Classifying paint colour using acoustic data from laser ablation

JAMES A. GRANT-JACOB,[1,*] MICHALIS N. ZERVAS,[1,†] AND BEN MILLS[1,†]

[1]*Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK*
[†]*These authors contributed equally.*
**J.A.Grant-Jacob@soton.ac.uk*

Abstract: We present a novel approach for identifying the paint colour and tone on a surface by using the acoustic signal collected during laser ablation. We trained convolutional neural networks to classify the colour (from 8 different colours) and the tone (the percentage of black in the paint). The colour was predicted with ~91% accuracy and the tone with an R-value of 0.95. This technique has significant potential for supporting real-time optimisation in laser-material processing, particularly for high-precision laser cleaning, as well as broader applications where direct visual observation of the sample is not feasible.

1. Introduction

Lasers have become ubiquitous across manufacturing [1], with applications including metal welding [2], machining of metals and ceramics [3], crystalline material deposition [4], fabric patterning [5], and rust and surface material removal [6]. However, depending on the process and the laser used, plasma can form during ablation due to the high laser intensity, which ionizes the material or the surrounding air at the ablation site[7–11]. Due to the plasma produced during laser ablation, it is often impossible to directly observe the surface and verify whether the desired ablation has occurred. Developing the ability to monitor and understand the ablation process without visual feedback could transform the laser processing industry by enabling more precise machining control and improving overall efficiency.

Existing methods for detecting paint colour during laser ablation primarily rely on optical techniques such as laser-induced breakdown spectroscopy (LIBS). LIBS is widely used for elemental analysis by measuring the emitted light spectrum from the plasma generated during ablation, providing insights into material composition, including pigments responsible for colour differentiation [12] .

Additionally, hyperspectral imaging has been utilized to analyse paint composition by capturing both spatial and spectral information, thereby allowing differentiation between paint layers and subtle colour variations [13].

In addition, Raman spectroscopy has also been employed for pigment analysis, demonstrating its potential for non-destructive characterization of artistic materials. For instance, González-Vidal et al. [14] developed an automatic classification system for Raman spectra that enables rapid identification of pigment compositions, thereby facilitating efficient analysis in conservation science. Similarly, Nevin et al. [15] applied Raman spectroscopic methods to pigment analysis, highlighting the ability to distinguish subtle chemical differences among pigments.

However, these methods often require complex and expensive optical setups or can be time consuming, making real-time and cost-effective solutions challenging to implement. The integration of deep learning, particularly convolutional neural networks (CNNs), offers a promising alternative for automated and robust paint classification based on acoustic signals generated during ablation.

In recent years, deep learning, particularly convolutional neural networks (CNNs) [16], has gained significant attention in science due to advances in GPU computing power [17]. CNNs can classify features in data, such as images [18,19], with high accuracy and speed, often providing results in milliseconds. These networks have been applied to various tasks, including image classification of bird songs[20], labelling cars [21], and analysing magnetic resonance

imaging (MRI) scans [22]. In other applications of deep learning for material classification, 3D convolutional neural networks (3D-CNNs) have been applied to hyperspectral images for material classification [23]. The authors analyse how 3D-CNN models process spatial information and compare their performance in classifying hyperspectral images, which contain detailed spectral data about materials. The study provides insights into the role of deep learning in material classification beyond traditional imaging techniques.

The application of deep learning has also been applied in real-time classification of enamel paint coatings on a conveyor line [24]. The authors develop a system that classifies different object types and surface conditions of enamel paint using computer vision. The research highlights how CNN-based methods can be used for automated quality control in industrial settings. More specific to lasers, CNNs have been applied to welding [25], particulate matter detection [26], laser scanning [27], additive manufacturing [28], laser powder bed fusion [29], subtractive machining [18], and laser induced breakdown spectroscopy in laser cleaning [31].

