-parallel laser-machined lines, and no relationship was discovered for consistent directionality. From the binary topography arrays, surrounded in red, there is no way to determine whether northwest to southeast was laser-machined first or last, and therefore no way for a network to determine which way the cells will align, either.

Even while details of the cellular structure proved too variable for the network to produce a single clear output, areas of likely cellular activity were still produced, which showed a similar scenario as for parallel lines. An area of skeletal cell activity in low density (Figure 52 f) was carried through in higher densities (Figure 52: Two topographical inputs were used to test the network output for varying densities. g and h) as the most probable cell position for that input topography. A lower density input provided the most statistically likely cell position without additional less-likely cell positions that were generated at higher densities. As a higher number of cells in network output images would result in a higher likelihood of matching cell position in experimental images, lower density inputs were used for statistical testing in later section 6.4.6, in which the network was statistically validated, so as not to artificially inflate the possible success of the network. However, medium density inputs provided a clear cell structure for relatively difficult topographies, while still providing a good number of cells to inspect, and therefore was used for further visual inspection.

6.4.4 TOPOGRAPHIES FOR ALIGNMENT

The last input to be tested independently was topography, with density and time-point remaining constant (Figure 54). Alignment behaviour is important in development of angiogenesis, biological neural networks, and the growth plate (chondrocyte columns) in the growth of long bones, so was an important focus for testing. Parallel lines of increasing separation, 20 μm , 90 μm , 120 μm and 290 μm , were chosen for the input topographies, which were not included in the training data. The exact mechanisms underlying cell alignment, such as cytoskeletal organization and interaction with other cells, has been observed as a response to various topographical cues [9, 16, 30, 72, 75, 79-81, 86, 88, 242-248]. At a wound interface, cells have a migratory phenotype switched on rather than a proliferative (division) phenotype and these cells influence the adjacent cells, sending signals to migrate to the wound space. It is therefore crucial to determine which topographies result in cell alignment and which do not.

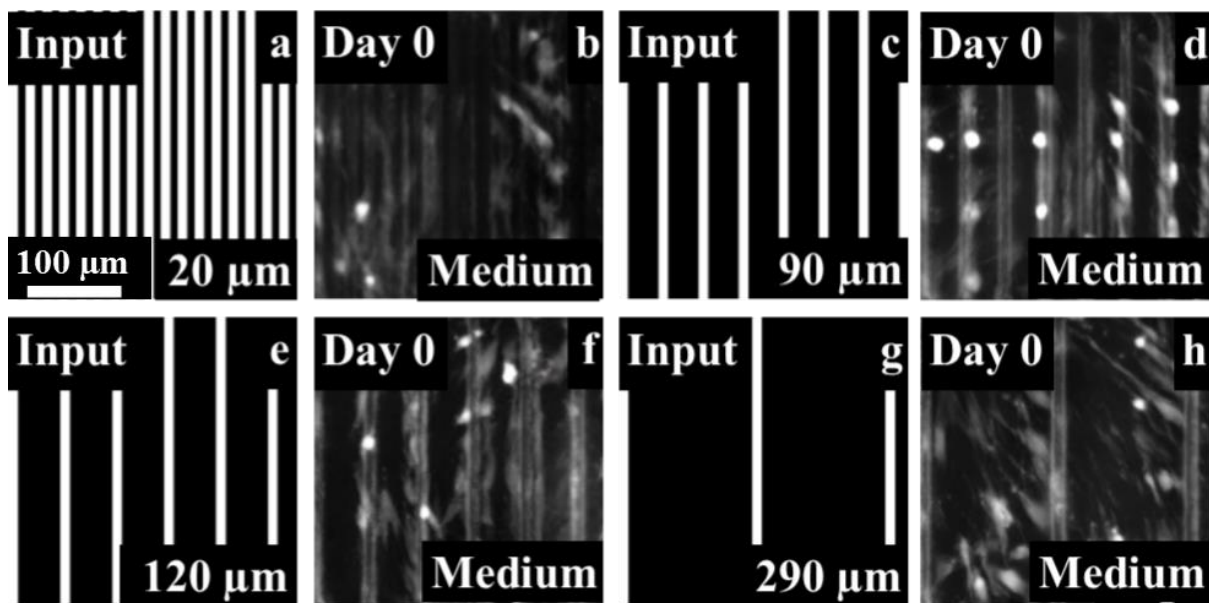


Figure 54: The generated output (b, d, f, h) for each line separation varies depending on the input topography (a, c, e, g). Time and density input remained unvaried at day 0 and medium density with respect to confluency. Input lines (a, c, e, g) were parallel and maintained a constant width of 25 μm , with separation (the spacing from mid-point of the line width to neighbouring line min-point) varied from 20 μm (a) to 290 μm (g). Images (b, d, f, h) are network generated. Image reproduced from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

Separation between parallel lines was measured from midpoint of one line to midpoint of the neighbouring line, and increased from 20 μm (Figure 54 a) to 290 μm (Figure 54 g). Line separations were chosen using a Mersenne Twister PRNG with a maximum value of 350 μm in intervals of 10 μm . Network output changed with each increase in line separation. At 20 μm (Figure 54 b), the cells both aligned and stretched across the lines, depending on which area of the image is analysed. For 90 μm separation (Figure 54 d), the cells were mostly

adhered to the lines, with only a couple of cells aligned alongside the lines in parallel. In experimental results, cells adhered to the machined lines can release signals to the unconditioned cells to align parallel [11], and the network has simulated these effects. The network predicted preference for machined areas, supported by experimental data (Figure 28), which implied that the microscale texturing of the laser-machined glass promoted adhesion and/or reduced cell mobility. Cell response was relatively unchanged when separation was increased to 120 μm (Figure 54 f), yet significantly altered at 290 μm (Figure 54 h), where there was no observed alignment.

6.4.5 EVALUATING VERSATILITY AND LIMITATIONS

Changes to input topography in all previous tests resulted in a change in predicted cell positioning and cell behaviour modelling (adherence, alignment, and visual structure). However, topographies consisting of a wider scope of features, instead of restricted to straight lines ranging from 10 μm to 25 μm in width, needed to be tested to visually assess whether the network was capable of extrapolating cell response. This is necessary for application as a model for skeletal stem cell behaviour, as one of the purposes of a network model is experimentation on topographies that have not been physically created. Consequently, the input topographies were varied to include concentric circles, parallel curves, small filled and unfilled circles, and intricate lines of changing widths (Figure 55). Such topographies were expected to show limitations of the network versatility, as only straight lines of 7.5 μm to 25 μm thickness were used in training.

The first topography, concentric circles, consisted of a line width similar to that which the network had used for training, and involved curves, which were not used during training (Figure 55 1a). The resulting network output (v 1b) contained both well-defined (cell shape visible and whole) and poorly-defined (unclear cellular mass or split unrealistically) cells, shown by the green and red circles respectively. There was some predicted cellular alignment and some areas of cellular activity crossing multiple lines, and more cells on or within proximity to laser-machined areas than blank areas. This was consistent with straight parallel lines at similar line separation (Figure 54 b) and previously predicted cell position preference for laser-machined areas (Figure 54 f).

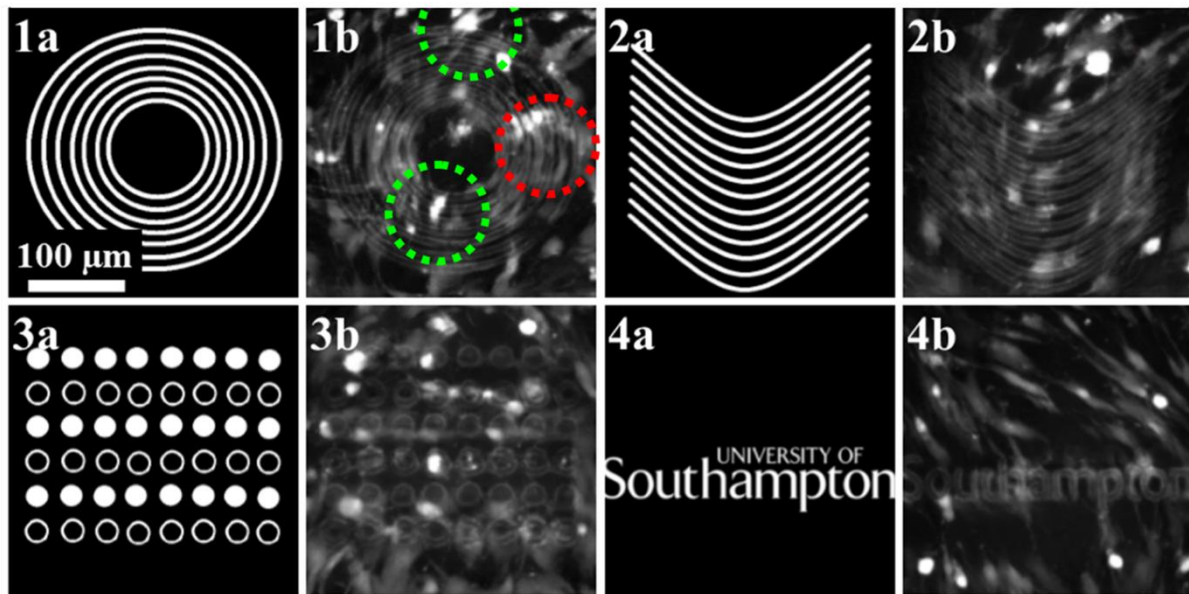


Figure 55: The limits to successfully generated output for extreme topographical input. Concentric circles (1), parallel curved lines (2), filled and unfilled circles of cellular size (3), and alphanumeric characters for line variety (4) were input into the network. Each (a, b) pair consists of a topographical input and corresponding network generated output. Well-defined cells are circled in green while badly-defined cellular mass is circled in red. Image modified from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

The second topography of curved parallel lines (Figure 55 2a) produced a similar output (Figure 55 2b). The third topography varied the most from the training data, consisting of multiple small circles (Figure 55 3a), which were either small continuously machined areas or thin lines. Neither design was included in training data, testing the network interpolation and generation. The network output (Figure 55 3b) was surprising as, while there were areas of unclear cell structure, there were also areas of clearly predicted cell response to the complex topography. Another promising feature was the asymmetry of the output. When a symmetrical topography was input to the network, the distribution and positioning of cells was not symmetrical, which would have been unrealistic as cells did not behave symmetrically in experimental data (Figure 50 and Figure 53). This could be due to the non-uniformity of the cell density input channel, or information gained from the asymmetry in training data.

The fourth topography consisted of alphanumeric characters of thin, varying line width, relative to lines used in training data (Figure 55 4a). Cell structure around the characters was not clear (Figure 55 4b), showing the network could not accurately predict cell behaviour around simulated laser-machined areas. However, the bright fluorescence of the larger characters correctly predicts cell preference for laser-machined areas and, interestingly, the finer characters were predicted to possess no adhesive qualities, as they are not visible. This implies the existence of a minimum linewidth for adhesion, in sub-micron range, where there

is no cellular response for a non-uniform lined topography. While the network had not been trained on sub-micron topographies, this prediction could be correct, as skeletal stem cells have been shown to have greater response to regulated and partially-regulated sub-micron topographies than to random topographies [9, 70]. Another reason for an absence of adhesion could be a lack of laser-machining in these regions. The finer lines (the word “university”) were not visibly machined, which could be evidence that the network had learned some properties of diffraction, which has been achieved by a similar network architecture in recent years [182]. This law of physics limits potential line-width unless specific adjustments are made to the experimental set-up – adjustments that were not in use during creation of the training data. The visible lack of “university” would be consistent with the laws of diffraction, and may be further proof that the network had been trained to generate predictions for both cell behaviour and laser-machining simultaneously.

6.4.6 STATISTICAL VALIDATION

Predicting realistic-appearing output images was a success for the network, but statistically evaluating the accuracy of these output images was vital. It was important that data used to test the network was completely independent from the training data, or the validity of the network could not be assessed. Training accuracy may only show how well the network had “memorised” the data, optimising weights to best fit particular inputs and reach the specific target output image, whereas testing accuracy should determine how accurate the network is in real-world application. Testing data was created using the same style datasets as for training data, but from a (new) unseen patient, on unseen topographies, imaged at an unseen cell density, including at unseen time-points.

Each of these testing inputs resulted in a network output that could be directly compared to experimentally obtained target outputs. A low density input was used to gain the most statistically likely cell positioning, as discussed in section 6.4.3. Four topographies were used for statistical validation, consisting of straight lines to minimise unclear cell structure, and differing patient and time-point combinations for larger scope (Figure 56). Breaking down the steps to create a visual indication of testing success, one topography (Figure 56 a) and accompanying time and density data was input into the network, resulting in a network generated output image predicting cell response to these features (Figure 56 b), which could be compared to the experimentally obtained fluorescent image for the same features, the target image (Figure 56 c), to create a comparison image (Figure 56 d). The network output was in blue, the positions of target cell positions were in red and the overlapping cell

positions were in green. Comparison images for all four topographies were created with the input topography overlaid in grey to show cell positioning relative to laser-machined areas (Figure 56 e-h). By looking at the areas of green pixels, there were cells correctly predicted at the end of laser-machined lines (Figure 56 e) at day 0, adhered to and elongated along boundary lines at day 8 and day 15 (Figure 56 f and g), and positioned off-centre of a singular line (Figure 56 h).

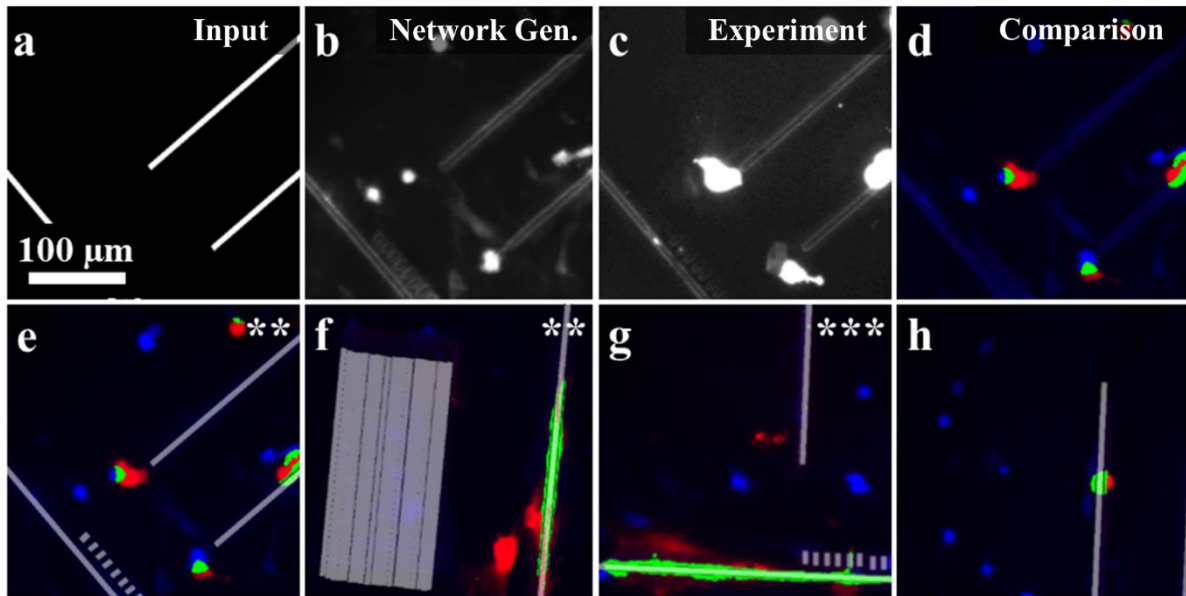


Figure 56: Statistical testing required analytical differences between target outputs and network generated outputs. A topography consisting of laser-machined lines (a) resulted in a network generated output (b) and target created from experimental cell culture (c). The difference between the two (d) was shown through three colours: the blue channel showed the network output, the red channel showed the positioning of cells in the target output, and the green channel showed where there were both predicted and experimentally imaged cells. The topography was overlaid in transparent white lines (e-h) for showing cell positioning in relation to laser-machined areas. Four different input topographies were analysed (e-h) with varying statistical significance, including $P > 0.05$ (no stars), $P < 0.01$ (**), and $P < 0.001$ (***). Scale bar in (a) applies to (a-h). Image reproduced from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

The abundance of blue showed that the network generated cell density prediction was larger than for the target images, which was due to the relatively low cell density within the testing images compared with the training images, and possible fluorescence variation between training data and testing data collection, which were generated from different experiments. A hypergeometric distribution probability mass function was therefore used to determine statistical significance of the correct network predictions, as the higher number of predicted cell positions compared to correct target cell positions could be accounted for [249]. The probability P can be calculated using:

$$P = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}} \quad (6.1)$$

Where N is the total number of cell-sized sections in the image, K is the number of target cells, n is the number of predicted cells, and k is the number of successfully predicted cells. Approximate cell-sized sections were used rather than pixel-to-pixel comparison because previous tests showed the network could generate realistic-appearing cells. The probability calculated from this method would therefore be the likelihood of correctly positioning those cells in comparison to corresponding areas on the target image. For the four topographies, the values of K , n , and k were calculated computationally, to avoid any human bias.

Three of the topographies were found to be statistically significant (Figure 56 e-g), with two topographies (Figure 56 e-f) achieving $P < 0.01$ and one (Figure 56 g) achieving $P < 0.001$. Interestingly, while visual inspection would suggest predicting the correct position of a single cell on a straight line would be statistically unlikely, the final topography (Figure 56 h) achieved only $P = 0.055$. This was due to the higher number of cells produced by the network output in comparison to the single experimental cell, resulting in increased probability from binomial adjustments. However, this higher number of cells did not prevent accurate placement of the sole target cell, showing that perhaps cell positioning is not inherently random (even when adhering to a single straight line) and, when other topographies are taken into consideration, can be accurately predicted.

With such promising statistical analysis, it was important to test negative controls such as inputs that contain no cells or no topographies, to confirm the absence of overfitting. In contrast to statistical analysis tests, where the network output and target output similarities were a success, similarities between network output and target output for negative control testing was considered a failure of the network. This is because, for example, the network output should be random without topographical guidance, and therefore different from the target output.

Two inputs of topographies with negligible/extremely low cell density input produced two low density outputs (Figure 57 1a and 2a), which were compared to target images (Figure 57 1b and 2b). Even with a cell density of virtually null, there were variations in the level of fluorescent particles that built up on the laser-machined lines over time during cell culture, which the network was given no way to account for – the network could not differentiate between an input of low fluorescence due to an extremely sparse density and a high build-up of fluorescence. This resulted in (successful) differences between the network and experimentally obtained images, with one network output incorrectly assigning a single cell

as the cause of fluorescence (Figure 57 1a and 1b), and another network output incorrectly assuming lower background fluorescence and greater particle build-up (Figure 57 2a and 2b).

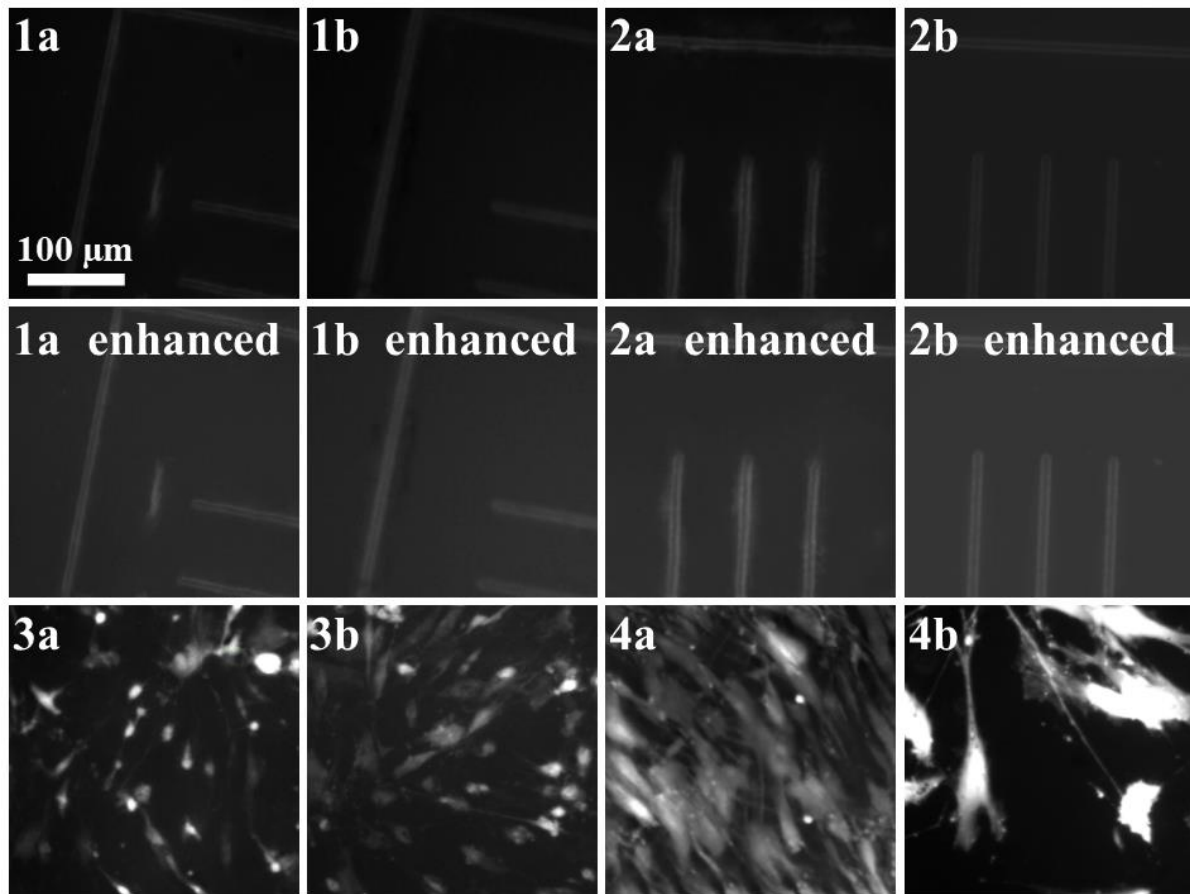


Figure 57: The network was tested for overfitting from a range of inputs that would produce a random, unpredictable, output. Network outputs (a) were compared to target outputs (b) of imaged experimental cell culture. A cell density close to zero was tested (pair 1 and 2), enhanced with altered brightness and contrast for easier visual analysis. A blank topography was also tested (pair 3 and 4) to confirm differing, random, output. Scale bar in (1a) applies to (1a-4b). Image reproduced from Mackay et al. under Open Access Creative Commons license [11] (© 2020 Mackay et al. Published by Elsevier Ltd.).