Recent advancements have demonstrated the integration of acoustic signals and deep learning techniques in various laser-based material analysis applications. For instance, in laser paint removal, the acoustic emissions generated during the process provide valuable information about the state of paint removal. One study introduced a real-time acoustic monitoring method using deep learning to analyse these signals, enabling precise process control [32]. In laser beam welding, acoustic process monitoring combined with machine learning has been applied to analyse recorded acoustic signals to aid in defect detection and optimization of welding parameters, improving weld quality [33].

For additive manufacturing, particularly selective laser melting (SLM), a method combining acoustic signals with deep learning was proposed to monitor spatter behaviours. Acoustic data recorded by a microphone was analysed alongside high-speed camera footage, demonstrating that deep learning could effectively track spatter in a cost-efficient manner [34]. In laser ablation, both acoustic and plasma emissions have been analysed using deep learning models to classify tissue types, identify laser pulse energy, and predict the appearance of the ablated sample, improving the understanding of laser-tissue interactions [35]. Similarly, Alperovich et al. [36] demonstrated real-time tissue classification by analysing acoustic emissions during pulsed ultraviolet laser ablation, applying machine learning for automated tissue identification.

Other recent results including predictive visualisation of the sample surface during ablation [37], tissue classification during laser ablation [38], and for laser weld depth classification [39]. This paper introduces a novel approach for classifying paint colour and tone, using acoustic data collected in real-time during laser ablation. More specifically, we use a microphone to collect acoustic signal during laser ablation of paint, and use deep learning to classify the colour (from 8 different colours) and the tone (the percentage of black in the paint) of the ablated paint. Traditional methods in material processing often rely on visual inspection, which can be limited by accessibility, surface contamination, or material transparency. In contrast, this method uses CNNs to analyse acoustic signals, offering a non-visual, indirect measurement technique that operates effectively even when optical viewing of sample is restricted.

2. Experiment and method

2.1 Sample fabrication

Two microscope slides (25 mm × 75 mm × 1 mm thick soda-lime glass) from Thermo Fisher Scientific, each featuring ten wells, were used as the substrates for the paint. Daler Rowney Georgian oil paints were used as the samples in this study and were diluted in plastic wells with 60 ml of Winsor & Newton artist white spirit. The paints, referred to by simplified names for clarity throughout the work were Lamp Black (black), French Ultramarine (blue), Sap Green (green), Lime Yellow (lime), Yellow Ochre (ochre), Cadmium Red (orange), Crimson Alizarin (red), and Titanium White (white).

The diluted paints were deposited onto each substrate using a 20 mL Eppendorf Research plus micropipette at a volume of 10 mL in each well. The first slide consisted of the eight colours. The second slide featured varying black/white ratios, starting with a 0/100 ratio (0% black) in the first well, which was entirely white, followed by a 10/90 ratio (10% black) in the next well, producing a weakly grey tone. This was repeated for all ten wells, with the final well having a 90/10 ratio (90% black). The first slide, which had pure black, was used to represent the 100/0 ratio (100% black). Each tone was created by depositing a total of 10 mL of paint mixture per well. For instance, 3 mL of black mixed with 7 mL of white resulted in a 30/70 ratio (30% black). The paints were left to dry over several days before they were used in the laser ablation setup.

2.2 Laser ablation setup

Single laser pulses (1 mJ, 1030 nm wavelength and ~190 fs pulse length) from a Light Conversion Pharos SP, were focused onto the surface of a painted substrate using a 20× magnification Nikon microscope objective lens (TU Plan ELWD, 0.40 NA) with a spot size of 30 µm (see figure 1 for diagram of setup). The microscope objective also enabled the surface ablation to be imaged and monitored with the use of a camera (Basler a2A5320-23ucPRO, 5320 × 5320 pixels). By including the camera, the sample position could be adjusted by the motorised stages so that each paint colour could be located, and each paint colour's top surface could be aligned to be in focus prior to ablation, for consistency throughout all colours. The glass slide samples were clamped onto a motorised linear 3-axis stage (Zaber, 3 × LSM050A-T4) that was computer-controlled and allowed accurate positioning of the sample (in 1 µm step size movement). A USB microphone (Adafruit, 22.2 mm × 18.3 mm × 7.0 mm) was placed approximately 1 mm below the lower surface of the glass slide and 2 mm away from the laser axis. The microphone was connected to a Windows 10 workstation with an Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz 3.60 GHz, 64.0 GB RAM, and an NVIDIA Titan Xp (12 GB VRAM).

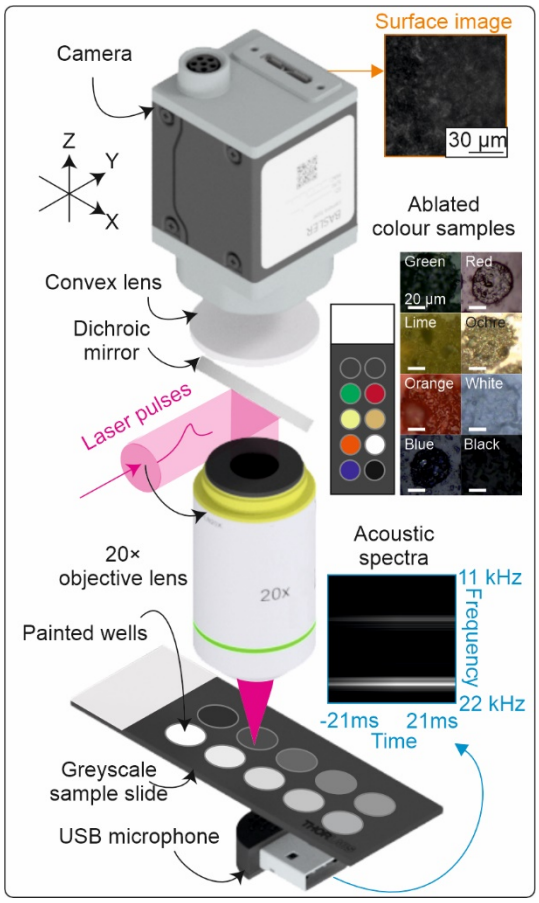

Fig. 1. Femtosecond laser pulses were focussed onto the surface of a dried painted region and the acoustic signal produced as a result of ablation was recorded using a USB microphone. The acoustic signals were then transformed into spectral images that were then used as the input to a neural network that classified the colour or the tone. The white bar indicates 20 µm in the inset ablated colour samples.

2.3 Data collection

Python code was used on the Windows workstation to automate the data collection process, controlling the 3-axis stage, laser pulses, microphone, and camera. Each sample category was positioned under the laser and single pulse ablation occurred in a 5 × 9 grid (total of 45 pulses), with 50 µm spacing between ablated spots. During each ablation, the acoustic signal was recorded using the USB microphone over a 3-second window. The temporal peak in the signal was identified, and the acoustic data was cropped to a 42.3 ms window, with the peak centered along the time axis (horizontal in the images in this work). The temporal peak in the signal refers to the maximum amplitude observed in the time domain of the acoustic waveform. This peak was identified to align the acoustic signal with the moment of laser-material interaction, ensuring that the extracted features were directly related to the ablation event. A sample period of 42.3 ms was chosen based on the duration of the observable acoustic signal, allowing for optimal feature extraction.

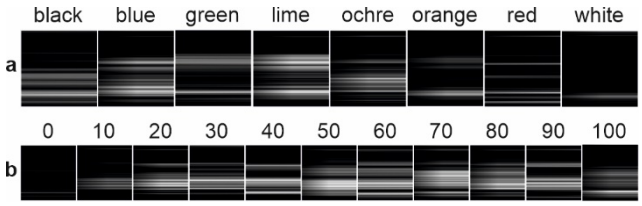

Fig. 2. Diagram showing examples of acoustic spectra for the different categories, for a) colour and b) tone (where 0 corresponds to white and 100 to black).