Differences between network and target output when there was no topography guiding cells were even greater, for both medium density (Figure 57 3a and 3b) and high density (Figure 57 4a and 4b). Cell shape and number were realistically created for medium density, but the positioning for each were (successfully) different. For high density, the limitations of the network regarding using fluorescence as an approximate density was evident, due to highly varied number of visible cells. The network predicted a high number of closely-packed cells (Figure 57 4a), whereas the true source of the high fluorescence was caused from a higher than expected fluorescence from relatively few cells (Figure 57 4b), leading to overexposure and a high average level of fluorescence. Yet this difference was also clear evidence of a lack of overfitting to training data given the inability of the network to predict the target image, the network did not produce “memorised” data. Instead, a (possibly random) likely outcome was

produced, which did not match the target image. Due to the lack of overfitting, which would have altered the true probability of a successful prediction, the network was confirmed as producing statistically significant results, and therefore validated as a potential model of cell response to topography.

6.4.7 NETWORK-MODEL PREDICTION

The purpose of the network model is to perform experimentation from a standard computer that would otherwise require a laboratory-based environment. While the network had been tested for accuracy when predicting cell response to topography, the network also required testing when used specifically as a model. This required obtaining the variable for minimum separation of laser-machined lines for cell alignment, using data from the network model, which was then verified using experimental data (Figure 58).

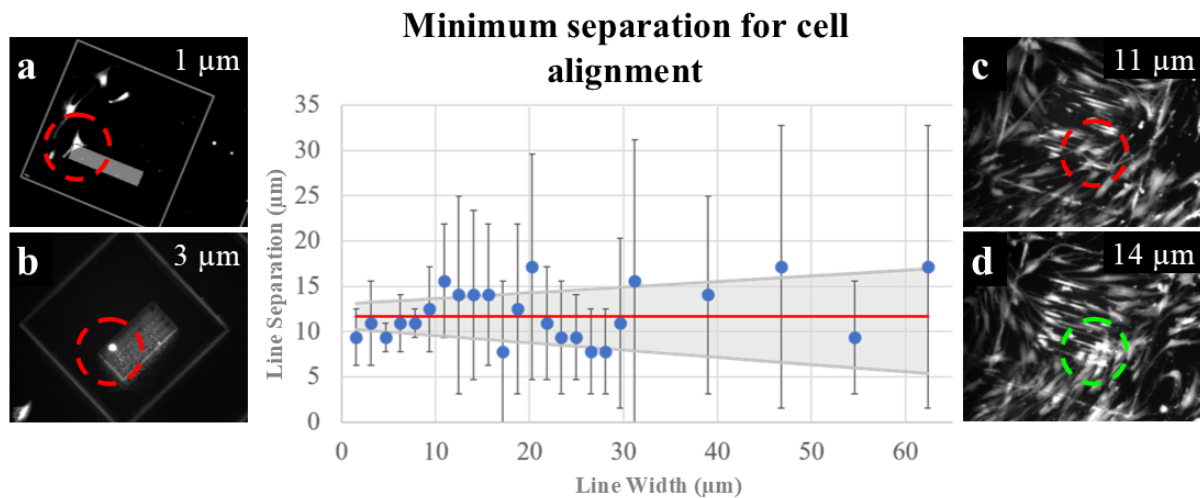


Figure 58: Four images of cell response to laser-machined parallel lines at increasing line separation, from 1 μm to 14 μm , and a graph of minimum separation required for cell alignment with data points (blue) taken from network output.

No relationship was found between the line width and separation between the lines for minimum separation required for cell alignment (red linear trend line with error as grey cone). Some of the experimentally obtained fluorescent images show cell alignment (within green dashed circle) and some fluorescent images do not (within red dashed circle).

The variable under investigation was the minimum line separation for cell alignment, as previous testing showed the possibility for such a value (Figure 54 on page 136) and also due to the importance of cell alignment for healing. The relationship between minimum line separation and line width was also investigated (Figure 58) to determine whether it was only the distance between the lines which caused cell alignment, or the distance from one line to another including laser-machined areas. Investigation involved producing repeated topographical input arrays, not laser-machined or existing in reality, of varying line separation and thickness, and combining each with a random distribution of density and time-point. For each line width, the range of line separations was analysed to determine majority

cell alignment (over half the cells aligned $\pm 10^\circ$ to the topographical lines), with repeat readings due to variations in density and time-point for each topography.

The average minimum separation for cell alignment was plotted as blue circular markers (Figure 58), with error bars covering the range of possible minimum separations at a specific line width. With only network output used, no relationship was found between line separation and line width, as shown by a red linear trend line (error as a grey cone) with a gradient of 0.0 ± 0.1 [$\mu\text{m}/\mu\text{m}$] (Figure 58). The extracted minimum separation was 11.7 ± 1.3 μm , which was then compared to four experimentally obtained images of cell response to laser-machined lines:

- i. One topography (Figure 58 a) contained line separation of 1 μm , small enough that there is limited real separation between laser-machined areas when observed under SEM microscopy (Figure 26).
- ii. The second topography (Figure 58 b) contained line separation below the minimum value at 3 μm , far below the majority of minimum line separation lower bounds for individual points, which typically averaged above at 5 μm .
- iii. The third topography (Figure 58 c) contained line separation similar to the minimum value at 11 μm separation.
- iv. The fourth topography (Figure 58 d) contained line separation of 14 μm , which was above the 11.7 ± 1.3 μm value, while still low enough that maximum line separation for cell alignment need not be considered.

Visual analysis showed no alignment in the first two topographies (separation below 11.7 ± 1.3 μm), while the third topography showed areas of alignment and areas of no alignment, which is unsurprising due to the complex variables involved in cell response, and 11 μm is within the error of 11.7 ± 1.3 μm . However, the fourth topography (separation above 11.7 ± 1.3 μm) had clear cell alignment in all areas of parallel laser-machined lines, including an area which did not produce alignment in the third topography, which was distinct and almost perpendicular to positioning of cells outside of this area. Observational analysis verified that the network can perform successfully as a model for cell response to topographies.

The 11.7 ± 1.3 μm minimum separation was likely specific to both cell-type and laser-machined topographies on a glass sample, as different cells respond to different topographical cues [30]. This leads to potential use as a morphological identifier in cell colonies cultured on surfaces with laser-machined parallel lines, without the use of invasive dyes and staining.

However, care should be taken that nanotopographical cues are not activated, as these surface cues interact with different cell signalling mechanisms and would interfere with the microscale cues [250].

6.5 SUMMARY

Exploiting biophysical cues such as microtopographies allowed for influencing cell response, such as positioning and elongation, as the controlled environment interacts with multiple signalling pathways of the cell. Investigation of every topography is inefficient and improbable due in part to the complexity and intricacy of these signalling pathways and the time it would take in experimentation alone. Application of a DNN as a network model, via training on 203 image-pairs, generated a model capable of producing statistically significant cell responses without the need for additional stem cell culture and imaging.

The first aim was to train a DNN capable of producing realistic-appearing fluorescent images of cell, which was successful. Initial testing of network model focused on the independence of the input channels and a lack of (inaccurate) generalisation between density and time-point. High density and high time-point were successfully independent and produced visibly different output images, where cell number was similar, but cell behaviour was not. Testing then moved onto topographies, where realistic alignment potential was visually evaluated. Compared to experimentally-obtained data, the network-generated images shared many realistic qualities, including the size, shape, and behaviour of the cells. When the versatility of the network was tested, limitations and success were visible for each topography (Figure 55), which showed that the network displayed versatility beyond the scope of the training data, but also that further training would be necessary to apply the network to all conceivable microscale topographies.

The second and third aims were focussed on whether the model could be used to simulated cell behaviour correctly. The second aim was to determine the statistical significance of generated images compared to experimentally created images. This was achieved using a hypergeometric distribution probability mass function, with one test image managed a statistical significance $P < 0.001$. The third aim was to derive a variable solely from network-generated images. The model had the potential to derive the minimum line separation required for stem cell alignment, $11.7 \pm 1.3 \mu\text{m}$, which was verified by experimental data, successfully completing the final aim. The network output for sub-micron input also showed a possibility to expand this method into nanoscale topographical response

Modelling The Cell Response to Surface Topographies

successfully. Through combinations of nanoscale and microscale alignment cues, interaction of different signalling mechanisms from a range of topographical cues could be investigated, possibly leading to a universal stem cell network model.

Through successful completion of all aims, the hypothesis has been proved correct.

This model could be linked to other networks as part of a larger system, which promotes clinician/researcher-AI integration and provides greater opportunity for testing and validation, and generating greater scientific understanding, than large “black box” AI systems [40]. For example, the previous network on identification and labelling target cells (chapter 5) would allow a combination of network models to focus on how specific cell types respond to environmental cues, reducing complexity and therefore necessary computational power, time spent training the network, and greater identification of problem areas, when the network model is applied to more complex colonies. A separate network that predicts cell response from biochemical cues, which will be investigated in the following chapter, could increase the ability of the system to predict cell response to any novel biomaterial. Therefore, the success of this network model may not only be viewed as the success of a single neural network for a specific set of inputs, but also as a successful potential first-step towards a larger AI-based system for enhanced biomaterial experimentation.

Linked to a larger system or as a separate network, creating an accurate model from limited training data provides a way for equal-access science, where economic or physical limitation and geographical location do not prevent biomedical experimentation. Such approaches offer new methodologies against, for example, pandemic lockdowns of the future given experiment-based laboratory work could continue using a network model, with promising results followed-up experimentally as appropriate.

7 PREDICTING BIOCOMPATIBILITY OF 3-DIMENSIONAL SCAFFOLDS

A background on AI application to reduce animal testing was covered in a literature review published in Biomedical Physics and Engineering Express [40]. Research based on this chapter will be submitted for peer reviewed publication.

After networks have been shown capable of identifying single and multiple cells in a single image (chapter 5), and capable of predicting cell behaviour to diverse and unseen biophysical cues (chapter 6), the project advanced to whether a network can predict collective cell behaviour in response to biophysical and biochemical cues. From identification of singular cells in 3D, to predictions of multiple cells in 2D, this chapter focusses on the behaviour of cell systems on 3D structures using 2D imaging.

7.1 INTRODUCTION

In previous chapters, the benefits of using a DNN for a model were explained, and the circular nature of improvement in science, data, and DNN models was introduced. This cycle is not limited to stem cells on laser-machined topographies, and can be applied to virtually any area of science, including other areas of tissue engineering and biomaterial development [220, 228] (Figure 59). While the DNN and areas of science is constant, the data involved in this chapter involves *in vitro* pictures of cells on 3D scaffolds, instead of laser-machined surfaces, and *in vivo* images of scaffolds on chicks. Instead of Stro-1 positive isolated HBMSCs, HBMSCs are used without isolation and, instead of laser-machined topographies, multiple varied 3D scaffolds are investigated.

Biomaterials are an important area of medical research, as the reparative response of the body and self-healing are limited. Use of biomaterials, alongside other regenerative medicine methods, can enhance healing by overcoming or mitigating restrictions, and harnesses various properties across a combined field of material science, chemical engineering, and cellular biology, among other disciplines [251]. An important difference between materials and biomaterials is biocompatibility, which is related to the behaviour of materials in various contexts and defined as “the capability of coexistence with living tissues or organisms without causing harm” in the Collins English Dictionary, determined from both *in vitro* and *in vivo* testing. Cell culture, *in vitro*, tests alone cannot determine biocompatibility of a material for *in situ* application due to poor correlation, especially for bone regeneration [157].

Therefore, *in vivo* tests are typically used to determine the biocompatibility of biomaterials designed for *in vivo* use.

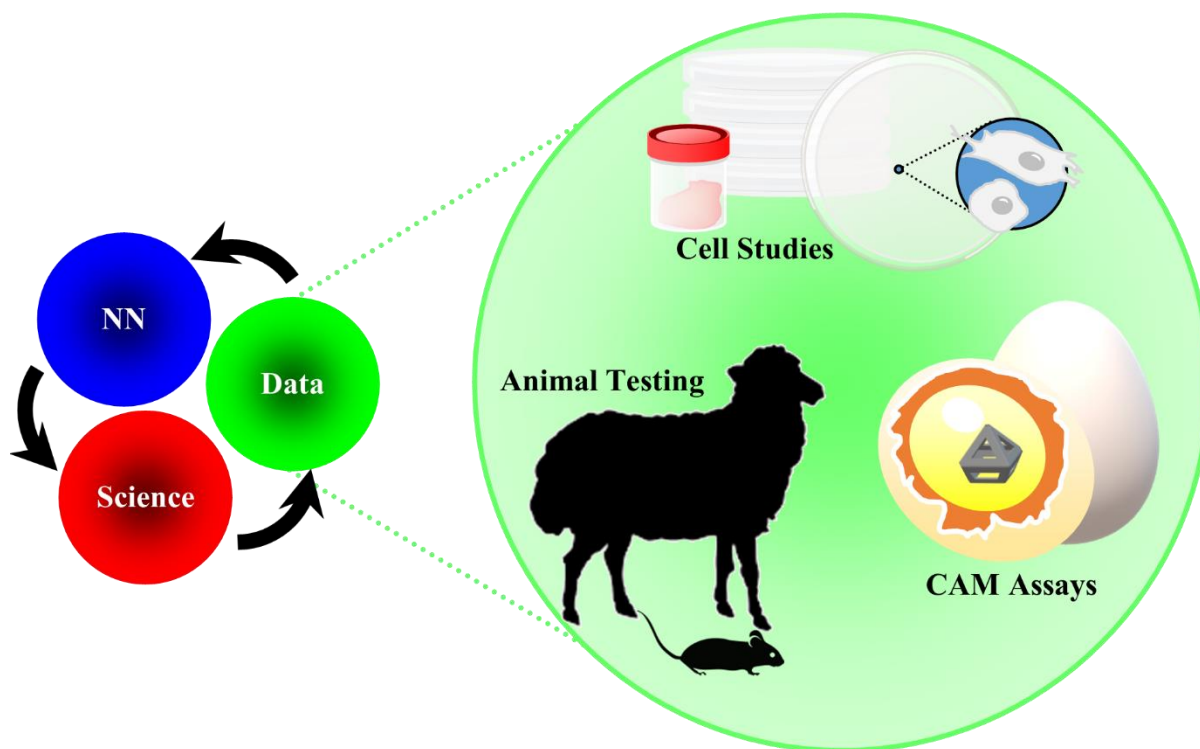


Figure 59: The cycle of improvement to science, data, and neural networks (NNs) can include many types of data. Data improves NNs, because network design can be optimised to the data, and the outputs of the model depend on data quality during training. Models can be used to improve scientific understanding, as they are capable of generating statistically probabilistic results to scenarios much faster than is possible in traditional experimentation, including for variables which could not be investigated ethically and/or physically in real world experimentation. This improved scientific understanding then improves data acquisition, and the whole system is improved further. The data in this system has few limitations, and can include data ranging from cell images (seen in chapter 4) to animal testing, including CAM assay images.

Animals have been used as a necessary (and strictly regulated) testing platform in medical research to evaluate biocompatibility of a materials, specifically in environments with strong correlation to the desired *in situ* location, without being unethical [252]. In a post Covid-19 pandemic environment, health science is increasingly on the public agenda, and possible discrepancies in ethical expectations of the public and the scientific community should be mitigated, as there have been studies with promising medical advances halted due to public concerns [253]. Guidance from the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) aims to prevent important medical advances being lost, while keeping animal suffering to a minimum, and therefore staying ahead of shifting, and improving, ethical expectations of the public and scientific communities alike. In chapter 6, the use of a DNN, a NN with multiple layers, could be used to reduce the

amount of *in vitro* laboratory experimentation required for stem cell research. This work could be advanced to not only reduce the amount of *in vitro* laboratory experimentation required for biocompatibility research, but also reduce the number of *in vivo* experiments [40]. If successful, a DNN could help to reduce, and possibly replace after enough training data is acquired for accurate generalisation, animal testing (2 of the 3 NC3R aims) by producing statistically likely biocompatibility predictions of untested biomaterials.

The training and subsequent evaluation of a DNN for biocompatibility prediction is complex, and encompasses multiple different sections, including the DNN itself. One DNN will be focussed on *in vivo* and the other *in vitro*, and then both used in a series to evaluate whether poor correlation can be overcome [157]. Section 2.4 covers the broad features of biomaterials not covered in chapter 6, biochemical characteristics and 3D scaffold characteristics, and a brief explanation of chick chorioallantoic membrane (CAM) assays. The hypothesis and aims are outlined in section 7.2 before the method is explained section 7.3; the results are analysed and discussed in section 7.4 and summarised in section 7.5.

7.2 HYPOTHESIS, AIMS AND OBJECTIVES

Deep learning could streamline screening of various biocharacteristic combinations without the use of further CAM assays. Mathematical modelling of a scaffold with different cell types interacting enables extension of results for a longer experimental period than is practical, due to chicken eggs hatching at day 21 of gestation, which allows further prediction of progression in a simulated environment [254]. DNNs could perform these predictions with data collected from necessary previous and current animal studies to predict outcomes of *in vivo* usage of various materials, and perhaps aid the design of new materials. The materials found to be biocompatible could then be taken forward to rodent studies and the unsuccessful candidates discounted at an earlier stage of the research pathway, leading to reduction in the number of animals used, refinement of materials being investigated to be more biocompatible, and simultaneously allowing modifications to enhance angiogenesis and osteogenesis results. This would effectively replace rodents with the CAM assay and computer simulations [40].

This chapter advances previous work in chapter 5 and chapter 6, using multiple DNNs for predicting cell response to both biochemical and 3D biophysical cues through generation of biocompatibility images (Figure 60). Potential success would provide significant savings to

Predicting Biocompatibility of 3-Dimensional Scaffolds

both costs and time spent conducting laboratory-based experimentation. Additionally, the methodology of this chapter could potentially result in reduced levels of animal testing.

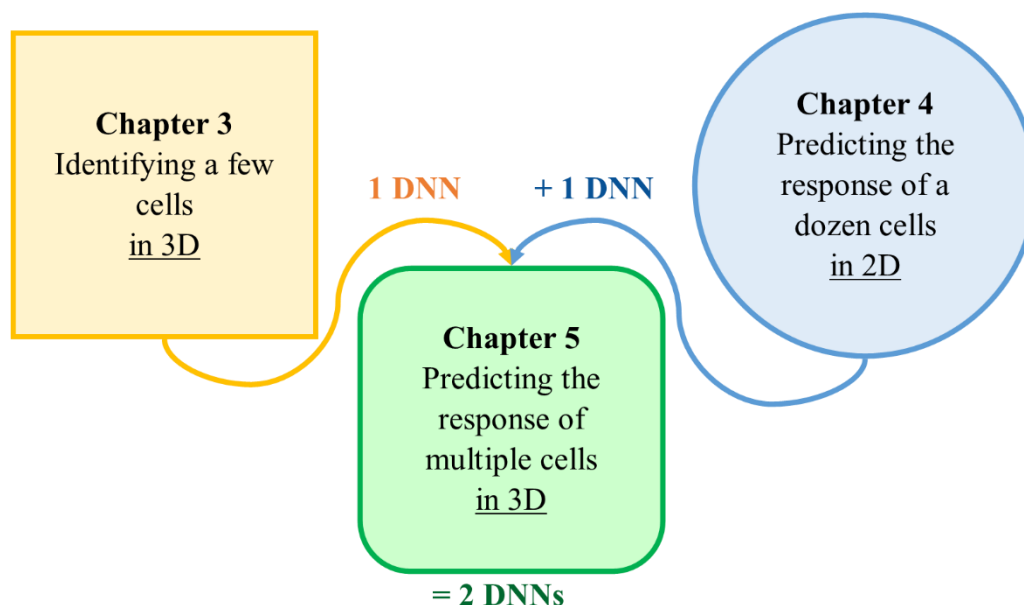


Figure 60: Chapter 5 advances upon work accomplished in chapters 3 and 4, with more DNNs likely required.

Work in chapter 3 was focussed on identification of single cells and endothelial cell structures in 3D. Chapter 4 predicted cell behaviour for cell numbers a magnitude higher and was mostly 2D, with limited depth variations in topographies. Chapter 5 will involve prediction of biocompatibility for cell numbers multiple magnitudes higher than in previous chapters on 3D structures. As chapter 3 and chapter 4 both used DNNs as large as computational power could sustain, the increased complexity of chapter 5 could require multiple DNNs.

7.2.1 HYPOTHESIS

CAM assay images for a scaffold of a specific material and coating combination can be generated that is indistinguishable from an experimentally obtained CAM assay image.

7.2.2 AIMS AND OBJECTIVES

There are three aims, which can be divided further into accompanying objectives.

- 1) Generate simulated fluorescent images of cells seeded on scaffolds.
 - a) Process and augment fluorescent images created by a collaborator for training of a DNN.
 - b) Test the trained DNN using scaffold material and coating combinations that were not included in the training dataset.
 - c) Determine the differences between network generated output images and laboratory obtained target images.
- 2) Generate a simulated CAM assay image of a scaffold on a chick membrane that has no statistically significant differences from laboratory obtained images.
 - a) Augment CAM assay images created by a collaborator for training a DNN.
 - b) Process the outputs from the trained network, which will be tested using scaffolds that were not included in the training dataset, for comparisons in key areas for CAM assay image analysis.
 - c) Calculate whether the differences in key areas for output and target images are statistically significant.

- 3) Use two trained networks in a series to form a network-chain for biocompatibility prediction.
 - a) The output from the first network will be input into the second network without any modification or processing. The subsequent output from the second network will then be visually compared to target images to evaluate the success of biocompatibility prediction.

7.3 METHOD

As detailed in previous chapters, a recurring criticism for using NNs are the hidden layers, which gives NNs the “black box” moniker. Data is converted from inputs to outputs through computational transformations at each layer – transformations that are not easily visualised, even mathematically, after the initial layers [255, 256]. For the networks used throughout this thesis, networks contained up to 17 layers in the U-Net and 34 layers in the W-Net, making visualisation virtually meaningless [255, 256]. Therefore, two networks were used instead of one to convert information on cell scaffolds (only the material and coating) into predictions of biocompatibility, via images of CAM assays (Figure 61), allowing for a check-point in the form of fluorescent images of cell cultures on the scaffolds.

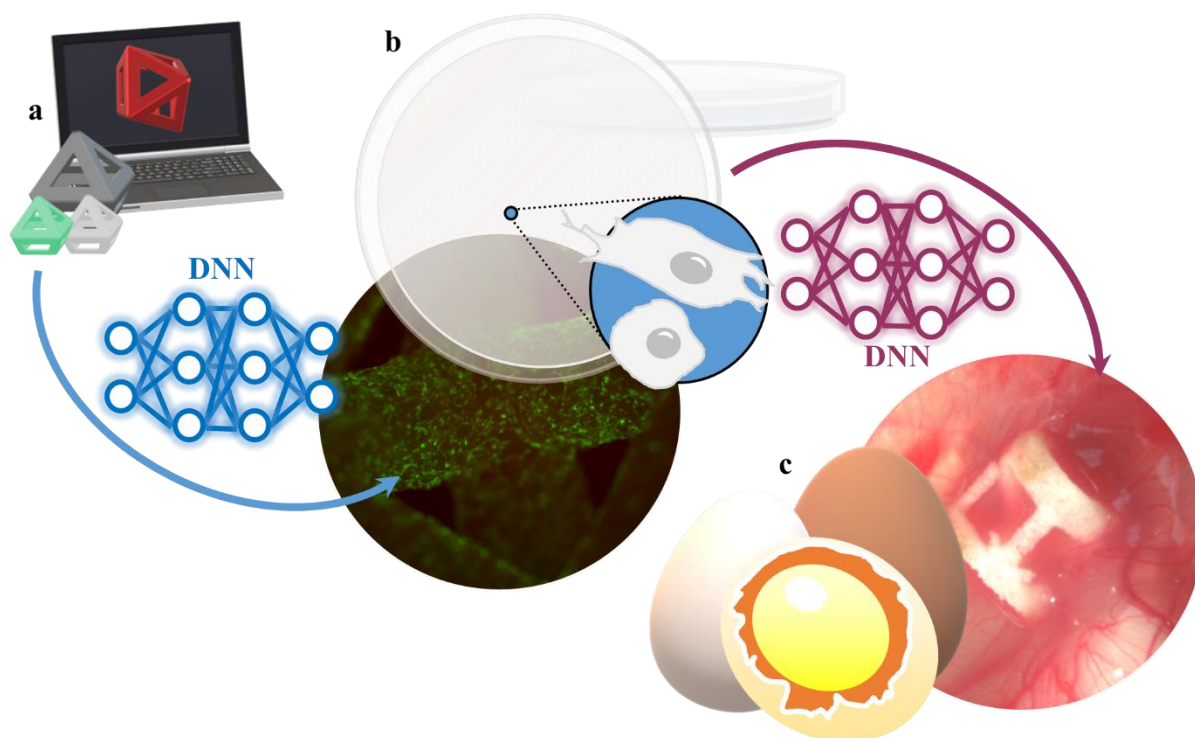


Figure 61: Illustration of two networks used in a series for biocompatibility prediction. The first DNN (blue) transforms information on the scaffold biochemistry to fluorescent images of HBMSCs after 14 days of culture/. The second DNN (purple) transforms fluorescent images into images of a CAM assay after 14 days of growth. (a) provides both biochemical and biophysical cues to HBMSCs; (b) provides information on the biocompatibility of the material in vitro via the approximate quantity and health of the cells; and (c) provides information on the biocompatibility in vivo via analysis of the number and branching of blood vessels, among other metrics.

Predicting Biocompatibility of 3-Dimensional Scaffolds

The two networks needed to be trained individually, and tested individually, before they could be used in a series for transforming scaffold information (the material and coating combination) into *in vitro* biocompatibility data (fluorescent images) and, more importantly, *in vivo* biocompatibility data (CAM assay images).

Work in chapter 6 illustrated the potential for a network to produce predictions on cell behaviour, through predicting fluorescent images, when input information of environmental cues. However, the network was as large as could be trained on the computational equipment available, and the output had limitations, where large alterations to topographies used in training resulted in poor quality images. By using two networks in a series, there are effectively two W-Net generators and two discriminators, with hyperparameters optimised for both halves of the data transformation, creating a larger network than could be trained otherwise. The method for the first half of the data transformation, from scaffold material and coating to fluorescent images (Figure 61 a-b), is covered in 4.3.1 and 7.3.1. The method for the second half, from fluorescent images of cell culture to CAM assay images (Figure 61 b-c), is covered in 4.3.2 and 7.3.2.

7.3.1 FLUORESCENT IMAGE ANALYSIS

Analysis of the network output predictions was performed together with analysis of target image (the images collected experimentally) for comparison of variables important in fluorescent microscopy. Additionally, a pixel comparison was performed between the output and target images to determine the visual difference between network prediction and experimental images (Figure 62).

The comparison was completed for every green and red pixel in shared positions so, for example, a pixel in the top left of an output image would be compared to the pixel in the top left of a target image. The difference in value for each position in the 256 x 256 image (65,536 pixel positions) were averaged to create a mean difference for each output-target image pair. As the difference alone does not account for the variability in the dataset, or provide useful information on biocompatibility, analysis on output images were analysed in the same way experimental images would be analysed to determine biocompatibility, with target images analysed the same way to provide comparisons for further network evaluation. This included calculation of the mean green pixel value (proportional to the number of live cells) per image and the mean red pixel value per image (proportional to the number of cells with damaged/porous membranes, which are considered dead or dying), alongside the

standard deviation for each. Mean pixel value was used rather than the mean number of pixels above a determined value as this required no thresholding, which could bias comparisons. Furthermore, the size of a cell relative to a pixel was not constant, due to the 3D shape of scaffold design, so the mean number of pixels would not be directly proportional to the number of cells.

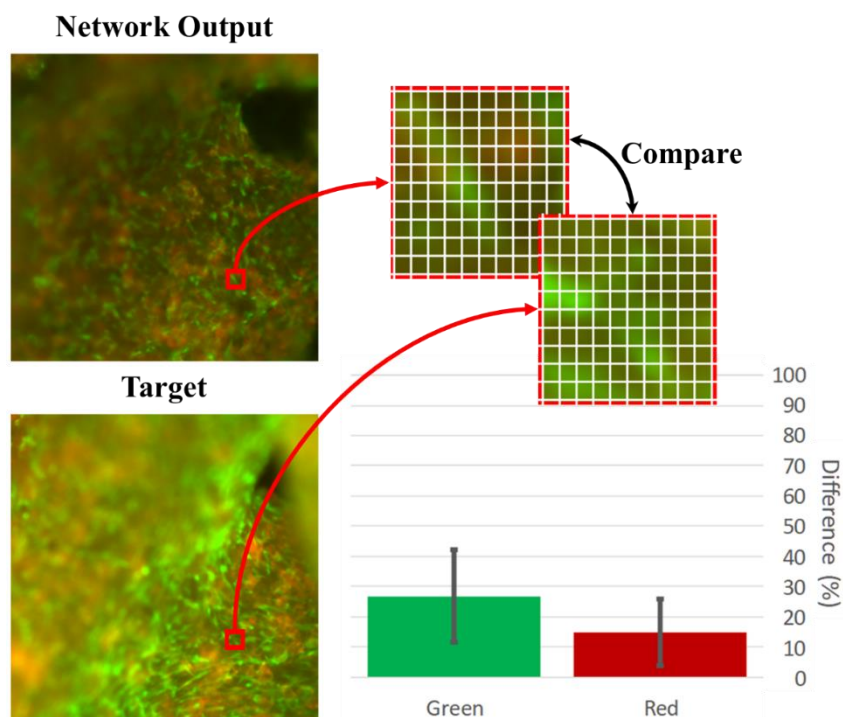


Figure 62: Analysis of material biocompatibility prediction using pixel comparisons. The difference in each pixel position for the network output and experimental target image is averaged across each image. This difference in pixel values is then converted to a percentage of maximum difference (maximum pixel value is 255) to derive a percentage difference for each output image. This was completed for green and red pixels individually, as both represent different fluorescent wavelengths/information on cell culture health.

7.3.2 BLOOD VESSEL ISOLATION AND ANALYSIS

To compare the biocompatibility of different scaffolds, three different metrics were analysed, including the quantity of vessels, the numbers of junctions (also known as branch points), and mean lacunarity (the average lacunarity across each image). Lacunarity is a geometric term used to measure the size and quantity of gaps between vessels, along with the heterogeneity of the image, where a low value (0.3 or lower) would be an organised, repetitive pattern, and a high value (1.0 or above) is heterogeneous, lacks symmetry, and has multiple (varied) gaps [257]. These metrics have been used throughout the medical field, and can be used as a quantitative value for comparing the pattern/distribution of blood vessels [257-273]. These metric values from network generated and laboratory obtained images were compared to

Predicting Biocompatibility of 3-Dimensional Scaffolds

determine whether the network could generate values that, for each scaffold type, were statistically insignificant from experimental values. If there were no statistically significant differences between the two samples, network output and laboratory target, the network has succeeded.

Statistical significance between the network generated and laboratory obtained values was calculated from a t-test, a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which could then be converted into a *P*-value, a measure of the probability that an observed difference could have occurred just by random chance. It was also important to determine whether there were statistically significant differences between different scaffold types and randomly generated images. This would determine whether the network was generating outputs that were dependent on the inputs or provided metrics uncorrelated to the inputs, even if the images appeared realistic visually. All values were calculated computationally and then checked using a respected online calculator at Social Science Statistics [274].

Before the quantity of vessels, the number of junctions, and the lacunarity could be calculated, the images needed processing for vessel extraction (Figure 63). Processing was automated to prevent any bias and decrease the time required to manually label every blood vessel. However, while vessel extraction can be tuned, through alteration to, for example, filter size, thresholding, and contrast, to accurately extract every blood vessel (Figure 11), this was difficult to perform across a whole dataset of experimental and network generated images. Therefore, the automated processing was focussed onto the larger (more than three pixels in diameter) blood vessels on the surface of the CAM assay images. As the same process was used for both network output and experimental target images, any systematic error would impact results equally and, as values were compared, be unproblematic for future analysis. This novel processing algorithm was based on vessel extraction by Mangir *et al* [14], combined with AngioTool from Zudaire *et al* [275].

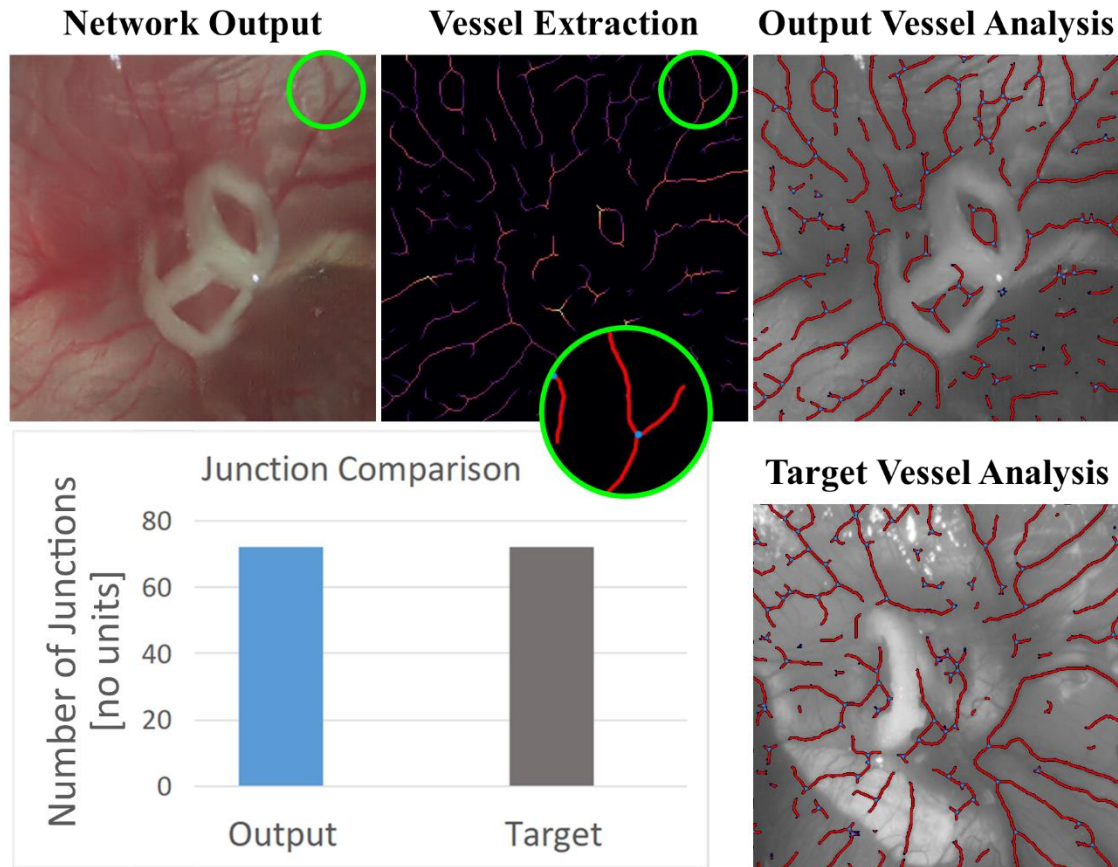


Figure 63: Processing of CAM assay images for comparative analysis of network output and target images. To obtain the quantity of vessels, number of junctions, and mean lacunarity for each image, vessel need to be extracted and skeletonised. Performed computationally, processing included Gaussian filters, local thresholding and colour extraction (Vessel Extraction). The resulting vessels were skeletonised so that metrics such as junctions (within the green circles), shown as blue dots, could be easily counted (Output Vessel Analysis). Small vessels, with a diameter less than three pixels, and vessels not close to the surface were not extracted. The metrics for both network output (Output Vessel Analysis) and experimental target images (Target Vessel Analysis) were compared via plots (Junction Comparison).

After vessel extraction and skeletonisation, the junctions were counted, the quantity of vessels as a percentage of total image area was computed, and the mean lacunarity was calculated. The values for each scaffold type were averaged, plotted and compared to randomly generated images. Networks were shown to produce realistic-appearing images in previous work, where images are considered realistic-appearing if neither a discriminator network nor a person can distinguish them from an image not generated by a network without context. Therefore, images were created with randomised vessel quantity, mean lacunarity, and number of junctions to determine whether the network was also creating randomised values, even though the appearance of the images could be realistic. To create these images for network comparison, the range for the number of junctions and width of vessels were taken from experimental target images, with a distribution centred in the central 100 pixels of the image.

7.4 RESULTS AND DISCUSSION

The two networks were tested using input data that had not been used during training, and then network generated output images were compared to laboratory obtained target images. Before processing and subsequent analysis, the images were visually evaluated to determine whether the networks were capable of generating images similar in appearance to the target images (Figure 64).

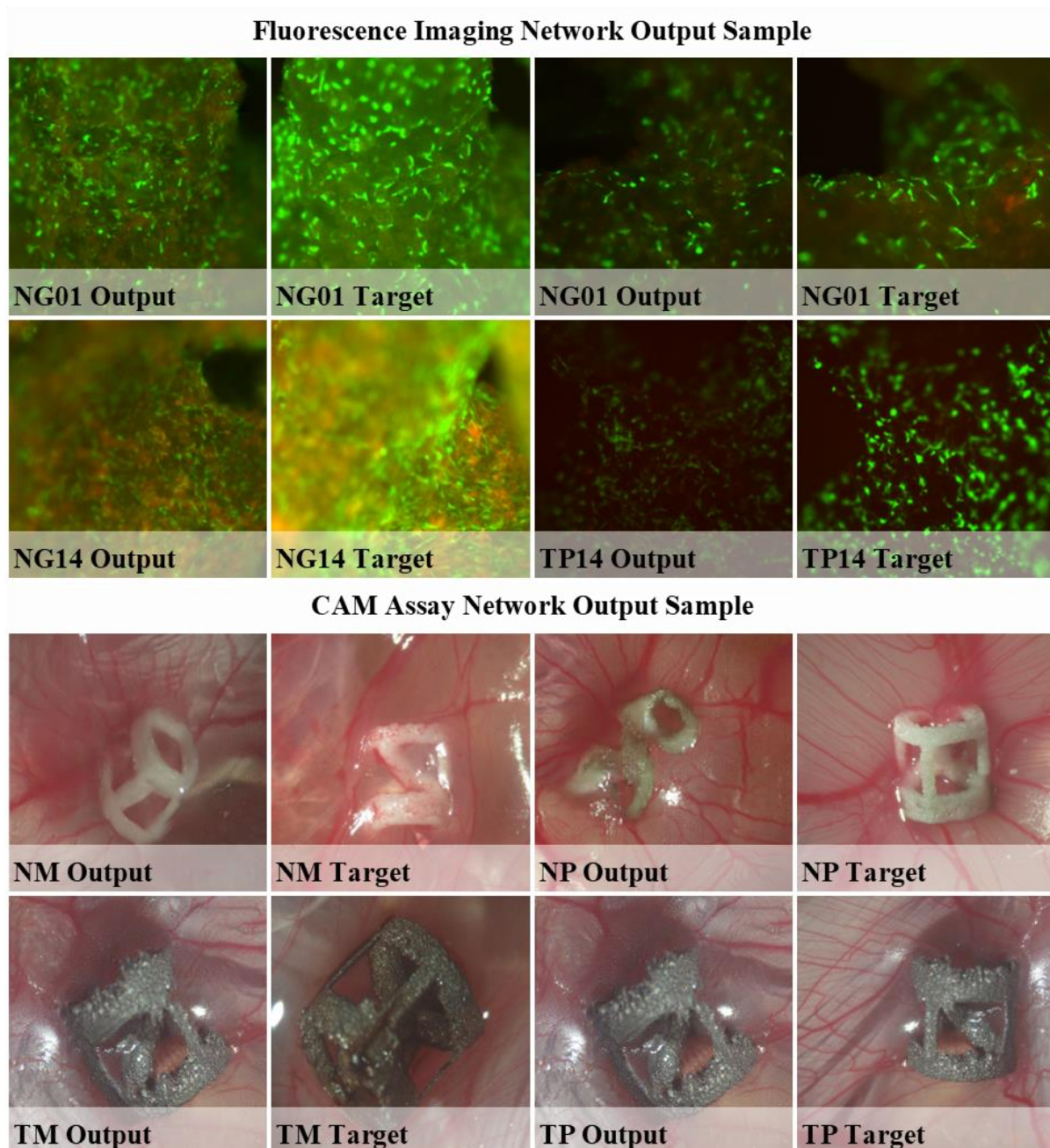


Figure 64: Randomly selected network generated output and experimental target images for multiple scaffold types. Three scaffolds were used in testing fluorescence image generation: Nylon scaffold with ‘Glasgow’ coating imaged one day after cell seeding (NG01); Nylon scaffold with ‘Glasgow’ coating imaged fourteen days after cell seeding (NG14); Titanium scaffold with no coating, left plain, imaged fourteen days after cell seeding (TP14). Four scaffolds were used in testing CAM assay image generation: Nylon scaffold with a mineral coating (NM); Nylon scaffold with a plain coating (NP); Titanium scaffold with a mineral coating (TM); Titanium scaffold with a plain coating (TP).

Images for visual evaluation were randomly selected from the multiple output/target image pairs that were generated/created for each scaffold type. The generated fluorescent image outputs were visibly less bright than the target images for each scaffold type. However, the 3D scaffold structure, shape and size of the cells, and the approximate health of the cells (the balance of green and red) was more successfully approximated. Both NG01 target images contained different levels of green and red pixels (there were variations in biocompatibility but also in imaging conditions between samples), which showed why averaging multiple outputs was necessary for network analysis. Analysis of a single output for each scaffold combination would not encompass the variations in laboratory images, and possible variations in network generated outputs. The randomly selected CAM assay output images showed mixed results. The outputs contained promising characteristics, such as the generated blood vessels and the scaffold designs in NM, TM, and TP outputs. Less promisingly, the output for TM and TP was the same, showing that the network was not always able to distinguish differences in coated and uncoated titanium. Further analysis would be necessary to determine if TM and TP outputs contained statistically significant differences over a larger sample.

Fluorescence imaging generation was analysed in section 7.4.1, where red and green pixels were separated to determine network success for both types of biocompatibility information (the amount of live cells and the amount of dead/dying cells independently). CAM assay image generation was analysed in section 7.4.2, with the quantity of vessels, number of junctions and mean lacunarity compared for output and target images, as well as statistically significant differences between the scaffold outputs. Finally, the networks were used in series in section 7.4.3, where the output from the fluorescent image network was used as an input to the CAM assay network.