A Fourier transform was used to convert the time-domain acoustic data into its frequency spectrum. The lower frequencies of the spectra were removed due to low signal within this range, apart from signal in the range 100-400 Hz, which was present within all spectra. As a result, the resultant spectra were split into 94 bands from 11.4 kHz to 22.1 kHz (since this was where most of the signal existed), and then resized to a square array of 256 × 256 pixels and duplicated across three channels to create a 256 × 256 RGB image. The images were generated by mapping the frequency intensities onto a 256-pixel y-axis, with each pixel representing approximately 43 Hz (calculated as the 11 kHz frequency range divided by 256 pixels). The intensities were normalized using min-max scaling to ensure consistency across samples. Frequency bands were evenly distributed along the y-axis, providing a clear spectral representation of the acoustic data over time.

Figure 2 shows examples of acoustic spectra for the categories. The data were normalised to the average maximum signal in the dataset from 0–255-pixel intensity (1 channel, greyscale), to allow for signal to cover the bit depth of the images, and thus increase the dynamic range of the signals, to allow for more accurate CNN network to be trained.

2.3 Neural networks

Two neural networks were trained using MATLAB 2024b (see Fig. 3 for concept). The first network was for classifying the colour, and was a CNN (Inception v3 [40]) with 314 layers and a classification output layer. The second network was designed to classify numbers from 0 to 100, where 0 represented a sample with 0% black paint and 100 represented a sample with 100% black paint. This CNN architecture consisted of 18 layers, including 4 convolutional layers (each followed by batch normalization and ReLU activation), 2 average pooling layers, a dropout layer, a fully connected layer, and a regression output layer. The neural network was trained using an initial learning rate of 0.0002, with a validation frequency of 200 iterations, a learning rate drop-out factor of 0.1, and the mini-batch size was set to 2. In addition, weight decay (L2 regularization) was incorporated to penalize overly large weights and reduce overfitting. No adjustments were made the structure of the model, apart from the image input size, which changed from the default size of 299 × 299 × 3 to 256 × 256 × 3, where 3 indicates RGB channels.

Training took 22 minutes (20 epochs) for the first network and 2 minutes (20 epochs) for the second network, with both networks showing a training loss plateau during this time. Both neural networks were trained and tested on a Windows 10 computer workstation with two NVIDIA RTX A6000 (48 GB VRAM each).

In total, 328 acoustic spectra images were used for training the colour neural network and 32 (4 from each category) were used for testing the neural network. The tone neural network was trained on 365 spectra images, and were tested on 40 spectra (test data 1) randomly chosen images not used in training, in the quantity (8, 6, 3, 4, 7, 3, 5, 3, 1) from the 0, 10, 20, 40, 50, 60, 80, 90, 100 categories, respectively, that were used in training, and 45 spectra for each 30 and 70 category (test data 2) not used in training.

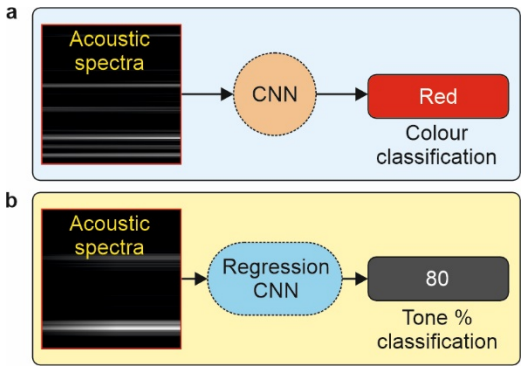

Fig. 3. The concept of using neural networks to identify (a) colour and (b) tone, from the acoustic spectra.