Statistical significance was determined using the *t*-test, also known as Student's *t*-test [276]. This provides whether the difference between two samples could have happened by chance. For the *t*-test to be considered valid, the test statistic must follow a normal distribution. Therefore, the D'Agostino-Pearson Test for Normality [277] was applied to both the neural network output and the data generated for a comparison baseline, which will be detailed in the relevant sections below.

Predicting Biocompatibility of 3-Dimensional Scaffolds

7.4.1 FLUORESCENCE IMAGING GENERATION ANALYSIS

As detailed in section 7.3.1, the network generated output and laboratory obtained target fluorescence images were compared for all corresponding pixel positions, with green and red pixel values analysed independently. The difference in value between output and target images were averaged over the whole image, and the process was repeated for every output/target image pair. These differences were then averaged over each scaffold type to gain insight into which scaffolds were easier to predict (lower difference) and which were harder for the network to predict (higher difference). Rather than keep the difference values in pixels, the difference was plotted as a percentage difference, where the maximum difference possible, 255, equated to 100% and the minimum difference, 0, equated to 0%.

To form a baseline for analytical comparison, the scaffolds were plotted alongside randomly generated images of red and green pixels (Figure 65).

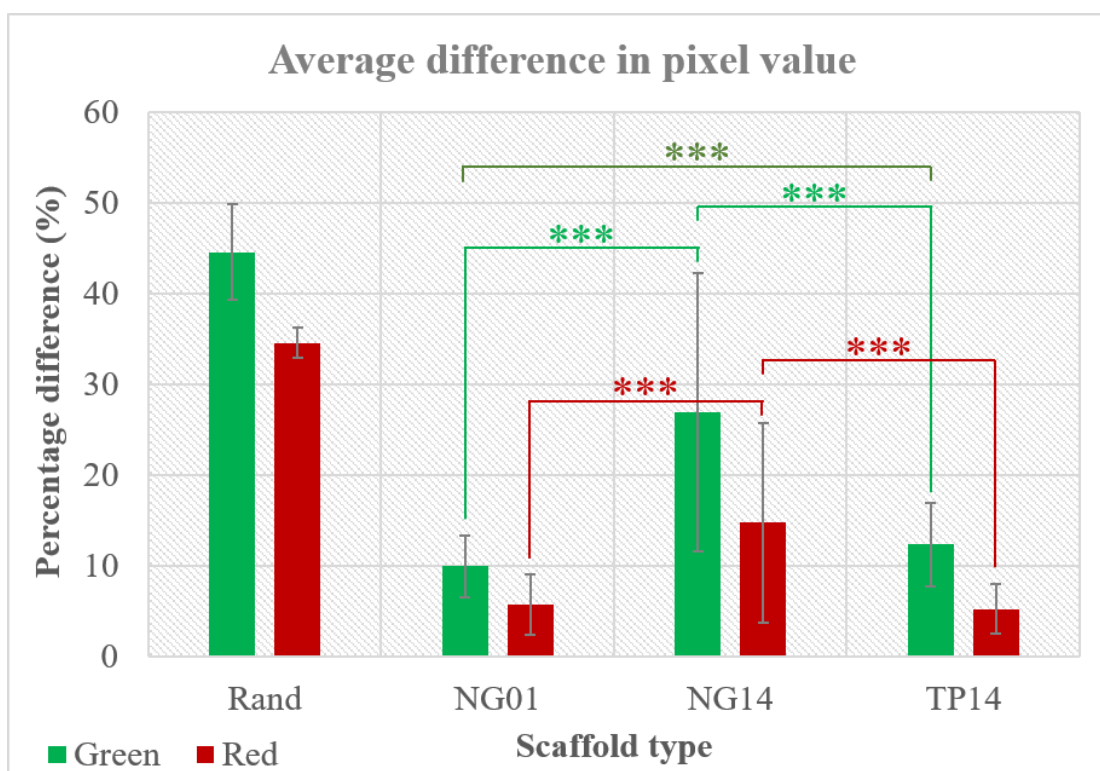


Figure 65: Plot of average percentage difference in green and red pixel values for testing scaffolds and randomly generated images.

NG01, NG14 and T14 were the three testing scaffolds. Additionally, randomly generated red and green images (Rand) were used for an analysis baseline. Green bars show the mean average of green pixel difference for all images of that scaffold type, and red bars show red pixel difference. Standard error is shown as vertical grey error bars. *** shows statistically significant differences between datasets of $P < 0.01$. Not all statistically significant differences are displayed.

This showed that, for all scaffolds, the network generated outputs were more similar to target images than if it were randomly generating images. Consequently, it showed that the network had used the input information of scaffold material and coating to create more accurate

fluorescent images than would be possible without this information. It follows that the network could be capable of determining the biocompatibility of different scaffold material-coating combinations that had not been included in the training data.

Each of the scaffold types contained statistically significant differences from another scaffold type, so the source of errors is unlikely to be only systematic, with a *t*-test providing $P < 0.01$ across both tested scaffold types, and the differences in each scaffold type could be analysed independently. The network generated images for scaffold NG01 showed less than 10% green pixel difference to laboratory obtained images, which was impressive as the network had not seen the combination of a nylon scaffold with the Glasgow coating. The cell response would likely be relatively consistent across multiple scaffold-coating combinations after only one day of culture, providing fewer likely biocompatibility scenarios for the network to predict. The higher percentage difference average for NG14, 27% green pixel difference, supports this theory, as the biocompatibility of a scaffold after fourteen days of culture could be harder to predict, and a large standard deviation in both green and red differences indicated a large range in the accuracy of the network. However, the percentage difference for TP14 was 12% green pixel difference and only 5% red pixel difference. Unlike for NG01 and NG14, this scaffold-coating pairing had been used in training data after one day of cell culture (TP01). The network was therefore not limited to only predicting *in vitro* biocompatibility of unseen scaffold-coating combinations, but could also be used as a time-predictor for two weeks of cell culture. Both uses would dramatically reduce the time taken to determine *in vitro* biocompatibility and optimise the design of novel biomaterials.

One of the visual comparisons between network generated output images and the corresponding target images was that the output images were not as bright as the target images (Figure 64), so lower pixel values would be expected. To determine whether this was true across the whole dataset, the green and red pixel values for each image were averaged over the whole image for each output and target image. Additionally, in analysis of fluorescent images created through laboratory experiments, the mean green pixel value would correlate to the size and health of the cells, and the mean red pixel value would correlate to the proportion of dead or dying cells. These are therefore important markers for the success of network generation, and output values were plotted beside target values for comparative analysis (Figure 66).

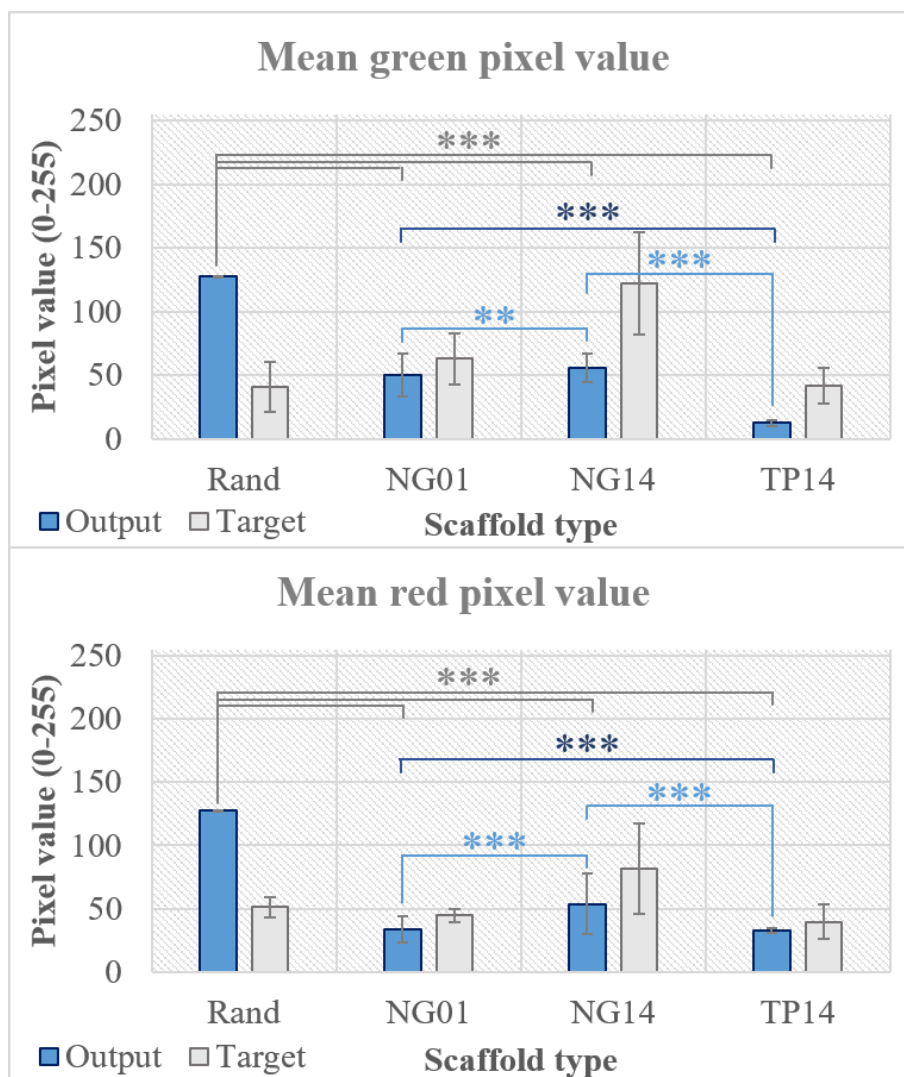


Figure 66: Plots of the mean green pixel value and mean red pixel value for each scaffold type. The mean green pixel value (top) and mean red pixel value (bottom) for each scaffold type (NG01, NG14, and TP14) and randomly generated images (Rand) against minimum to maximum pixel values (0-255). Blue bars show the averaged values for output images and grey bars show the average values for target images. Target for Rand images were randomly selected images from NG01, NG14 and TP14. Standard error is shown as vertical grey error bars. *** shows statistically significant differences between datasets of $P < 0.01$. ** shows statistically significant differences between datasets of $P < 0.05$. Not all statistically significant differences are displayed.

All network generated outputs were significantly ($P < 0.01$) different from random images, and significantly different for each scaffold type, which was determined statistically, as differences are difficult to visually determine graphically. Thus, output images were not random and depended on the input data, with differences to output for both changed scaffold material-coating pairings (NG14 and TP14) and changed timescales for cell culture before imaging (NG01 and NG14). The statistically significant difference between the randomly generated images and corresponding target images (images selected from all testing scaffolds) was promising, as it confirmed that randomly generated images could not match laboratory obtained fluorescent images. Less promisingly, there were also statistically

significant differences between the output and target for NG01, NG14 and TP14, which showed that the network had not generated images that matched laboratory obtained images.

Differences in red and green pixels shown in Figure 65 were visible in the combined differences between output and target values for all scaffold types in Figure 66. However, while the output and target images did not match value-to-value, the mean pixel values were always lower in output images than for target images, which was seen during visual inspection in Figure 64, where all the output images were comparable to target images but were all less bright in appearance (lower pixel values). Therefore, the differences in mean green and red pixel values may not be a result of unsuccessful predictions, just a poor approximation of total fluorescence. When target images for NG01, NG14 and TP14 were compared, NG14 displayed an increase in pixel values in comparison to NG01, with 93.0% increase in mean green pixel value and 82.9% increase in mean red pixel value. TP14 displayed a decrease in pixel values in comparison to NG01, with 33.4% decrease in mean green pixel value and 10.7% decrease in mean red pixel value. Importantly, these changes relative to NG01 were evident in the generated network output images, too, with NG14 increasing in pixel values and TP14 decreasing. The predicted percentages change for red pixels was more accurate than for green pixels, with as little as 8.9% difference between the predicted and real decrease for TP14.

7.4.2 CAM ASSAY IMAGE GENERATION ANALYSIS

Rather than counting vessels, which can be computationally complex with non-uniform images, the quantity of vessels was calculated in a simpler manner that still allowed for comparison in vessel numbers between scaffolds. The vessels were extracted and skeletonised as part of the CAM assay image processing (Figure 63), so the quantity of skeletonised vessels as a percentage of the total image area correlated to the number of vessels. The number of junctions, or branch points, were calculated computationally from areas of intersecting surface-vessels, using *AngioTool* [275].

For output and target images, the quantity of vessels and the number of junctions were averaged for each scaffold type, as well as for randomly generated images, and plotted (Figure 67). These random images were based on the average number of junctions for all target images, and the vessel-width. Together with the set range of vessel-width and number of junctions, the random vessel-like lines were coded to branch out from a centralised region, where a solid circle imitated the position of a scaffold, with the same diameter as a scaffold

but without the varied designs. This resulted in images that were not a collection of random pixels, but instead contained several boundaries that would ensure realism. While this provided a more difficult benchmark for network output analysis, it was important for rigorous evaluation of the generated images.

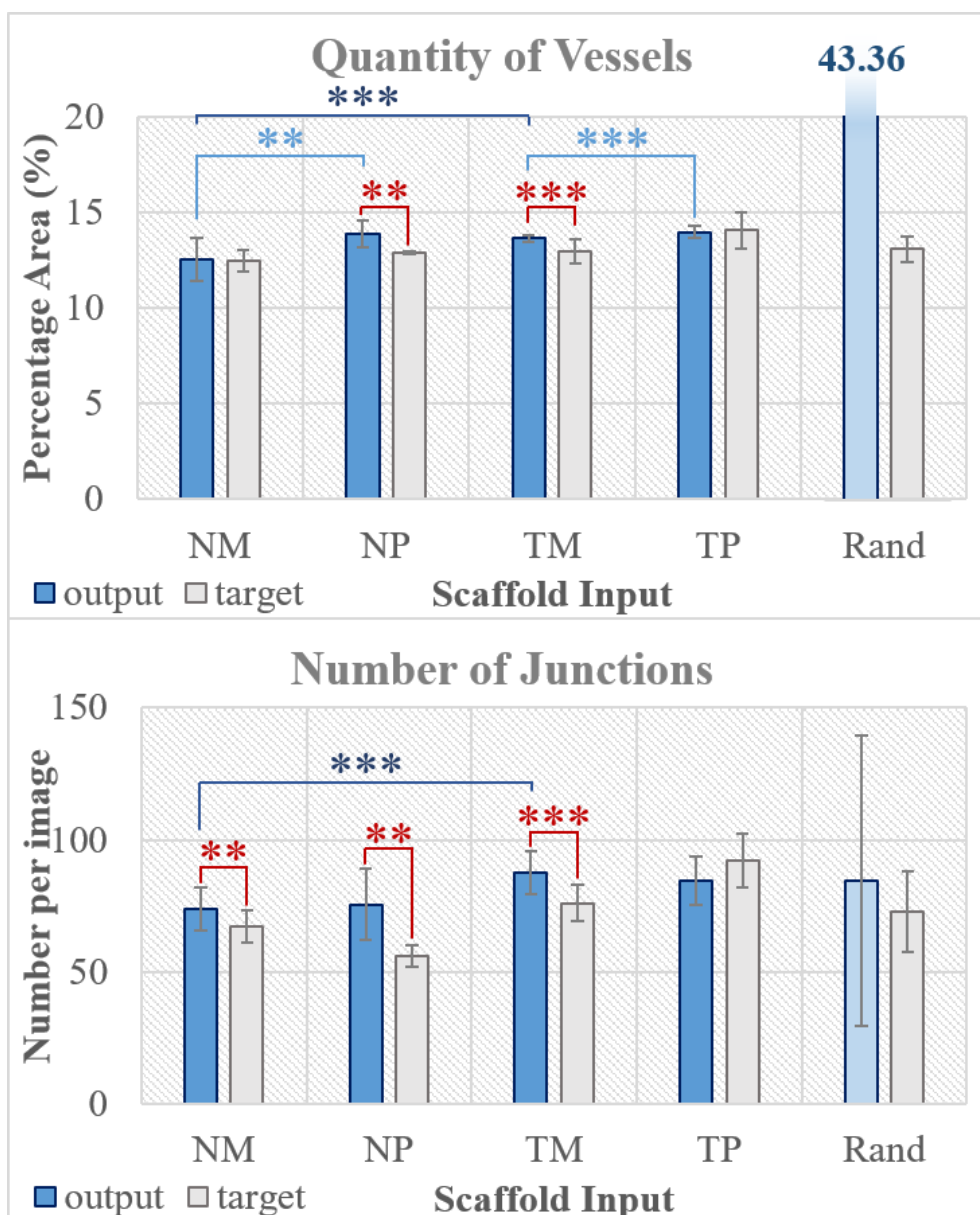


Figure 67: Plots of the mean quantity of vessels and mean number of junctions for each scaffold type. The mean quantity of vessels (top) and mean number of junctions (bottom) for each scaffold type (NM, NP, TM, and TP) and randomly generated images (Rand). Blue bars show the averaged values for output images and grey bars show the average values for target images. Target for Rand images were randomly selected images from NG01, NG14 and TP14. Standard error is shown as vertical grey error bars. *** shows statistically significant differences between datasets of $P < 0.01$. ** shows statistically significant differences between datasets of $P < 0.05$. The statistically significant differences between Rand and all output scaffold inputs were not displayed.

The quantity of vessels was similar for all scaffolds, between 12.5% and 14.0% for target images, and between 12.5% and 14.0% for output images. Since visual comparisons were

difficult, each scaffold was mathematically compared for statistically significant differences between output and target images. Two of the scaffolds, NP and TM, contained significant differences, but the scaffolds NM and TP had no significant differences between the output and target images.

Interestingly, that showed that the network was not better at generating a particular scaffold material, as nylon was not more or less successful than titanium, or a particular scaffold coating, as plain scaffolds were not more or less successful than mineral coated scaffolds. Additionally, NP and TP data was included in both training and testing data, although the images were from separate experiments, suggesting that additional training data was of limited effect. The network was trained on some NP and TP scaffold images from a separate experiment than the images included in testing, while mineral coatings were not included in the training data. However, the network was less successful with generating the correct number of junctions. Only TP output and target images had no significant images, although there were significant differences between the number of junctions for nylon and titanium scaffolds (there were more junctions for titanium than nylon scaffolds), which was seen in both network generated output and laboratory obtained target images.

Generated outputs more closely matched the mean lacunarity (λ) of target images than for the number of junctions (Figure 68), showing the network was more successful in this area. Two of the four scaffolds had no statistical differences between output and target images, meaning that the network generated images were statistically similar to target images for both NM and TP scaffolds, but not for NP and TM. However, there were visual similarities between both successful and less successful output-target pairs (Figure 69). Importantly, the lacunarity for each output average was close to the target average ($< 0.05 \lambda$), while the difference between randomly generated images and the average of all target images was almost 0.3λ . This, along with the statistically significant differences between scaffold materials and coatings, showed that the network was not randomly generating images, and that the output images depended on the fluorescent image inputs.

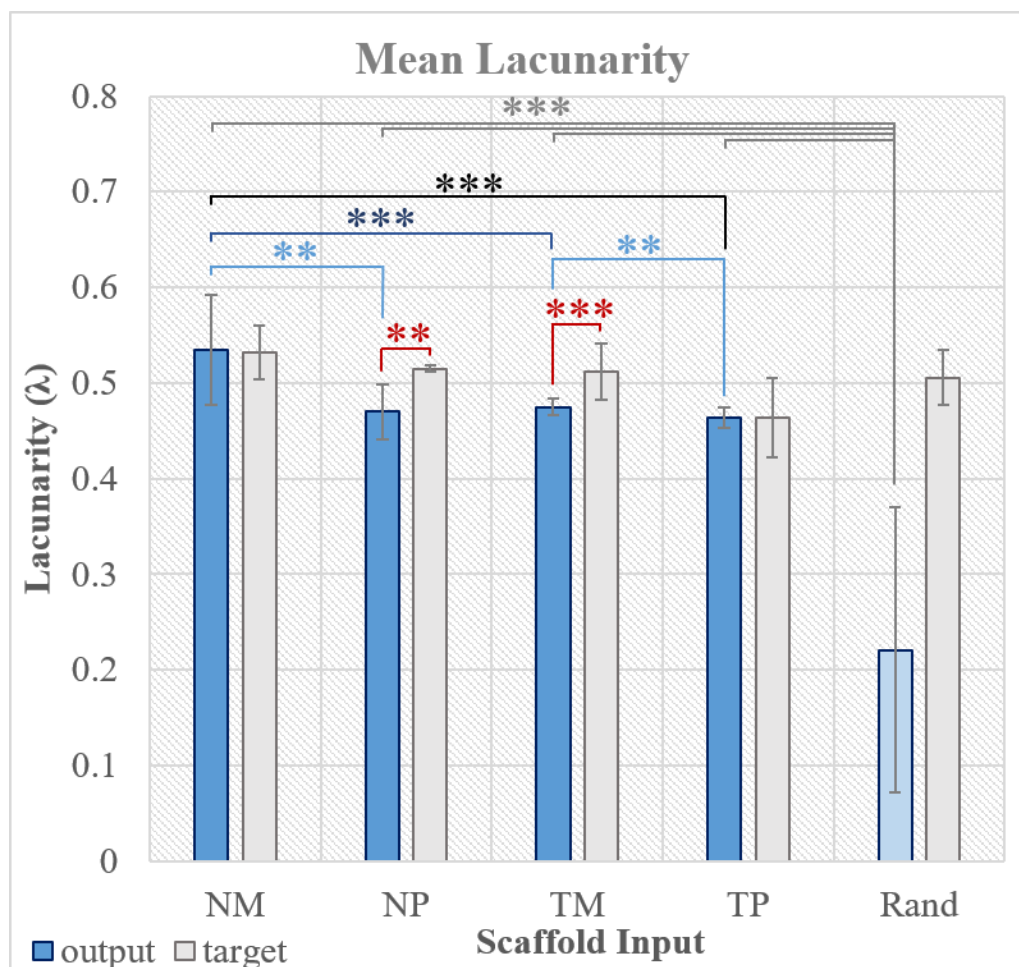


Figure 68: Plot of the mean lacunarity for each scaffold type and randomly generated images. Blue bars show the averaged values for output images and grey bars show the average values for target images. Standard error is shown as vertical grey error bars. *** shows statistically significant differences between datasets of $P < 0.01$. ** shows statistically significant differences between datasets of $P < 0.05$. Not all statistically significant differences are displayed.