3. Results and discussion

Fig. 4(a) shows the confusion matrix for the results of testing the neural network on spectra from the 8 different coloured paints. The neural network was able to correctly predict 29 of the 32 randomly chosen test samples. Fig. 4(b) shows the application of Grad-CAM (Gradient-weighted Class Activation Mapping [41]), which is a technique to visualise and therefore interpret the decisions made by CNNs. Grad-CAM was applied with the reduction layer being the penultimate layer (SoftMax layer) before the classification output, for 1 of each of the different classes and the results are visualized as an overlay (50% alpha) of corresponding spectra. The figure presents a Grad-CAM visualisation applied to different coloured paint samples, highlighting the spectrum's most influential area in the CNN's classification process. Each sub-image corresponds to a specific paint colour, with an overlaid heat map indicating areas of high and low activation. The intensity scale at the bottom of each sub-image represents the strength of the activation, where brighter regions (e.g., yellow) signify higher importance in the model's decision-making, while darker regions (e.g., blue) indicate less influence. This visualisation aids in interpreting the CNN's spectral region of focus, aiding in understanding of how the neural network distinguishes between different paint colours.

As shown in the images, the neural network associates distinct spectral regions with different classifications. For example, in the white spectrum, the upper region has the greatest influence on the classification, while in the blue spectrum, the middle region is most influential, and in the green spectrum, the lower region has the strongest impact. This indicates that the neural network has identified spectral features (or a lack of spectral features) that are correlated with each colour.

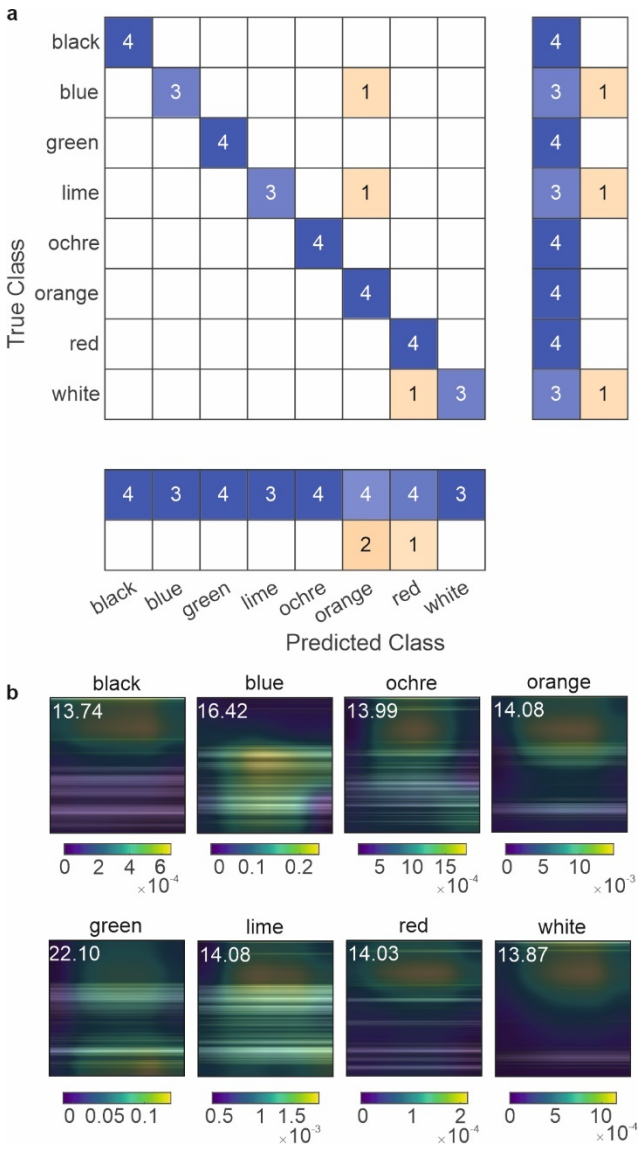

Fig. 4 (a) Confusion matrix showing the true class vs predicted class for the colour neural network being applied to acoustic spectra from the coloured samples. The blue shows the correct predictions and the orange the incorrect predictions, with the number of images tested on labeled in the corresponding squares. (b) Grad-CAM overlay of spectra from different samples, with the peak frequency (kHz) labeled in the inset in white.