The (statistically significant) differences between each output scaffold are comparable to the same differences between target scaffolds, with NM having a higher mean lacunarity than NP and TM, and with TP having the lowest mean lacunarity. Visually, this corresponds to CAM assay images for TP scaffolds being more heterogeneous and with larger gaps between vessels, compared to NM scaffolds. Interestingly, these differences were generated by the network with no input data on the scaffold material or coating, as the only input was fluorescent images of the scaffolds after *in vitro* culture, which implies a link between *in vitro* and *in vivo* biocompatibility was made by the network.

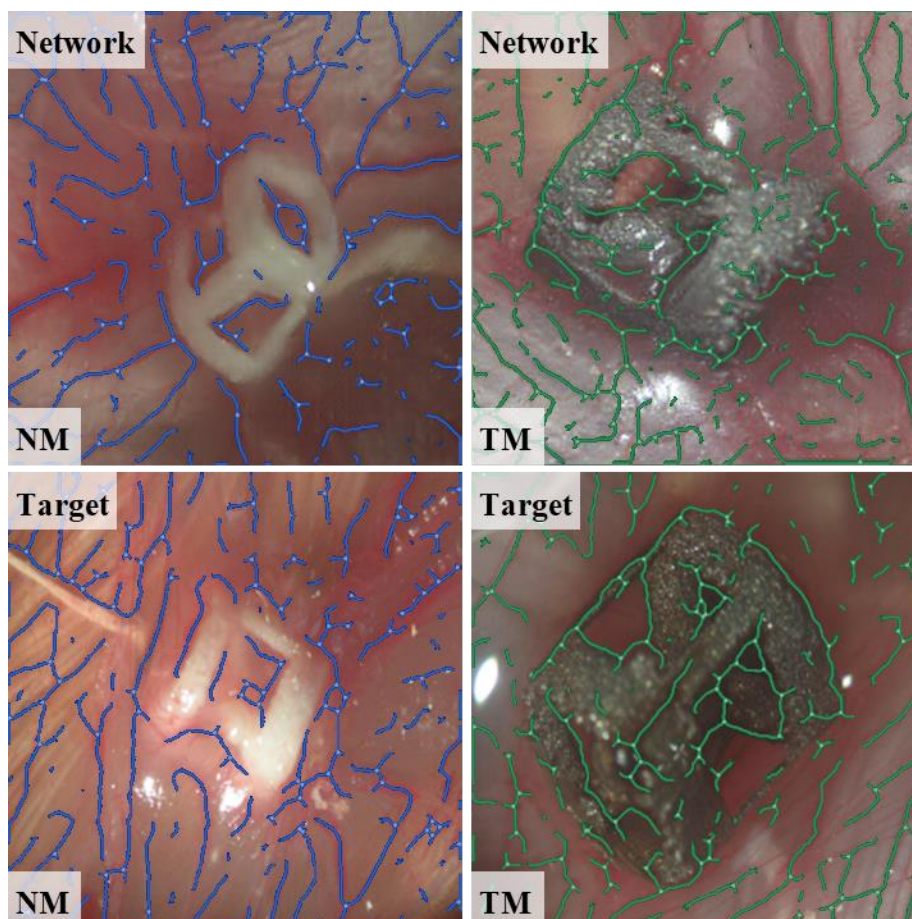


Figure 69: Two scaffold predictions and accompanying target images overlaid with the extracted blood vessels. Two scaffold types, NM (blue) and TM (green), have the network generated and laboratory created image displayed beneath the extracted blood vessels. Extracted blood vessels were large enough and close enough to the surface to be labelled during the automated extraction process. NM mean lacunarity was successfully predicted by the network, and shows multiple clear long blood vessels, while TM mean lacunarity was unsuccessfully predicted, and vessel extraction appears more fractured.

7.4.3 LINK-UP OF MULTIPLE NETWORKS

The final application of the network included using the two networks, one for generating fluorescent images and the other for generating CAM assay images, in a series (Figure 61). This resulted in a data transformation from the scaffold material and coating combination to a CAM assay image for that scaffold, with an additional step in the middle for checking that the transformation was visually credible, which was the output from the first network and subsequent input to the second.

The input was the same as for section 7.4.1, where the scaffold material, coating, and relative depth was input to the first network (no additional chemical or physical knowledge was provided). Relative depth was computationally calculated from the focus detection algorithms, based on convolving the image with a Laplacian kernel to determine the blurriness of different sections, with a pixel value assigned to the level of focus. This limited 3D information was necessary for a biocompatibility prediction: When a 2D depth profile

Predicting Biocompatibility of 3-Dimensional Scaffolds

was input into the trained network (a uniform focus across the image), no recognisable fluorescent image prediction was generated. This could be a result of overfitting to a 3D training dataset, or the effect of 3D scaffold design on biocompatibility, which future inclusion of 2D images could help determine.

For this test, two scaffold combinations were used, NG and TP. NG and TP had not been used in training of the first network, but had been used in training the second network. These two scaffolds were chosen because the scaffolds had both previously been analysed, with NG the least successfully generated and TP the most successfully generated in testing. The check step, where the first network output is used as the input for the second network with no additional processing or alteration, created a way to control the path of, and visually evaluate, the data transformation. The output for the data transformation was the CAM assay image for the input scaffold, which could then be compared to target images (laboratory obtained CAM assay images for that scaffold).

Two variations for each scaffold, NG and TP, were used during the network in series test, to account for variations in possible outputs, seen previously in Figure 64. The input appears as many oblong sections with varying brightness, which corresponds to the 3D information as explained earlier, and colour, corresponding to the scaffold material and coating. Oblongs at a range of rotations were used instead of uniform squares to allow for greater variety in 3D profile inputs in application. For NG-1 (the first NG scaffold input variation) and NG-2 (the second) in Figure 70, the check step showed a realistic appearing fluorescent images with high levels of green and moderate to high levels of red pixels, which matched observations of experimental data in Figure 64. TP-1 and TP-2 showed similarly promising transformations in the check step. Importantly, when these network generated images were input into the network, the images were similar enough to laboratory obtained images that the second network was able to utilise the data to form a CAM assay image outputs. This was confirmation that the output images from the first network were realistic at both a visual and numerical level.

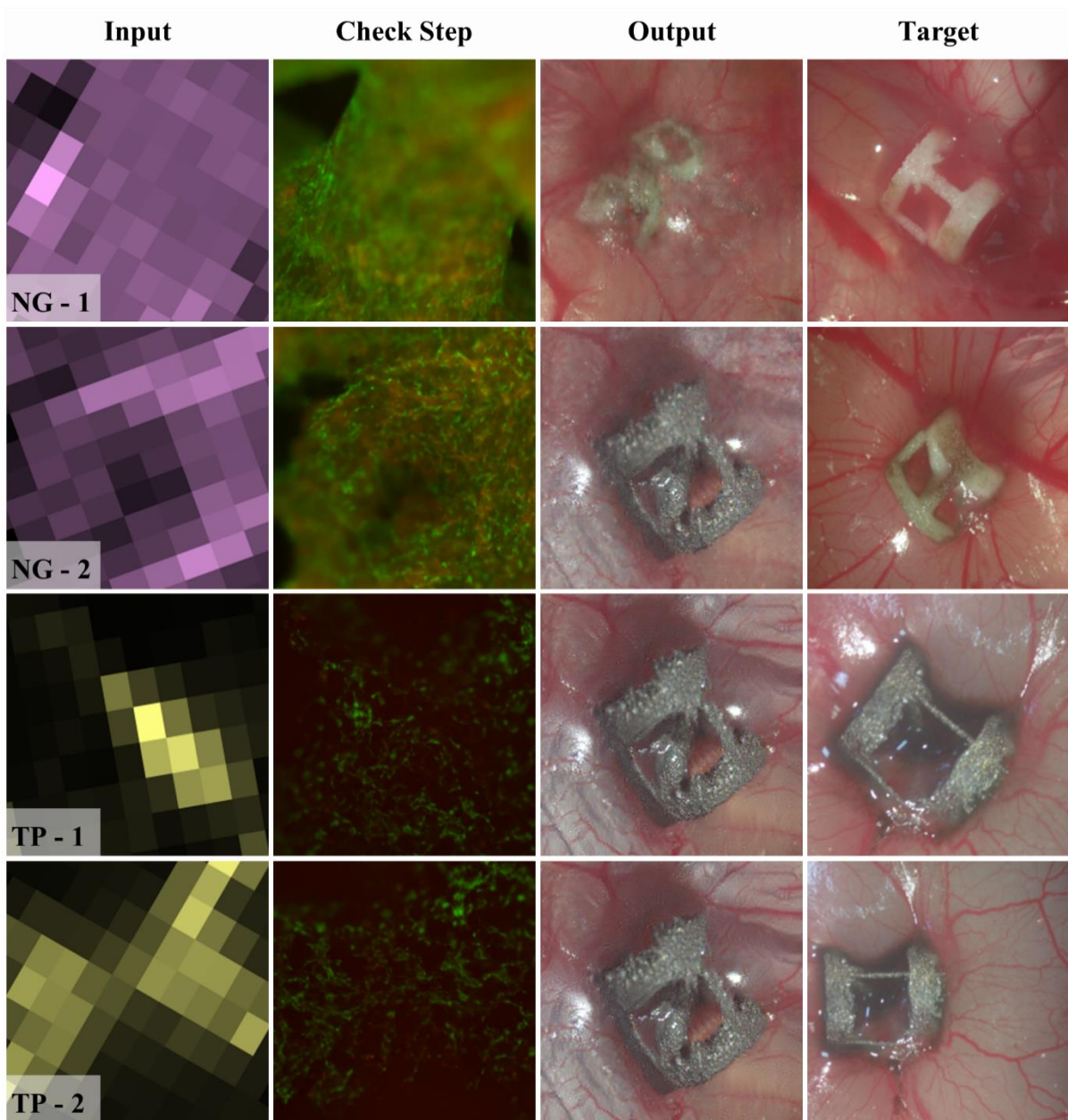


Figure 70: Results of two networks used in a series for biocompatibility prediction. Two variations of each scaffold, NG and TP, were input into the first network through information on material and coating. The output from this network was then input into the second network with no alteration, providing a visual check step in the form of a fluorescent image. The output for the second network, and the network used in a series, was a network generated CAM assay image. These outputs were then compared to target images, which were laboratory obtained CAM assay images for the same scaffold.

The output from the networks, the generated CAM assay images, had mixed results. The least promising output was for NG-2, where the scaffold material was clearly titanium and not nylon. Only NG-1 produced an approximation to a nylon scaffold, but with substantial deformities to the 3D scaffold design. On closer inspection of both the NG-1 and NG-2 outputs, both were heavily pixelated and the vessels are not as clearly defined as the outputs for TP-1 and TP-2. This may show the lack of confidence in the network output for this material and coating combination, which could be due to the variation of possible NG

Predicting Biocompatibility of 3-Dimensional Scaffolds

biocompatibility (Figure 66). The outputs for TP-1 and TP-2, in contrast, were promising as the vessels were clearly defined, and the scaffolds were almost identical to the laboratory obtained target images of plain titanium scaffolds, with minimal structural deformities. Without using any laboratory obtained images, the trained networks had successfully transformed simple information on a scaffold material and (lack of) coating into visually comparable CAM assay images.

An additional test was also performed, to determine if the assumption that two networks in a series would be necessary (with the GPU and computing power available) was correct. A W-Net was therefore trained on the same CAM assay image data as used previously, with an input of the scaffold material and coating. Hyperparameters, number of epochs, and other variables were maintained for an accurate comparison between outputs from a single network and a series network. There was no *in vitro* data used for training the single network, only scaffold information and *in vivo* images (Figure 71).

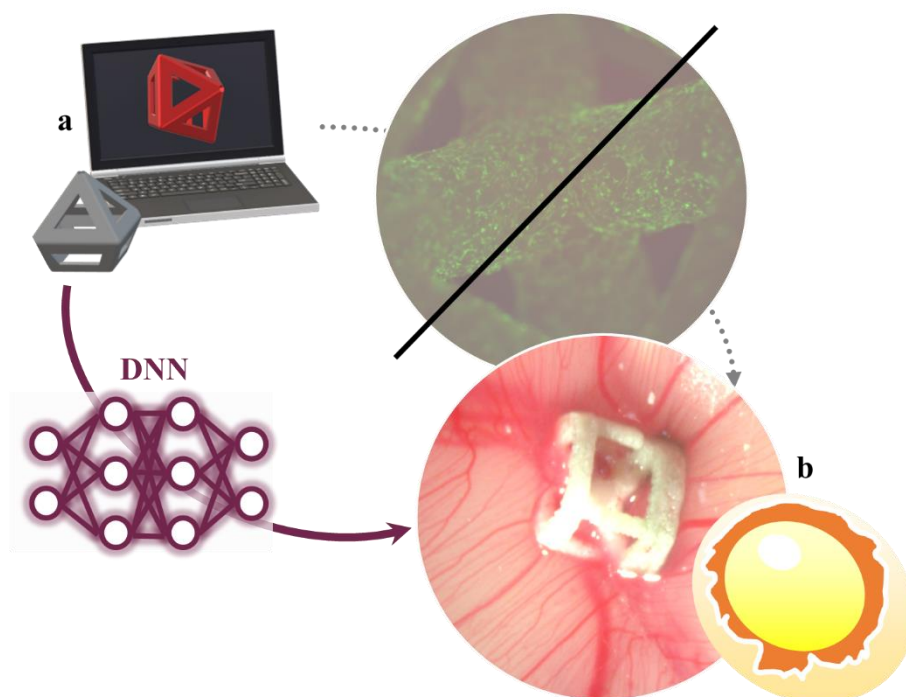


Figure 71: Using a DNN to go from scaffold information to CAM assay images without fluorescent images. The DNN (purple arrow) directly links the scaffold information (a) and CAM assay images (b). Additional data (grey dotted arrow) was not included, and neither was an additional DNN.

One of the similarities between outputs of the single network and series network, shown in Figure 72, was the unchanged TP output, for both TP-1 and TP-2. However, the single network also produced unchanged outputs for NG-1 and NG-2 because there was no variation in scaffold material and coating, which was the only information input to the network. Other

differences include the visual quality of the outputs, such as the ability to extract blood vessels visually and computationally, and the scaffold shape.

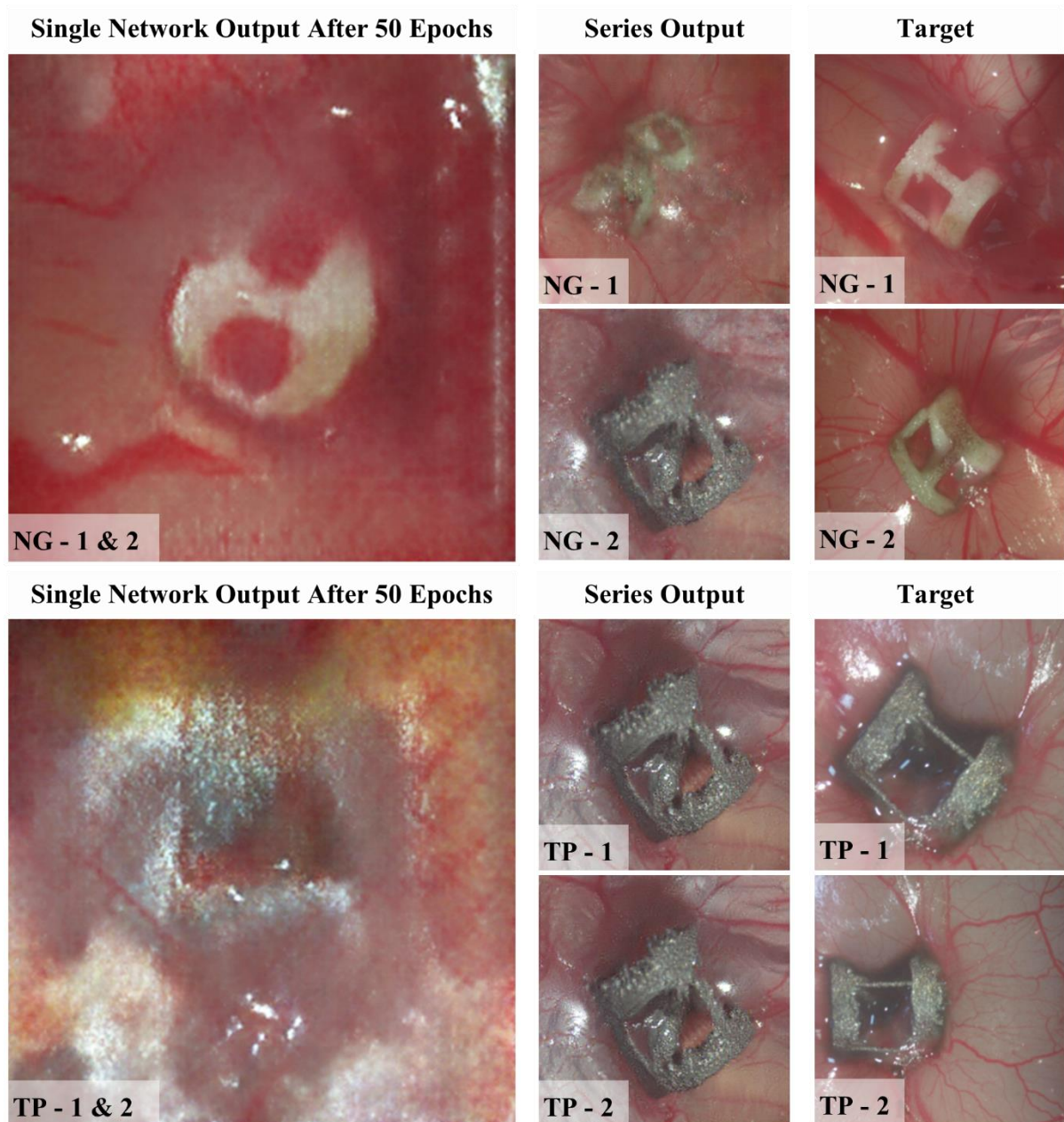


Figure 72: Outputs from a single network trained without in vitro data and a network trained in a series using in vitro data. An input of scaffold materials and coatings, NG and TP, was input into two networks that were trained on the same CAM assay image data. The single network produced identical outputs for both NG and TP variations, The series network produced identical outputs for only TP variations. A target image obtained from laboratory experiments shows what the networks were trained to produce.

While both networks struggled with NG, TP output from the single network was pixelated, provided no clear scaffold structure, and had no extractable vessels for analysis. Comparisons of the quantity of vessels, number of junctions, and mean lacunarity was not possible for a single network, as the output quality was too poor to obtain meaningful results. This also meant that the single network, unlike the networks used in a series, was not fit for purpose, as

Predicting Biocompatibility of 3-Dimensional Scaffolds

no important variable could be obtained from the output images, and it could not be applied for any form of prediction or experimentation.

While the testing input data remained the same for both the single and series networks, the training data was different for the series network than for the single network, as *in vitro* data used to train the series. This implies that there was a correlation between *in vitro* and *in vivo* data that could be extracted by a network for biocompatibility prediction. If this work were to be expanded in the future, it is possible that networks in a series could be used to predict biocompatibility results for more complex biological systems, such as large animal studies or even human *in situ* application. As using multiple networks in a series increased the quality of predicted output than when using a single network alone, this series approach could be applied to other areas of AI-integration where GPU and computation power are not yet sufficient for successful application.

7.5 SUMMARY

The research in this chapter focussed on advancing the work presented in chapter 6, where cell response to physical cues was successfully predicted by a DNN. Instead of a few cells cultured in 2D, this research targeted the cell layer cultured on a 3D scaffold, with DNNs generating predictions of collective cell response to both physical and chemical cues. The predicted biocompatibility of scaffolds was divided into *in vitro* and *in vivo* sections, with one network trained to generate fluorescent images and another trained to generate CAM assay images. These two networks could then be used in a series to predict both *in vitro* and *in vivo* biocompatibility of scaffold material and coating combinations that are potentially untested in a laboratory.

When tested individually, the networks had areas of success and areas that need further improvement. The first network generated fluorescent image predictions capable of only 5% red pixel difference to laboratory images. However, green pixel difference was over 20% for other scaffold types, which could be improved with a larger and more varied training dataset. Similarly, the second network produced images with a mean lacunarity within 10% of the target image values, but also had statistically significant differences in all analysed values for at least one scaffold type.

Regarding the first aim, correction to the uniformity of fluorescence and dataset normalisation may overcome the comparative dimness of generated images. When the lower mean green and red pixel values are ignored, the comparisons between different scaffold

types for output images matched those seen with target images, with highest fluorescence (mean pixel value) for NG14 and lower fluorescence for TP14. Additionally, all results were significantly different ($p < 0.01$) from randomly generated images of green and red pixels, and there were significant differences between different scaffold outputs ($p < 0.05$). This showed success in generating comparative *in vitro* biocompatibility data between scaffolds, which is important for evaluating different scaffold material and coating combinations, and confirmed that the network was not randomly generating fluorescent images.

The second aim was the same as the first, but related to analysis of CAM assay images. This was partially achieved, as the tests showed areas of success and areas where further work was required. A less successful area was the number of junctions, where only TP had no statistically significant differences between the output and target images, but the quantity of vessels was more promising, with both TP and NM showing no differences. Importantly, both TP and NM also showed no significant differences between output and target images for mean lacunarity, which gives a quantitative value to the layout of blood vessels. That means that, for two scaffolds of different material and coating, the network generated realistic blood vessel patterns. As mineral coated nylon was not included in the training dataset, the results are impressive, and suggest that, with enhancements to the dataset, the generated output can be improved.

While the first two aims were not fully achieved, there were some promising results in red pixel difference and mean lacunarity, therefore work continued towards completion of the third aim. Building on these areas of success, the third aim was achieved, as the two networks were successfully used in a series to generate biocompatibility predictions of scaffold material and coating combinations. The outputs for NG scaffolds were not as good as for TP scaffolds, but this could be due to the increased variability in possible biocompatibility outcomes for NG. Additional NG images in the training dataset, among other scaffolds, could improve this result, or further information may be needed to increase network confidence in biocompatibility predictions.

The hypothesis that two independently trained networks can be used in a series to generate CAM assay images for a scaffold of specified material and coating was not proven, as the success of the independent networks (and subsequent series application) was mixed. Overall, this work has implications in the reduction of animal testing, but further work is required before it should be applied in a biomaterial research environment. This project focussed

Predicting Biocompatibility of 3-Dimensional Scaffolds

solely on a select number of scaffold material and coating combinations, and *in vitro* data is limited to fluorescent image data and *in vivo* data is limited to CAM assay images after 8 days of growth. Further studies should include larger and more varied datasets, and animal testing kept to a minimum by collecting images from a range of separate biomaterial research experiments where data collection has been completed.

8 CONCLUSION

There were three aims linked to the hypothesis of this project, with one for each chapter. These were naturally split into two areas – using AI for labour-intensive tasks and using AI for modelling. Each of these chapters contain research that can be used for either area, but chapter 5 focussed more on the former while chapter 6 and chapter 7 focussed more on the latter. The outcome of each aim is summarised below, with the first covered in section 8.1, the second in section 8.2, the third in section 8.3, and finally a note on future application of this research in section 8.4.

8.1 CHAPTER 5

The labelled images were SBFSEM stacks of treated placental tissue, containing several different cell types and acellular material. If deep learning could complete the lengthy task of labelling each 2D slice within a 3D SBFSEM stack, it could enable more quantitative analytical approaches, and eradicate subjective labelling at boundary regions. For a single cell, there was an effective error of less than 1% across all networks used, including those trained on minimal data (less than 20% of the SBFSEM stack). Less data used in training undoubtedly leads to overfitting and lower output quality, but the trade-offs in terms of time management and resources make finding a minimum an important achievement, as a trained network can process an image in under a second. The networks not only identified singular cells, but also distinguished specific cell types and labelled large cellular structures with an average error less than 2%.

8.2 CHAPTER 6

Investigation of every topography is inefficient, as previously explained. Keeping to less than 1000 images when the potential feature space is infinite could have resulted a DNN unable to output reliable approximations. However, application of a DNN trained on only 203 image-pairs generated a model capable of producing statistically significant cell responses. Visually compared to experimentally-obtained data, the network-generated images contained many realistic qualities, including the size, shape, and behaviour of the cells. Mathematically compared, using a hypergeometric distribution probability mass function, one test image managed a statistical significance $P < 0.001$. Additionally, the minimum line separation required for stem cell alignment, $11.7 \pm 1.3 \mu\text{m}$, which was verified by experimental data, was derived only using network-generated data.

Conclusion

8.3 CHAPTER 7

The research in this chapter focussed on advancing the work presented in previous chapters, but targeted the thousands of cells cultured on a 3D scaffold, with DNNs generating predictions of collective cell response to both physical and chemical cues. Unlike in previous chapters, multiple networks were used in a series, with one network trained to generate fluorescent images and another trained to generate CAM assay images. These two networks were then able to predict both *in vitro* and *in vivo* biocompatibility of scaffold material and coating combinations that were potentially untested in a laboratory. The generated *in vitro* biocompatibility data differed between scaffolds, which is an important step for evaluating different scaffold material and coating combinations, and confirmed that the network was not randomly generating fluorescent images. However, the success of the independent networks and subsequent series application was mixed. Overall, this work has implications in the reduction of animal testing, but further work is required before it should be applied in a biomaterial research environment.

8.4 A NOTE ON FUTURE RESEARCH

This project started with the hope of using AI for both labour-intensive tasks and data modelling. While both have been achieved for tissue engineering purposes, only the former was applied for placental research. However, since the completion of this postgraduate project, the research into this area has continued, with supervisor Ben Mills and collaborator Rohan Lewis (both University of Southampton) advancing the methods used in Chapter 3 for SBFSEM image generation. Initial results have been promising, with successful image generation based on i.e. the size and positioning of singular or multiple different cells. This work could revolutionise placental modelling, with implications to both prenatal and lifelong medical research. Some of this work has recently been submitted to BIOSTEC 2022 (James A Grant-Jacob, Ben Mills, Benita Mackay, Rohan Lewis, Bram Sengers. “Modelling of 3D Placental Cell Features using Deep Learning”). For more information please contact Ben Mills (bm602@orc.soton.ac.uk).

8.5 RESEARCH CAVEATS

While the research undertaken throughout this project has potential in future application of DNNs across many areas, the scope of specific results are less broad. Less than two thousand images from three placenta SBFSEM stacks were used for the training and testing of cell labelling. Training on additional materials would be required for cell labelling to be achieved in other 3D imaging techniques. Similarly, the modelling of stem cells was accurate on a

system of one topographical material laser-machined by a single laser set-up, but results for stem cell prediction on, for example, titanium would require additional training materials beyond the roughly two hundred used here. The system for predicting biocompatibility is limited in the same way by the low quantity of images within the training dataset. Each of these networks would likely not produce such impressive outputs when provided inputs from different data sources (until they are trained further on datasets with greater variation and a larger number of feature combinations).

8.6 FINAL CONCLUSION

DNNs were capable of labelling, modelling, and predicting cell biocompatibility. This can lead to increases in efficiency of 3D-image processing, which could help researchers analyse how cellular and acellular structures in placental tissue affects lifelong health. DNNs may also potentially help increase understanding of stem cell behaviour through accurate modelling of stem cell response to biophysical and biochemical cues, which may also lead to a reduction in animal testing. If this research is progressed, DNN application to regenerative medicine could one day help reduce levels of chronic poor health in later life.

APPENDIX

STRO-1 ISOLATION METHOD AND PROTOCOL

The procedure, step by step, of isolating Stro-1 positive antibody cells from bone marrow.

Bone Marrow Cell Isolation Passage 0 (P0) – 1 hour

- Remove marrow tubes from the fridge
 - Label any tubes with ALL marrow sample details and date
- Open marrow tube and spare universals
 - Pipette 5-7 mL of α MEM into marrow tube and shake vigorously
 - Tubes don't have a proper seal on them. Wrap up well with tissue when shaking
 - Transfer liquid into universal
 - Repeat several times, using small pipettes to crunch the bone if needed. Continue until bone chips are pinky white
 - Discard remaining marrow into virkon
 - Add virkon to tube and discard in blue box
- Centrifuge tube at 1200rpm (400G) for 5 minutes
- Carefully discard supernatant and fat into virkon without risking the pellet
- Wash the cells in α MEM
 - Wash 1
 - Re-suspend pellet in 10-20 mL of α MEM
 - Centrifuge tube at 1200rpm for 5 minutes
 - Discard supernatant and fat (carefully)
 - Wash 2
 - Repeat
- Re-suspend the pellet in 25 mL of α MEM per intended Lymphoprep
 - If sample too blood rich than separation will not work
 - For dense samples re-suspend the pellet in 50 mL of α MEM and prepare two tubes of Lymphoprep after straining
- Filter the cell suspension through a cell strainer

Appendix

- Open the cell strainer and HOLD BY THE EDGE
- Place on top of new 50 mL tube
- Pipette cell suspension through strainer, gently scraping bottom of filter to prevent clogging
 - Safest to pipette by hand and slowly. Splashback is likely otherwise.

Lymphoprep with Human Marrow Cells – 2 hours

- Pipette 20 mL Lymphoprep into a freshly labelled tube
 - 20 mL of Lymphoprep per 10 mL of cell suspension
 - If have large bloody sample/lots of cells, dilute cell suspension up to 20 mL and split into 2 tubes and maintain above ratio
- Using a 3 mL disposable pipette, tilt the tube, slowly and gently layer the cell suspension over the surface on the Lymphoprep and DO NOT RUSH
 - Build up a layer of cell suspension on top of Lymphoprep by holding the pipette close to the side of the tube and releasing a slow stream of the cell suspension down the side of the tube.
- Centrifuge at 2200rpm (800G) for 20 minutes at 18°C with the brake off.
 - This will take 40 minutes – 20 minutes spinning and 20 minutes slowing down to a stop
 - At the end of the spinning there should be a distinct layer of media at the top and then the BMNC (Bone Marrow Mononuclear Cells) layer beneath at interface with Lymphoprep
 - Called the “buffy coat”. Red blood cells and other granulocytes will be pelleted at the bottom.
- Use pipette to carefully remove all cells from the interface “buffy coat” layer and transfer to a new tube. Try not to take the Lymphoprep – taking media is better.
- Make up to 50 mL with α MEM to wash and pellet the cells
 - Wash 1
 - Centrifuge tube at 1200rpm for 4 minutes
 - Carefully discard supernatant into virkon
 - Re-suspend the pellet in 50 mL of α MEM
 - Wash 2
 - Centrifuge tube at 1200rpm for 4 minutes

- Pour off supernatant into virkon

MACS (Magnetically Activated Cell Sorting) Stro-1 Isolation – over 2 hours

- Re-suspend in 2 mL of MACS Blocking Buffer
- Incubate for 30 minutes at 4°C
 - MIX REGULARLY or use the rocker in the cold room
 - Get a tube of Stro-1-hybridoma supernatant
 - Collect from -80°C freezer and thaw in fridge
 - Sign out number of vials used in the black folder (top drawer next to the sink)
- Take tube from the fridge and make up to 10 mL with MACS buffer (chilled)
 - Remember to put MACS buffer back into fridge during incubations
 - Centrifuge at 1200rpm for 5 minutes
- Re-suspend in 0.5 mL aliquot of Stro-1-hybridoma supernatant (use neat)
- Incubate for 30 minutes at 4°C (mix regularly)
- Take tube from the fridge and wash the cells three times with MACS buffer (chilled)
 - Add <20 mL of MACS buffer
 - Wash 1
 - Centrifuge tube at 1200rpm for 5 minutes
 - Carefully discard supernatant into virkon
 - Add <20 mL of MACS buffer
 - Wash 2
 - Centrifuge tube at 1200rpm for 5 minutes
 - Carefully discard supernatant into virkon
 - Re-suspend the pellet in 5 mL MACS buffer
 - Cell Count for determination of number of MACS IgM microbeads required
 - Dilute small volume from cell suspension in a microfuge tube (25x or 50x)
 - Mix 1:1 with Trypan Blue
 - Count cells

Appendix

- Wash 3
 - Add <20 mL of MACS buffer
 - Centrifuge at 1200rpm for 5 minutes
 - Discard supernatant in virkon
- Re-suspend in appropriate volumes of MACS Rat a-mouse IgM Microbeads and MACS buffer based on the cell count:
 - 80 μ L per 10⁷ cells of MACS buffer.
 - 20 μ L per 10⁷ cells of MACS Rat a-mouse IgM Microbeads
 - 200 μ L of MACS Rat a-mouse IgM Microbeads is the maximum that can be used for 1 column, for higher cells numbers divide accordingly.
- Incubate in the Microbeads for 15 minutes at 4°C (mix regularly)
 - Collect the Magnet stand and columns and spray into hood ready
- Wash the cells three times with MACS buffer (chilled)
 - Wash 1
 - Add <20 mL of MACS buffer
 - Centrifuge tube at 1200rpm for 5 minutes
 - Carefully discard supernatant into virkon
 - Wash 2
 - Repeat
 - Wash 3
 - Repeat
- Re-suspend pellet in 2 mL MACS buffer (use 4 mL for 2x10⁸ cells)

MACS Magnetic Separation – over 1 hour

- Remove the column and place in magnetic holder, push to secure into bracket.
 - Leave the plunger in the sterile packet until required in the last step
- Label universal tubes with the marrow ID and (-) and (+). Place the (-) tube in a rack underneath the columns to catch liquid that drips through
- Add 3 mL MACS buffer to wet the column
 - Let it drip through until the level reaches the top of the black

- Add 2 mL cell suspension to the column with a pastette
 - Drip slowly and avoid bubbles
- Add 3 x 3 mL of MACS buffer to wash out all (-) cells into collecting tube
 - Remove (-) tube
- Remove the column from the bracket and place in (+) tube
- Add 5 mL MACS buffer to the column
- Place plunger at top of tube and press down firmly and swiftly to push (+) cells out
- Wash twice in plain α MEM
 - Wash 1
 - Add <20 mL of α MEM
 - Centrifuge at 1200rpm for 5 minutes
 - Discard supernatant in virkon
 - Cell Count
 - Resuspend pellet in 5 mL MACS buffer
 - Mix ie 10 μ L 1:1 with Trypan Blue in microfuge tube
 - Wash 2
 - Add <20 mL of α MEM
 - Centrifuge at 1200 rpm for 5 minutes
 - Discard supernatant in virkon
- Re-suspend in COMPLETE α MEM (+ 10% FCS + 1% P/S) and seed into 12 well plate
 - Additional 1 mL of FCS always helps
 - If cell count too sparse seed into flask as they will not grow (not everything will stick down)
 - Leave in culture for 1 week before washing with PBS and changing the media
- Incubate cells under normal conditions (37°C, 5% CO₂) for minimum three days

Change the media and transfer to machined substrate when approaching confluence.

REFERENCES

1. Li, J., et al., *Biophysical and Biochemical Cues of Biomaterials Guide Mesenchymal Stem Cell Behaviors*. Front Cell Dev Biol, 2021. **9**: p. 640388.
2. Harris, S.E., et al., *Pericytes on placental capillaries in terminal villi preferentially cover endothelial junctions in regions furthest away from the trophoblast*. Placenta, 2021. **104**: p. 1-7.
3. Miyachi, A. and J. Jenoptik, *Femtosecond Laser Processing of Metal and Plastics*. MANUFACTURING & PROTOTYPING, 2015.
4. Jain, R. *3 Types of Gradient Descent Algorithms for Small & Large Data Sets*. 2017; Available from: <https://www.hackerearth.com/blog/developers/3-types-gradient-descent-algorithms-small-large-data-sets/>.
5. Isola, P., et al., *Image-to-Image Translation with Conditional Adversarial Networks*. arXiv, 2016: p. arXiv:1611.07004v3.
6. Dubochet Center for Imaging. *Structural Studies of Neurodegeneration*. [cited 2021 09/03/2021]; Available from: <https://www.c-cina.org/stahlberg/research/neuro/>.
7. Waldrop, M.M., *News Feature: What are the limits of deep learning?* Proc Natl Acad Sci U S A, 2019. **116**(4): p. 1074-1077.
8. Xia, X. and B. Kulis, *W-Net: A Deep Model for Fully Unsupervised Image Segmentation*. arXiv, 2017: p. arXiv:1711.08506v1.
9. Teixeira, A.I., et al., *Epithelial contact guidance on well-defined micro- and nanostructured substrates*. J Cell Sci, 2003. **116**(Pt 10): p. 1881-92.
10. Choubisa, T., et al. *Direction and gender classification using convolutional neural network for side-view images captured from a monitored trail*. in *2017 Fourth International Conference on Image Information Processing (ICIIP)*. 2017.
11. Mackay, B.S., et al., *Modeling adult skeletal stem cell response to laser-machined topographies through deep learning*. Tissue and Cell, 2020. **67**: p. 101442.
12. McNamara, L.E., et al., *Nanotopographical control of stem cell differentiation*. J Tissue Eng, 2010. **2010**: p. 120623.
13. Heath, D.J., et al., *Machine learning for 3D simulated visualization of laser machining*. Optics Express, 2018. **26**(17): p. 21574-21584.
14. Mangir, N., et al., *Using ex Ovo Chick Chorioallantoic Membrane (CAM) Assay To Evaluate the Biocompatibility and Angiogenic Response to Biomaterials*. ACS Biomater Sci Eng, 2019. **5**(7): p. 3190-3200.
15. Kumar, G., et al., *The determination of stem cell fate by 3D scaffold structures through the control of cell shape*. Biomaterials, 2011. **32**(35): p. 9188-96.
16. Dalby, M.J., N. Gadegaard, and R.O. Oreffo, *Harnessing nanotopography and integrin-matrix interactions to influence stem cell fate*. Nat Mater, 2014. **13**(6): p. 558-69.
17. Marshall, K.M., J.M. Kanczler, and R.O. Oreffo, *Evolving applications of the egg: chorioallantoic membrane assay and ex vivo organotypic culture of materials for bone tissue engineering*. J Tissue Eng, 2020. **11**: p. 2041731420942734.
18. Wolf, A.D. and B. Pleger, *Burden of major musculoskeletal conditions*. Bull World Health Organ, 2003. **81**: p. 646-656.
19. Ackermann, I.N., et al., *Substantial rise in the lifetime risk of primary total knee replacement surgery for osteoarthritis from 2003 to 2013: an international, population-level analysis*. Osteoarthritis and Cartilage, 2017. **25**(4): p. 455-461.
20. Inacia, M.C.S., et al., *Increase in Total Joint Arthroplasty Projected from 2014 to 2046 in Australia: A Conservative Local Model With International Implications*. Clin Orthop Relat Res, 2017. **475**(8): p. 2130-2137.
21. Bianco, P. and P. Gehrom Robey, *Stem cells in tissue engineering*. Nature, 2001. **414**: p. 118-121.
22. Cooper, C., P. Mitchell, and J.A. Kanis, *Breaking the fragility fracture cycle*. Osteoporos Int, 2011. **22**(7): p. 2049-50.
23. Hilgsmann, M., et al., *Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO)*. Semin Arthritis Rheum, 2013. **43**(3): p. 303-13.
24. Ackerman, I.N., et al., *Substantial rise in the lifetime risk of primary total knee replacement surgery for osteoarthritis from 2003 to 2013: an international, population-level analysis*. Osteoarthritis Cartilage, 2017. **25**(4): p. 455-461.
25. Inacio, M.C.S., et al., *Increase in Total Joint Arthroplasty Projected from 2014 to 2046 in Australia: A Conservative Local Model With International Implications*. Clin Orthop Relat Res, 2017. **475**(8): p. 2130-2137.
26. Montoya, C., et al., *On the road to smart biomaterials for bone research: definitions, concepts, advances, and outlook*. Bone Res, 2021. **9**(1): p. 12.
27. Kuznetsov, S.A., et al., *Single-colony derived strains of human marrow stromal fibroblasts form bone after transplantation in vivo*. J Bone Miner Res, 1997. **12**(9): p. 1335-47.
28. Muraglia, A., R. Cancedda, and R. Quarto, *Clonal mesenchymal progenitors from human bone marrow differentiate in vitro according to a hierarchical model*. J Cell Sci, 2000. **113** (Pt 7): p. 1161-6.
29. Bianco, P., et al., *The meaning, the sense and the significance: translating the science of mesenchymal stem cells into medicine*. Nat Med, 2013. **19**(1): p. 35-42.
30. Bettinger, C.J., R. Langer, and J.T. Borenstein, *Engineering substrate topography at the micro- and nanoscale to control cell function*. Angew Chem Int Ed Engl, 2009. **48**(30): p. 5406-15.