Since a regression CNN was applied to train and test acoustic spectra from paint samples with varying tone, each prediction was a numerical value. Figure 5 shows the plot of actual values versus predicted values for the test data. The results of the test data from categories used in the training are shown as red circles in the figure, giving an R-value of 0.9915 and RMSE (root mean-squared error) of 4.9307. The results of the data from categories not used in training (30% and 70%) are shown as blue circles, giving an R-value of 0.9476, and a mean and RMSE of 18.1 ± 15.3 for 30% black and 71.1 ± 8.33 for 70% black. The ability for the neural network to identify the tone from the acoustic signal can be understood by looking at the example spectra in Fig. 2(b), which show visually distinct differences in the spectra. For example, 0% and 10% black have few visible frequency (horizontal lines of high intensity) components compared to

higher percentage such as 70-100% black. Future work could involve exploring mixtures of colours as well as other types of paints or materials.

This technique has significant potential for enhancing optical material processing, especially where visual assessment is difficult or impossible. For example, in high-precision laser cleaning, real-time acoustic-based classification can dynamically adjust laser parameters, improving selectivity and reducing substrate damage. In heritage conservation, it offers non-invasive insights without visual access. In industrial coating processes, it enables rapid, in-situ quality control by identifying paint colour, tone, rust or material in real-time. Additionally, this method supports adaptive laser processing, where feedback from acoustic signals allows for intelligent control systems, therefore broadening the utility of laser technologies in various environments, including industrial, manufacturing and cleanroom environments.

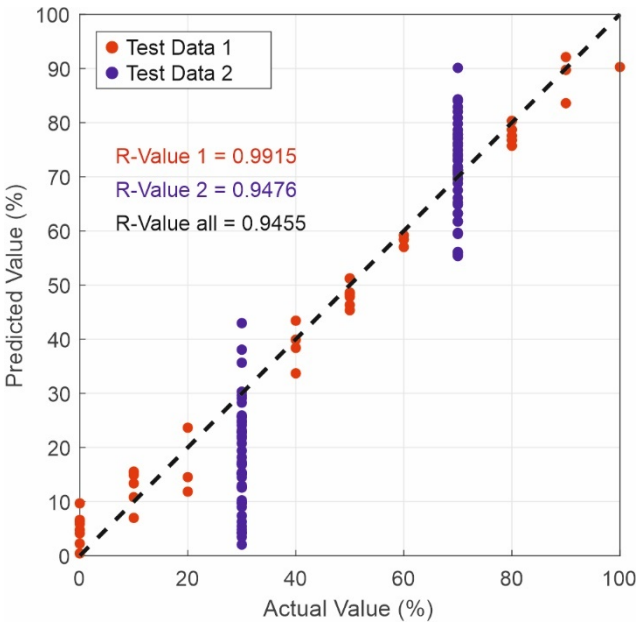

Fig. 5 Testing of the tone neural network on data not used in training. Red circles represent test data 1, covering the full range from 0 to 100, while blue circles represent test data 2, comprising the entire categories of 30 and 70 that were excluded from training.

4. Conclusion

In conclusion, this study demonstrates the ability of neural networks to classify both the color and the tone of paint from their acoustic spectra during laser ablation, achieving approximately 91% accuracy in color classification and an R-value of 0.95 for tone prediction. This technique holds potential for supporting real-time optimisation in laser-material processing, especially for high-precision laser cleaning, and it could be applied more broadly in situations where direct visual observation of the sample is not possible.

Funding

Engineering & Physical Sciences Research Council (EP/W028786/1, EP/T026197/1, EP/Z002567/1).

Data availability

Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Competing interests

The authors declare no competing interests.