References

31. Bianco, P., et al., *Bone marrow stromal stem cells: nature, biology, and potential applications*. Stem Cells, 2001. **19**(3): p. 180-92.
32. Curtis, A.S.G. and M. Varde, *Control of Cell Behaviour: Topological Factors*. Journal of the National Cancer Institute, 1964. **33**(1): p. 15-26.
33. Stevens, M.M. and J.H. George, *Exploring and Engineering the Cell Surface Interface*. Science, 2005. **310**(5751): p. 1135-1138.
34. Curtis, A. and C. Wilkinson, *Nanototechniques and approaches in biotechnology*. Trends in Biotechnology 2001. **19**(3): p. 97-101.
35. Dalby, M.J., N. Gadegaard, and R.O.C. Oreffo, *Harnessing nanotopography and integrin–matrix interactions to influence stem cell fate*. Nature Materials, 2014. **13**: p. 558-569.
36. Heath, D.J., et al., *Closed-loop corrective beam shaping for laser processing of curved surfaces*. Journal of Micromechanics and Microengineering, 2018. **28**(12).
37. Heath, D.J., et al., *Sub-diffraction limit laser ablation via multiple exposures using a digital micromirror device*. Applied Optics, 2017. **56**(22): p. 6398-6404.
38. Mills, B., et al. *Image-based monitoring of high-precision laser machining via a convolutional neural network*. in *SPIE Photonics West*. 2019. San Francisco.
39. Lewis, R.M., J.K. Cleal, and M.A. Hanson, *Review: Placenta, evolution and lifelong health*. Placenta, 2012. **33** **Suppl**: p. S28-32.
40. Mackay, B.S., et al., *The future of bone regeneration: integrating AI into tissue engineering*. Biomed Phys Eng Express, 2021.
41. Cooper, C., P. Mitchell, and J.A. Kanis, *Breaking the fragility fracture cycle*. Osteoporosis International, 2011. **22**(7): p. 2049-2050.
42. Hilgsmann, M., C. Cooper, and N. Arden, *Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO)*. Semin Arthritis Rheum, 2013. **43**: p. 303-313.
43. Bjarnason, I., et al., *Side effects of nonsteroidal anti-inflammatory drugs on the small and large intestine in humans*. Gastroenterology, 1993. **104**(6): p. 1832-47.
44. Rainsford, K.D., *Profile and mechanisms of gastrointestinal and other side effects of nonsteroidal anti-inflammatory drugs (NSAIDs)*. Am J Med, 1999. **107**(6A): p. 27S-35S; discussion 35S-36S.
45. Rainsford, K.D., et al., *Effects of the NSAIDs meloxicam and indomethacin on cartilage proteoglycan synthesis and joint responses to calcium pyrophosphate crystals in dogs*. Vet Res Commun, 1999. **23**(2): p. 101-13.
46. Cartagena Farias, J., et al., *Prescribing patterns in dependence forming medicines*. 2017, NatCen: London.
47. da Costa, B.R., et al., *Oral or transdermal opioids for osteoarthritis of the knee or hip*. Cochrane Database Syst Rev., 2014. **17**(9): p. CD003115.
48. (NIBIB), N.I.o.B.I.a.B. *Tissue Engineering and Regenerative Medicine*. Science Education [cited 2020 05/12/2020]; Available from: <https://www.nibib.nih.gov/science-education/science-topics/tissue-engineering-and-regenerative-medicine>.
49. Stevens, M.M., *Biomaterials for bone tissue engineering*. Materials Today, 2008. **11**(5): p. 18-25.
50. Rawal, B.R., A. Yadav, and V. Pare, *Life estimation of knee joint prosthesis by combined effect of fatigue and wear*. Procedia Technology, 2016. **23**: p. 60-67.
51. Tang, D., et al., *Biofabrication of bone tissue: approaches, challenges and translation for bone regeneration*. Biomaterials, 2016. **83**: p. 363-82.
52. Armstrong, J.P.K. and M.M. Stevens, *Emerging Technologies for Tissue Engineering: From Gene Editing to Personalized Medicine*. Tissue Eng Part A, 2019. **25**(9-10): p. 688-692.
53. Oryan, A.A., S.; Moshiri, A.; Maffulli, N, *Bone regenerative medicine: classic options, novel strategies, and future directions*. Journal of Orthopaedic Surgery and Research, 2014. **9**(18).
54. Clarke, B., *Normal bone anatomy and physiology*. Clin J Am Soc Nephrol, 2008. **3** **Suppl 3**: p. S131-9.
55. Petite, H., et al., *Tissue-engineered bone regeneration*. Nat Biotechnol, 2000. **18**(9): p. 959-63.
56. Bueno, E.M. and J. Glowacki, *Cell-free and cell-based approaches for bone regeneration*. Nat Rev Rheumatol, 2009. **5**(12): p. 685-97.
57. Dimitriou, R., et al., *Bone regeneration: current concepts and future directions*. BMC Med, 2011. **9**: p. 66.
58. Gong, T., et al., *Nanomaterials and bone regeneration*. Bone Res, 2015. **3**: p. 15029.
59. Grayson, W.L., et al., *Stromal cells and stem cells in clinical bone regeneration*. Nat Rev Endocrinol, 2015. **11**(3): p. 140-50.
60. Gaharwar, A.K., I. Singh, and A. Khademhosseini, *Engineered biomaterials for in situ tissue regeneration*. Nature Reviews Materials, 2020. **5**: p. 686-705.
61. Dawson, J.I. and R.O. Oreffo, *Bridging the regeneration gap: stem cells, biomaterials and clinical translation in bone tissue engineering*. Arch Biochem Biophys, 2008. **473**(2): p. 124-31.
62. Lin, G., et al., *Tissue distribution of mesenchymal stem cell marker Stro-1*. Stem Cells Dev, 2011. **20**(10): p. 1747-52.
63. Oreffo, R.O.C., C.; Mason, C.; Clements, M., *Mesenchymal Stem Cells: Lineage, Plasticity, and Skeletal Therapeutic Potential*. Stem Cell Reviews, 2005. **1**: p. 169-178.
64. Maman, D. *The FDA Takes Action to Protect Patients from Risk of Certain Textured Breast Implants; Requests Allergan Voluntarily Recall Certain Breast Implants and Tissue Expanders from the Market: FDA Safety Communication*. 2019 [cited 2020 19/08/2020]; Available from: <https://www.fda.gov/news-events/press-announcements/fda-takes-action-protect-patients-risk-certain-textured-breast-implants-requests-allergan>.

65. Johnson, L., et al., *Breast implant associated anaplastic large cell lymphoma: The UK experience. Recommendations on its management and implications for informed consent*. Eur J Surg Oncol, 2017. **43**(8): p. 1393-1401.
66. Jones, J.L., et al., *Breast implant-associated anaplastic large cell lymphoma (BIA-ALCL): an overview of presentation and pathogenesis and guidelines for pathological diagnosis and management*. Histopathology, 2019. **75**(6): p. 787-796.
67. Rose, F.R. and R.O. Oreffo, *Bone tissue engineering: hope vs hype*. Biochem Biophys Res Commun, 2002. **292**(1): p. 1-7.
68. Dziki, J.L., et al., *Immunomodulation and Mobilization of Progenitor Cells by Extracellular Matrix Bioscaffolds for Volumetric Muscle Loss Treatment*. Tissue Eng Part A, 2016. **22**(19-20): p. 1129-1139.
69. Sadtler, K., et al., *Developing a pro-regenerative biomaterial scaffold microenvironment requires T helper 2 cells*. Science, 2016. **352**(6283): p. 366-70.
70. Dalby, M.J., et al., *The control of human mesenchymal cell differentiation using nanoscale symmetry and disorder*. Nat Mater, 2007. **6**(12): p. 997-1003.
71. Ma, Q.L., et al., *Improved implant osseointegration of a nanostructured titanium surface via mediation of macrophage polarization*. Biomaterials, 2014. **35**(37): p. 9853-9867.
72. Gilchrist, C.L., et al., *Micro-scale and meso-scale architectural cues cooperate and compete to direct aligned tissue formation*. Biomaterials, 2014. **35**(38): p. 10015-24.
73. Vogel, V. and M. Sheetz, *Local force and geometry sensing regulate cell functions*. Nat Rev Mol Cell Biol, 2006. **7**(4): p. 265-75.
74. Gamboa, J.R., et al., *Linear fibroblast alignment on sinusoidal wave micropatterns*. Colloids Surf B Biointerfaces, 2013. **104**: p. 318-25.
75. Hulander, M., et al., *Gradients in surface nanotopography used to study platelet adhesion and activation*. Colloids Surf B Biointerfaces, 2013. **110**: p. 261-9.
76. Koons, G.L., M. Diba, and A.G. Mikos, *Materials design for bone-tissue engineering*. Nature Reviews Materials, 2020. **5**(8): p. 584-603.
77. Yunus Basha, R., T.S. Sampath Kumar, and M. Doble, *Design of biocomposite materials for bone tissue regeneration*. Mater Sci Eng C Mater Biol Appl, 2015. **57**: p. 452-63.
78. Ventre, M. and P.A. Netti, *Engineering Cell Instructive Materials To Control Cell Fate and Functions through Material Cues and Surface Patterning*. ACS Appl Mater Interfaces, 2016. **8**(24): p. 14896-908.
79. Ventre, M., F. Causa, and P.A. Netti, *Determinants of cell-material crosstalk at the interface: towards engineering of cell instructive materials*. J R Soc Interface, 2012. **9**(74): p. 2017-32.
80. Christo, S.N., et al., *The Role of Surface Nanotopography and Chemistry on Primary Neutrophil and Macrophage Cellular Responses*. Adv Healthc Mater, 2016. **5**(8): p. 956-65.
81. Biggs, M.J., et al., *The use of nanoscale topography to modulate the dynamics of adhesion formation in primary osteoblasts and ERK/MAPK signalling in STRO-1+ enriched skeletal stem cells*. Biomaterials, 2009. **30**(28): p. 5094-103.
82. Goriainov, V., et al., *Harnessing Nanotopography to Enhance Osseointegration of Clinical Orthopedic Titanium Implants-An in Vitro and in Vivo Analysis*. Front Bioeng Biotechnol, 2018. **6**: p. 44.
83. Greer, A.I.M., et al., *Nanopatterned Titanium Implants Accelerate Bone Formation In Vivo*. ACS Appl Mater Interfaces, 2020. **12**(30): p. 33541-33549.
84. Plank Collaboration, *Plank 2018 results. VI. Cosmological parameters (v2)*. arXiv, 2018: p. arXiv:1807.06209v2 [astro-ph.CO].
85. Curtis, A.S. and M. Varde, *Control of Cell Behavior: Topological Factors*. J Natl Cancer Inst, 1964. **33**: p. 15-26.
86. Stevens, M.M. and J.H. George, *Exploring and engineering the cell surface interface*. Science, 2005. **310**(5751): p. 1135-8.
87. Curtis, A. and C. Wilkinson, *Nanotechnology and approaches in biotechnology*. Trends Biotechnol, 2001. **19**(3): p. 97-101.
88. Clark, P., et al., *Cell guidance by ultrafine topography in vitro*. J Cell Sci, 1991. **99** (Pt 1): p. 73-7.
89. Khang, D., et al., *Role of subnano-, nano- and sunimicron-surface features on osteoblast differentiation of bone marrow mesenchymal stem cells*. Biomaterials, 2012. **33**(26): p. 5997-6007.
90. Guida, L., et al., *Biological response of human bone marrow mesenchymal stem cells to fluoride-modified titanium surfaces*. Clin Oral Implants Res, 2010. **21**(11): p. 1234-1241.
91. Zhang, W., et al., *The synthetic effect of hierarchical micro/nano-topography and bioactive ions for enhanced osseointegration*. Biomaterials, 2013. **34**(13): p. 3184-3195.
92. Coelho, P., et al., *Nanometer-scale features on micrometer-scale surface texturing: a bone histological, gene expression, and nanomechanical study*. Bone, 2014. **65**: p. 25-36.
93. Galli, S., et al., *Local release of magnesium from mesoporous TiO2 coatings stimulates the peri-implant expression of osteogenic markers and improves osteoconductivity in vivo*. Acta Biomater, 2014. **10**(12): p. 5193-5201.
94. Ren, N., et al., *Nanostructured titanate with different metal ions on the surface of metallic titanium: a facile approach for regulation of rBMSCs fate on titanium implants*. Small., 2014. **10**(15): p. 3196-3180.
95. Yang, J., et al., *Nanotopographical induction of osteogenesis through adhesion, bone morphogenic protein signaling, and regulation of microRNAs*. ACS Nano, 2014. **8**(10): p. 9941-9953.
96. Kim, H., et al., *Surface Engineering of Nanostructured Titanium Implants with Bioactive Ions*. Journal of Dental Research, 2016. **95**(5): p. 558-565.
97. Park, J.-W., S.-H. Han, and T. Hanawa, *Effects of Surface Nanotopography and Calcium Chemistry of Titanium Bone Implants on Early Blood Platelet and Macrophage Cell Function*. BioMed Research International, 2018. **2018**: p. Article ID 1362958, 10 pages.

References

98. Saidak, Z., et al., *Strontium ranelate rebalances bone marrow adipogenesis and osteoblastogenesis in senescent osteopenic mice through NFATc/Mad and Wnt signaling*. *Aging Cell*, 2012. **11**(3): p. 467-474.
99. Saidak, Z. and P. Marie, *Strontium signalling: molecular mechanisms and therapeutic implications in osteoporosis*. *Pharmacol Ther.*, 2012. **136**(2): p. 216-226.
100. Dalby, M.J., et al., *The control of human mesenchymal cell differentiation using nanoscale symmetry and disorder*. *Nature Materials*, 2007. **6**: p. 997-1003.
101. Scaglione, S., et al., *Order versus Disorder: in vivo bone formation within osteoconductive scaffolds*. *Scientific Reports*, 2012. **2**(274): p. 1-6.
102. Allan, C., et al., *Osteoblast response to disordered nanotopography*. *Journal of Tissue Engineering*, 2018. **9**: p. 1-7.
103. Gong, T., et al., *Nanomaterials and bone regeneration*. *Bone Research*, 2015. **3**: p. 15029.
104. Christo, S.N., et al., *The Role of Surface Nanotopography and Chemistry on Primary Neutrophil and Macrophage Cellular Responses*. *Advanced Healthcare Materials*, 2016. **5**: p. 956-965.
105. Goodman, S.L., P.A. Sims, and R.M. Albrecht, *Three-dimensional extracellular matrix textured biomaterials*. *Biomaterials*, 1996. **17**(21): p. 2087-2095.
106. Bettinger, C.J., R. Langer, and J.T. Borenstein, *Engineering Substrate Topography at the Micro- and Nanoscale to Control Cell Function*. *Angewandte Chemie International Edition*, 2009. **48**(30): p. 5406-5415.
107. Ben-Shlomo, I., et al., *Signalling Receptome: A genomic and evolutionary perspective of plasma membrane receptors involved in signal transduction*. *Science's STKE*, 2003. **2003**(187): p. re9.
108. Wolf, K., et al., *Amoeboid shape change and contact guidance: T-lymphocyte crawling through fibrillar collagen is independent of matrix remodeling by MMPs and other proteases*. *Blood*, 2003. **102**(9): p. 3262-3269.
109. Loesberg, W.A., et al., *The threshold at which substrate nanogroove dimensions may influence fibroblast alignment and adhesion*. *Biomaterials*, 2007. **28**(27): p. 3944-3951.
110. Abrams, G.A., et al., *Nanoscale topography of the basement membrane underlying the corneal epithelium of the rhesus macaque*. *Cell Tissue Res.*, 2000. **299**: p. 39-46.
111. Dalby, M.J., et al., *Investigating the limits of filopodial sensing: a brief report using SEM to image the interaction between 10 nm high nano-topography and fibroblast filopodia*. *Cell Biology International*, 2004. **28**(3): p. 229-239.
112. Clark, P., et al., *Cell guidance by ultrafine topography in vitro*. *Journal of Cell Science*, 1991. **99**: p. 73-77.
113. Dunn, G.A. and H.J. P., *A new hypothesis of contact guidance in tissue cells*. *Exp Cell Res*, 1976. **101**: p. 1-14.
114. Izadpanahi, M., et al., *Nanotopographical cues of electrospun PLLA efficiently modulate noncodingRNA network to osteogenic differentiation of mesenchymal stem cells during BMP signaling pathway*. *Materials Science & Engineering*, 2018. **93**: p. 686-703.
115. Christo, S.N., et al., *The role of surface nanotopography and chemistry on primary neutrophil and macrophage cellular response*. *Adv. Healthcare Mat.*, 2016. **26**(5): p. 956-965.
116. Yim, E.K.F., S.W. Pang, and K.W. Leong, *Synthetic nanostructures inducing differentiation of human mesenchymal stem cells into neuronal lineage*. *Experimental Cell Research*, 2007. **313**(9): p. 1820-1829.
117. Engler, A.J., et al., *Matrix elasticity directs stem cell lineage specification*. *Cell*, 2006. **126**: p. 677-689.
118. Huebsch, N. and e. al, *Harnessing traction-mediated manipulation of the cell/matrix interface to control stem-cell fate*. *Nature Mater*, 2010. **9**: p. 518-526.
119. Kilian, K.A., et al., *Geometric cues for directing the differentiation of mesenchymal stem cells*. *Proc. Natl Acad. Sci. USA*, 2010. **107**: p. 4872-4877.
120. McBeath, R., et al., *Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment*. *Dev. Cell*, 2004. **6**: p. 483-495.
121. Dalby, M.J., et al., *Nanotopographical stimulation of mechanotransduction and changes in interphase centromere positioning*. *J Cell Biochem*, 2007. **100**(2): p. 326-338.
122. Dalby, M.J., et al., *Nanomechanotransduction and interphase nuclear organisation influence on genomic control*. *J. Cell. Biochem*, 2007. **102**(5): p. 1234-1244.
123. Tsimbouri, P., et al., *Nanotopographical Effects on Mesenchymal Stem Cell Morphology and Phenotype*. *J. Cell. Biochem*, 2013. **115**: p. 380-390.
124. McNamara, L. and T. Sjöström, *Skeletal stem cell physiology on functionally distinct titania nanotopographies*. *Biomaterials*, 2011. **32**: p. 7403-7410.
125. Khan, M., et al., *The epidemiology of failure in total knee arthroplasty: avoiding your next revision*. *Bone Joint J.*, 2016. **98-b**: p. 105-112.
126. *13th Annual Report 2016*. 2016, NJR.
127. Woodruff, M.A., et al., *Bone tissue engineering: from bench to bedside*. *Materials Today*, 2012. **15**(10): p. 430-435.
128. Bose, S., S. Vahabzadeh, and A. Bandyopadhyay, *Bone tissue engineering using 3D printing*. *Materials Today*, 2013. **16**(12): p. 496-504.
129. Rodgers, P., *What diffraction limit?* *Nature Nanotechnology*, 2009.
130. Andrew, T.L., H.Y. Tsai, and R. Menon, *Confining light to deep subwavelength dimensions to enable optical nanopatterning*. *Science*, 2009. **324**(5929): p. 917-921.
131. Scott, T.F., et al., *Two-Color Single-Photon Photoinitiation and Photoinhibition for Subdiffraction Photolithography*. *Science*, 2009. **324**(5929): p. 913-917.
132. Groves, T.R., *Electron Optical Lithography*. *Nanolithography*, 2014.
133. Kampherbeek, B.J. and M. Wieland, *How to save over \$100 mln per year on lithography cost?* *Maskless Electron Beam Lithography*, 2009. **3**.
134. Pronko, P.P., et al., *Machining of sub-micron holes using a femtosecond laser at 800nm*. *Optics Communications*, 1995. **114**(1-2): p. 106-110.

135. Kasai, M., R, et al., *The interaction of femtosecond and nanosecond laser pulses with the surface of glass*. J. Non-crys. Solids, 2003. **319**: p. 129-35.
136. Tamaki, T., W. Watanabe, and K. Itoh, *Laser micro-welding of transparent materials by a localised heat accumulation effect using a femtosecond fiber laser at 1558 nm* Opt. Express, 2006. **14**: p. 10460-10468.
137. Joglekar, A.P., et al., *A study of the deterministic character of optical damage by femtosecond laser pulses and applications to nanomachining*. Applied Physics B, 2003. **77**(1): p. 25-30.
138. Gattass, R.R. and E. Mazur, *Femtosecond laser micromachining in transparent materials*. Nature Photonics, 2008. **2**(4): p. 219-225.
139. Jiang, L., et al., *Electrons dynamics control by shaping femtosecond laser pulses in micro/nanofabrication: modeling, method, measurement and application*. Light: Science & Applications, 2018. **7**: p. 17134.
140. Fang, R., A. Vorobyev, and C. Guo, *Direct visualization of the complete evolution of femtosecond laser-induced surface structural dynamics of metals*. Light: Science & Applications, 2017. **6**: p. 16256.
141. Stuart, B.C., et al., *Nanosecond-to-femtosecond laser-induced breakdown in dielectrics*. Phys. Rev. B, 1996. **53**: p. 1749-61.
142. Schaffer, C.B., A. Brodeur, and E. Mazure, *Laer-induced breakdown and damage in bulk transparent materials using tightly-focused femtosecond laser pulses*. Meas. Sci. Technol., 2001. **12**: p. 1784.
143. Sugioka, K. and Y. Cheng, *Ultrafast lasers—reliable tools for advanced materials processing*. Light Sci. Appl., 2014. **3**: p. 149.
144. Konig, J., S. Nolte, and A. Tunnermann, *Plasma evolution during metal ablation with ultrashort laser pulses*. Opt. Express, 2005. **13**: p. 10597-10607.
145. Dumitru, G., et al., *Femtosecond ablation of ultrahard materials*. Appl. Phys. A, 2002. **74**: p. 729-739.
146. Wu, D., et al., *In-channel integration of designable microoptical devices using flat scaffold-supported femtosecond-laser microfabrication for coupling-free optofluidic cell counting*. Light: Science & Applications, 2015. **4**: p. 228.
147. Chen, Y., et al., *Laser writing of coherent colour centres in diamond*. Nature Photonics, 2017. **11**: p. 77-80.
148. Choi, I., et al., *Laser-induced phase separation of silicon carbide*. Nature Communications, 2016. **7**: p. 13562.
149. Li, D., et al., *In situ imaging and control of layer-by-layer femtosecond laser thinning of graphene*. Nanoscale, 2015. **7**: p. 3651-3659.
150. He, H., et al., *Manipulation of cellular light from green fluorescent protein by a femtosecond laser*. Nature Photonics, 2012. **6**: p. 651-656.
151. Heath, D.J., et al., *Single-pulse ablation of multi-depth structures via spatially filtered binary intensity masks*. Applied Optics, 2018. **57**(8): p. 1904-1909.
152. Klein-Wiele, J. and P. Simon, *Fabrication of periodic nanostructures by phase-controlled multiple-beam interference*. Applied Physics Letters, 2003. **83**(23): p. 4707-4709.
153. Nakata, Y., et al., *Fabricating a regular hexagonal lattice structure by interference pattern of six femtosecond laser beams*. Applied Surface Science, 2017. **417**: p. 69-72.
154. Nakata, Y., et al., *Design of interference using coherent beams configured as a six-sided pyramid*. Applied Optics, 2012. **51**(21): p. 5004-5010.
155. Yang, L., et al., *The Fabrication of Micro/Nano Structures by Laser Machining*. Nanomaterials, 2019. **9**(12): p. 1789.
156. Messaoudi, H., et al., *Distortion-free laser beam shaping for material processing using a digital micromirror device*. Production Engineering, 2017. **11**(3): p. 365-371.
157. Hulsart-Billstrom, G., et al., *A surprisingly poor correlation between in vitro and in vivo testing of biomaterials for bone regeneration: results of a multicentre analysis*. European Cells and Materials, 2016. **31**: p. 312-322.
158. Karslioglu, Y., N. Yigit, and O. Onguru, *Chalkley method in the angiogenesis research and its automation via computer simulation*. Pathol Res Pract, 2014. **210**(3): p. 161-8.
159. Mangir, N., et al., *Using ex Ovo Chick Chorioallantoic Membrane (CAM) Assay To Evaluate the Biocompatibility and Angiogenic Response to Biomaterials*. ACS Biomaterials Science & Engineering, 2019. **5**(7): p. 3190-3200.
160. Lewis, R.M. and J.E. Pearson-Farr, *Multiscale three-dimensional imaging of the placenta*. Placenta, 2020. **102**: p. 55-60.
161. Burton, G.J., *Scanning electron microscopy of intervillous connections in the mature human placenta*. J Anat, 1986. **147**: p. 245-54.
162. Mayhew, T.M., *Morphomics: An integral part of systems biology of the human placenta*. Placenta, 2015. **36**(4): p. 329-40.
163. Cahill, L.S., et al., *Feto- and utero-placental vascular adaptations to chronic maternal hypoxia in the mouse*. J Physiol, 2018. **596**(15): p. 3285-3297.
164. Wang, Y. and S. Zhao, *Cell Types of the Placenta*, in *Vascular Biology of the Placenta*. 2010, Morgan & Claypool Life Sciences: San Rafael (CA).
165. Alberts, B., et al., *Molecular Biology of the Cell*. 4 ed. 2002, New York: Garland Science.
166. Barreto, R.S.N., et al., *Pericytes in the Placenta: Role in Placental Development and Homeostasis*. Adv Exp Med Biol, 2019. **1122**: p. 125-151.
167. Bergers, G. and S. Song, *The role of pericytes in blood-vessel formation and maintenance*. Neuro Oncol, 2005. **7**(4): p. 452-64.
168. Palaiologou, E., et al., *Human placental villi contain stromal macrovesicles associated with networks of stellate cells*. J Anat, 2020. **236**(1): p. 132-141.
169. Kherlopian, A.R., et al., *A review of imaging techniques for systems biology*. BMC Syst Biol, 2008. **2**: p. 74.
170. Denk, W. and H. Horstmann, *Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure*. PLoS Biol, 2004. **2**(11): p. e329.

References

171. Kazemian, A., et al., *Evolutionary implications of fetal and maternal microvillous surfaces in epitheliochorial placentae*. J Morphol, 2019. **280**(4): p. 615-622.
172. Palaiologou, E., et al., *Serial block-face scanning electron microscopy reveals novel intercellular connections in human term placental microvasculature*. J Anat, 2020. **237**(2): p. 241-249.
173. Zachow, S., M. Zilske, and H. Hege. *3-D reconstruction of individual anatomy from medical image data: Segmentation and geom-etry processing*. in *Konrad-Zuse-Zentrum für Informationstechnik Berlin*. 2007. Berlin.
174. Pugin, E. and A. Zhiznyakov, *Histogram method of image binarization based on fuzzy pixel representation*, in *DYNAMICS OF SYSTEMS, MECHANISMS AND MACHINES 2017 (Dynamics)*, IEEE, Editor. 2017: Omsk.
175. Deerinck, T.J., et al., *Enhancing Serial Block-Face Scanning Electron Microscopy to Enable High Resolution 3-D Nanohistology of Cells and Tissues*. Microsc. Microanal., 2010. **16**(2): p. 1138-1139.
176. Simard, P.Y., D. Steinkraus, and J.C. Platt, *Best practices for convolutional neural networks applied to visual document analysis*. Seventh International Conference on Document Analysis and Recognition, Vols I and II, Proceedings, 2003: p. 958-962.
177. Wang, K., et al., *An automatic learning rate decay strategy for stochastic gradient descent optimization methods in neural networks*. International Journal of Intelligent Systems, 2022.
178. Zhang, Q.L. and C.S. Shu, *Performance Investigation of Learning Rate Decay in LMS-Based Equalization*. Ieee Photonics Technology Letters, 2021. **33**(2): p. 109-112.
179. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning*. Nature, 2015. **521**(7553): p. 436-44.
180. Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning*. 2016: MIT Press.
181. Hornik, K., M. Stinchcombe, and H. White, *Multilayer feedforward networks are universal approximators*. Neural Network, 1989. **2**(5): p. 359-366.
182. Mills, B., et al., *Predictive capabilities for laser machining via a neural network*. Opt Express, 2018. **26**(13): p. 17245-17253.
183. Rumerlhart, D.E., G.E. Hinton, and R.J. Williams, *Learning representations by back-propagating errors*. Nature, 1986. **323**: p. 533-536.
184. Shamout, F., T. Zhu, and D.A. Clifton, *Machine Learning for Clinical Outcome Prediction*. IEEE Rev Biomed Eng, 2021. **14**: p. 116-126.
185. Lo, S.C.B., et al., *Artificial convolution neural network for medical image pattern recognition*. Neural Networks, 1995. **8**(7-8): p. 1201-1214.
186. LeCun, Y. and Y. Bengio, *Convolutional Netowkrs for Images, Speech and Time-Series. The handbook of brain theory and neural networks*. 1995: MIT Press.
187. LeCun, Y., et al. *Handwritten digit recognition with a back-propagation network*. in *Advances in Neural Information Processing Systems 2 (NIPS 1989)*. 1989.
188. Vaillant, R., C. Monrocq, and Y. Lecun, *Original Approach for the Localization of Objects in Images*. Iee Proceedings-Vision Image and Signal Processing, 1994. **141**(4): p. 245-250.
189. Zhang, X. and Y. LeCun, *Text Understanding from Scratch*. arXiv, 2015: p. arXiv:1502.01710v5.
190. Donahue, J., et al., *Long-term Recurrent Convolutional Networks for Visual Recognition and Description*, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Editor. 2015. p. 2625-2634.
191. Dong, C., et al., *Image Super-Resolution Using Deep Convolutional Networks*. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2016. **38**(2): p. 295-307.
192. Nowlan, S. and J.C. Platt, *A convolutional neural network hand tracker*. Advances in Neural Information Processing Systems, 1995: p. 901-908.
193. Simoyan, K. and A. Zisserman, *Very deep convolutional networks for large-scale image recognition*. arXiv, 2015: p. arXiv:1409.155v6.
194. De Fauw, J., et al., *Clinically applicable deep learning for diagnosis and referral in retinal disease*. Nat Med, 2018. **24**(9): p. 1342-1350.
195. Hosseinzadeh Kassani, S. and P. Hosseinzadeh Kassani, *A comparative study of deep learning architectures on melanoma detection*. Tissue Cell, 2019. **58**: p. 76-83.
196. Xu, J., et al., *A Deep Convolutional Neural Network for segmenting and classifying epithelial and stromal regions in histopathological images*. Neurocomputing, 2016. **191**: p. 214-223.
197. Liang, L., M. Liu, and W. Sun, *A deep learning approach to estimate chemically-treated collagenous tissue nonlinear anisotropic stress-strain responses from microscopy images*. Acta Biomater, 2017. **63**: p. 227-235.
198. Esteva, A., et al., *A guide to deep learning in healthcare*. Nat Med, 2019. **25**(1): p. 24-29.
199. Dzobo, K., et al., *Integrating Artificial and Human Intelligence: A Partnership for Responsible Innovation in Biomedical Engineering and Medicine*. OMICS, 2020. **24**(5): p. 247-263.
200. Ching, T., et al., *Opportunities and obstacles for deep learning in biology and medicine*. J R Soc Interface, 2018. **15**(141).
201. Fukushima, K., S. Miyake, and T. Ito, *Neocognitron - a Neural Network Model for a Mechanism of Visual-Pattern Recognition*. Ieee Transactions on Systems Man and Cybernetics, 1983. **13**(5): p. 826-834.
202. Felleman, D.J. and D.C. Van Essen, *Distributed Hierarchical Processing in the Primate Cerebral Cortex*. Cerebral Cortex, 1991. **1**(1): p. 1-47.
203. Duffy, K.R. and D.H. Hubel, *Receptive field properties of neurons in the primary visual cortex under photopic and scotopic lighting conditions*. Vision Research, 2007. **47**(19): p. 2569-2574.
204. Hubel, D.H. and T.N. Wiesel, *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex*. J Physiol, 1962. **160**: p. 106-54.
205. Hubel, D.H. and T.N. Wiesel, *Republication of The Journal of Physiology (1959) 148, 574-591: Receptive fields of single neurones in the cat's striate cortex*. Journal of Physiology-London, 2009. **587**(12): p. 2722-2732.

206. Murphy, T., *The First Level of Super Mario Bros. is Easy with Lexicographic Orderings and Time Travel* 2013.
207. Farzad, A., H. Mashayekhi, and H. Hassanpour, *A comparative performance analysis of different activation functions in LSTM networks for classification*. *Neural Computing and Applications*, 2019. **31**(7): p. 2507-2521.
208. Wang, Y., et al., *The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition*. *Applied Sciences*, 2020. **10**(5): p. 1897.
209. Wang, S.H., et al., *Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling*. *J Med Syst*, 2018. **42**(5): p. 85.
210. Xu, B., et al., *Empirical Evaluation of Rectified Activations in Convolutional Network*. arXiv, 2015: p. 1505.00853.
211. Dong, C., et al., *Learning a Deep Convolutional Network for Image Super-Resolution*. *Computer Vision - Eccv 2014, Pt Iv*, 2014. **8692**: p. 184-199.
212. Xu, Z., et al., *Joint solution for PET image segmentation, denoising, and partial volume correction*. *Med Image Anal*, 2018. **46**: p. 229-243.
213. Zhao, H., et al., *Loss functions for image restoration with neural networks*. arXiv, 2018: p. 1511.08861.
214. Falk, T., et al., *Author Correction: U-Net: deep learning for cell counting, detection, and morphometry*. *Nat Methods*, 2019. **16**(4): p. 351.
215. Falk, T., et al., *U-Net: deep learning for cell counting, detection, and morphometry*. *Nat Methods*, 2019. **16**(1): p. 67-70.
216. Kohl, S.A.A., et al., *A Probabilistic U-Net for Segmentation of Ambiguous Images*. *Advances in Neural Information Processing Systems 31 (Nips 2018)*, 2018. **31**.
217. Ronneberger, O., P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*. *Medical Image Computing and Computer-Assisted Intervention, Pt Iii*, 2015. **9351**: p. 234-241.
218. Long, J., E. Shelhamer, and T. Darrell, *Fully Convolutional Networks for Semantic Segmentation*. 2015 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr), 2015: p. 3431-3440.
219. Marquez, E.S., J.S. Hare, and M. Niranjani, *Deep Cascade Learning*. *Ieee Transactions on Neural Networks and Learning Systems*, 2018. **29**(11): p. 5475-5485.
220. Mackay, B.S., et al., *Automated 3D Labelling of Fibroblasts and Endothelial Cells in SEM-Imaged Placenta using Deep Learning*. *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies, Vol 2: Bioimaging*, 2020: p. 46-53.
221. Mills, B., et al., *Sub-micron-scale femtosecond laser ablation using a digital micromirror device*. *Journal of Micromechanics and Microengineering*, 2013. **23**: p. 035005.
222. Heath, D.J., et al., *Closed-loop corrective beam shaping for laser processing of curved surfaces*. *Journal of Micromechanics and Microengineering*, 2018. **28**: p. 127001-127008.
223. Gentile, F., *Cell aggregation on nanorough surfaces*. *J Biomech*, 2021. **115**: p. 110134.
224. Howard, D., et al., *Immunoselection and adenoviral genetic modulation of human osteoprogenitors: in vivo bone formation on PLA scaffold*. *Biochem Biophys Res Commun*, 2002. **299**(2): p. 208-15.
225. Kanczler, J., et al., *Isolation, Differentiation, and Characterization of Human Bone Marrow Stem Cells In Vitro and In Vivo*. *Methods Mol Biol*, 2019. **1914**: p. 53-70.
226. Elsharkawy, S. and A. Mata, *Hierarchical Biomineralization: from Nature's Designs to Synthetic Materials for Regenerative Medicine and Dentistry*. *Adv Healthc Mater*, 2018. **7**(18): p. e1800178.
227. Cheng, Z.A., et al., *Nanoscale Coatings for Ultralow Dose BMP-2-Driven Regeneration of Critical-Sized Bone Defects*. *Adv Sci (Weinh)*, 2019. **6**(2): p. 1800361.
228. Mackay, B.S., et al., *Deep Learning for the Automated Feature Labelling of 3-Dimensional Imaged Placenta*, in *Biomedical Engineering Systems and Technologies*, X. Ye, et al., Editors. 2021, Springer International Publishing: Switzerland. p. 93-115.
229. Moen, E., et al., *Deep learning for cellular image analysis*. *Nat Methods*, 2019. **16**(12): p. 1233-1246.
230. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis*. *Nat Methods*, 2012. **9**(7): p. 671-5.
231. Grant-Jacob, J.A., et al., *A neural lens for super-resolution biological imaging*. *Journal of Physics Communications*, 2019. **3**(6).
232. Keeling, M.J., et al., *Predictions of COVID-19 dynamics in the UK: Short-term forecasting and analysis of potential exit strategies*. *PLoS Comput Biol*, 2021. **17**(1): p. e1008619.
233. AviSim. *Flight simulator capacity, how well are simulators utilized?* News 2018 12/12/2018 [cited 2021 18/06/2021]; Available from: <https://avi-sim.com/flight-simulator-capacity/>.
234. Rabel, K., et al., *Controlling osteoblast morphology and proliferation via surface micro-topographies of implant biomaterials*. *Sci Rep*, 2020. **10**(1): p. 12810.
235. Karras, T., et al., *Progressive Growing of GANs for Improved Quality, Stability, and Variation*. arXiv, 2017: p. 1710.101.96.
236. Karras, T., et al., *Training Generative Adversarial Networks with Limited Data*. arXiv, 2020: p. arXiv:2006.06676v2.
237. Miyato, T., et al., *Spectral Normalization for Generative Adversarial Networks*. arXiv, 2018: p. 1802.05957.
238. Miyato, T. and M. Koyama, *cGANs with Projection Discriminator*. arXiv, 2018: p. 1802.05637.
239. Noguchi, A. and T. Harada, *Image Generation From Small Datasets via Batch Statistics Adaption*, in *ICCV, C.V. Foundation*, Editor. 2019, IEEE: Seoul.
240. Radford, A., L. Metz, and S. Chintala, *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*. arXiv, 2016: p. 1511.06434.
241. Shorten, C. and T.M. Khoshgoftaar, *A survey on Image Data Augmentation for Deep Learning*. *Journal of Big Data*, 2019. **6**(1): p. 60.

References

242. Clark, P., et al., *Topographical control of cell behaviour. I. Simple step cues*. Development, 1987. **99**(3): p. 439-48.
243. Clark, P., et al., *Topographical control of cell behaviour: II. Multiple grooved substrata*. Development, 1990. **108**(4): p. 635-44.
244. Dunn, G.A. and A.F. Brown, *Alignment of fibroblasts on grooved surfaces described by a simple geometric transformation*. J Cell Sci, 1986. **83**: p. 313-40.
245. Oakley, C. and D.M. Brunette, *The sequence of alignment of microtubules, focal contacts and actin filaments in fibroblasts spreading on smooth and grooved titanium substrata*. J Cell Sci, 1993. **106** (Pt 1): p. 343-54.
246. Gill, E.L., et al., *Multi-length scale bioprinting towards simulating microenvironmental cues*. Biodes Manuf, 2018. **1**(2): p. 77-88.
247. Krishnamoorthy, S., Z. Zhang, and C. Xu, *Guided cell migration on a graded micropillar substrate*. Bio-Design and Manufacturing, 2020. **3**(1): p. 60-70.
248. Cutiongco, M.F.A., et al., *Predicting gene expression using morphological cell responses to nanotopography*. Nat Commun, 2020. **11**(1): p. 1384.
249. Rice, J.A., *Mathematical statistics and data analysis*. 3rd ed. 2007, Doncaster, United Kingdom: Duxbury Press.
250. Mingu, S., et al., *The behaviour of Neuro-2A cells on silicon substrates with various topographies generated by femtosecond laser micromachining*, in *BIOSTEC: BIOIMAGING*. 2020, IEEE: Valetta, Malta.
251. Ratner, B.D., *Biomaterials: Been There, Done That, and Evolving into the Future*. Annu Rev Biomed Eng, 2019. **21**: p. 171-191.
252. Festing, S. and R. Wilkinson, *The ethics of animal research. Talking Point on the use of animals in scientific research*. EMBO Rep, 2007. **8**(6): p. 526-30.
253. Drinkwater, E., E.J.H. Robinson, and A.G. Hart, *Keeping invertebrate research ethical in a landscape of shifting public opinion*. Methods in Ecology and Evolution, 2019. **10**(8): p. 1265-1273.
254. Lemon, G., et al., *Mathematical modelling of tissue-engineered angiogenesis*. Math Biosci, 2009. **221**(2): p. 101-20.
255. Barber, G., *Shark or Baseball? Inside the 'Black Box' of a Neural Network*, in *Wired*. 2019, Condé Nast Publications.
256. Carter, S., et al., *Exploring Neural Networks with Activation Atlases*. Distill, 2019.
257. Smith, T.G., G.D. Lange, and W.B. Marks, *Fractal methods and results in cellular morphology — dimensions, lacunarity and multifractals*. Journal of Neuroscience Methods, 1996. **69**(2): p. 123-136.
258. Zhou, W., et al., *Epileptic Seizure Detection Using Lacunarity and Bayesian Linear Discriminant Analysis in Intracranial EEG*. IEEE Trans Biomed Eng, 2013. **60**(12): p. 3375-81.
259. Zaia, A., et al., *MR imaging and osteoporosis: fractal lacunarity analysis of trabecular bone*. IEEE Trans Inf Technol Biomed, 2006. **10**(3): p. 484-9.
260. Yasar, F. and F. Akgunlu, *Fractal dimension and lacunarity analysis of dental radiographs*. Dentomaxillofac Radiol, 2005. **34**(5): p. 261-7.
261. Soares, F., et al., *3D lacunarity in multifractal analysis of breast tumor lesions in dynamic contrast-enhanced magnetic resonance imaging*. IEEE Trans Image Process, 2013. **22**(11): p. 4422-35.
262. Smitha, K.A., A.K. Gupta, and R.S. Jayasree, *Fractal analysis: fractal dimension and lacunarity from MR images for differentiating the grades of glioma*. Phys Med Biol, 2015. **60**(17): p. 6937-47.
263. Popovic, N., et al., *Fractal dimension and lacunarity analysis of retinal microvascular morphology in hypertension and diabetes*. Microvasc Res, 2018. **118**: p. 36-43.
264. Plotnick, R.E., et al., *Lacunarity analysis: A general technique for the analysis of spatial patterns*. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics, 1996. **53**(5): p. 5461-5468.
265. Palanivel, D.A., et al., *Multifractal-based lacunarity analysis of trabecular bone in radiography*. Comput Biol Med, 2020. **116**: p. 103559.
266. Karperien, A.L. and H.F. Jelinek, *Fractal, multifractal, and lacunarity analysis of microglia in tissue engineering*. Front Bioeng Biotechnol, 2015. **3**: p. 51.
267. Gudea, A.I. and A.C. Stefan, *Histomorphometric, fractal and lacunarity comparative analysis of sheep (Ovis aries), goat (Capra hircus) and roe deer (Capreolus capreolus) compact bone samples*. Folia Morphol (Warsz), 2013. **72**(3): p. 239-48.
268. Gould, D.J., et al., *Multifractal and lacunarity analysis of microvascular morphology and remodeling*. Microcirculation, 2011. **18**(2): p. 136-51.
269. Gilmore, S., et al., *Lacunarity analysis: a promising method for the automated assessment of melanocytic naevi and melanoma*. PLoS One, 2009. **4**(10): p. e7449.
270. Gereben, O., *Lacunarity analysis of atomic configurations: application to ethanol-water mixtures*. Phys Rev E Stat Nonlin Soft Matter Phys, 2015. **92**(3): p. 033305.
271. Dougherty, G. and G.M. Henebry, *Lacunarity analysis of spatial pattern in CT images of vertebral trabecular bone for assessing osteoporosis*. Med Eng Phys, 2002. **24**(2): p. 129-38.
272. Diaz, S., et al., *Progression of emphysema in a 12-month hyperpolarized 3He-MRI study: lacunarity analysis provided a more sensitive measure than standard ADC analysis*. Acad Radiol, 2009. **16**(6): p. 700-7.
273. Cordeiro, M.S., et al., *Fibrous Dysplasia Characterization Using Lacunarity Analysis*. J Digit Imaging, 2016. **29**(1): p. 134-40.
274. Stangroom, J. *Social Science Statistics T-Test Calculator*. [cited 2021 03.06.2021]; Available from: <https://www.socscistatistics.com/>.
275. Zudaire, E., et al., *A computational tool for quantitative analysis of vascular networks*. PLoS One, 2011. **6**(11): p. e27385.
276. Zabell, S.L., *On Student's 1908 article - "The Probable Error of a Mean"*. Journal of the American Statistical Association, 2008. **103**(481): p. 1-7.

277. Dagostino, R.B., A. Belanger, and R.B. Dagostino, *A Suggestion for Using Powerful and Informative Tests of Normality*. *American Statistician*, 1990. **44**(4): p. 316-321.

Contacting The Author

CONTACTING THE AUTHOR

There may be inconsistencies or errors in this thesis. If you find any, please contact me at b.mackay@soton.ac.uk.

If you wish to discuss the research contained in this project, please contact Dr Ben Mills at bm602@orc.soton.ac.uk.