References

1. N. B. Dahotre and S. Harimkar, *Laser Fabrication and Machining of Materials* (Springer Science & Business Media, 2008).
2. J. Blackburn, "Laser welding of metals for aerospace and other applications," in *Welding and Joining of Aerospace Materials* (Elsevier, 2012), pp. 75–108.
3. A. K. Dubey and V. Yadava, "Laser beam machining—a review," Int J Mach Tools Manuf 48(6), 609–628 (2008).
4. J. A. Grant-Jacob, S. J. Beecher, J. J. Prentice, D. P. Shepherd, J. I. Mackenzie, and R. W. Eason, "Pulsed laser deposition of crystalline garnet waveguides at a growth rate of 20 μm per hour," Surf Coat Technol (2017).
5. R. Nayak and R. Padhye, "The use of laser in garment manufacturing: an overview," Fashion and Textiles 3(1), 5 (2016).
6. G. Daurelio, G. Chita, and M. Cinquepalmi, "Laser surface cleaning, de-rusting, de-painting and de-oxidizing," Applied Physics A 69, S543–S546 (1999).
7. G. K. Lewis and R. D. Dixon, "Plasma monitoring of laser beam welds," Laser 13, 2 (1984).
8. D. J. Hwang, H. Jeon, C. P. Grigoropoulos, J. Yoo, and R. E. Russo, "Femtosecond laser ablation induced plasma characteristics from submicron craters in thin metal film," Appl Phys Lett 91(25), 251118 (2007).
9. Y. Kawahito, N. Matsumoto, M. Mizutani, and S. Katayama, "Characterisation of plasma induced during high power fibre laser welding of stainless steel," Science and Technology of Welding and Joining 13(8), 744–748 (2008).
10. A. Bogaerts and Z. Chen, "Effect of laser parameters on laser ablation and laser-induced plasma formation: A numerical modeling investigation," Spectrochim Acta Part B At Spectrosc 60(9–10), 1280–1307 (2005).
11. S. Amoruso, R. Bruzzese, C. Pagano, and X. Wang, "Features of plasma plume evolution and material removal efficiency during femtosecond laser ablation of nickel in high vacuum," Applied Physics A 89, 1017–1024 (2007).
12. L. Pagnin, L. Brunnbauer, R. Wiesinger, A. Limbeck, and M. Schreiner, "Multivariate analysis and laser-induced breakdown spectroscopy (LIBS): a new approach for the spatially resolved classification of modern art materials," Anal Bioanal Chem 412, 3187–3198 (2020).
13. A. Cosentino, "Identification of pigments by multispectral imaging; a flowchart method," Herit Sci 2(1), 8 (2014).
14. J. González-Vidal, R. Pérez-Pueyo, and M. J. Soneira, "Automatic classification system of Raman spectra applied to pigments analysis," Journal of Raman Spectroscopy 47(12), 1408–1414 (2016).
15. A. Nevin, I. Osticioli, D. Anglos, A. Burnstock, S. Cather, and E. Castellucci, "Raman spectra of proteinaceous materials used in paintings: a multivariate analytical approach for classification and identification," Anal Chem 79(16), 6143–6151 (2007).
16. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature 521, 436 (2015).
17. A. HajiRassouliha, A. J. Taberner, M. P. Nash, and P. M. F. Nielsen, "Suitability of recent hardware accelerators (DSPs, FPGAs, and GPUs) for computer vision and image processing algorithms," Signal Process Image Commun 68, 101–119 (2018).
18. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12 (Curran Associates Inc., 2012), pp. 1097–1105.
19. K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision* (2015), pp. 1026–1034.
20. B. Fazeka, A. Schindler, T. Lidy, and A. Rauber, "A multi-modal deep neural network approach to bird-song identification," arXiv preprint arXiv:1811.04448 (2018).
21. E. Talpes, D. D. Sarma, G. Venkataramanan, P. Bannon, B. McGee, B. Floering, A. Jalote, C. Hsiong, S. Arora, A. Gorti, and G. S. Sachdev, "Compute Solution for Tesla's Full Self-Driving Computer," IEEE Micro 40(2), 25–35 (2020).
22. A. Chattopadhyay and M. Maitra, "MRI-based brain tumour image detection using CNN based deep learning method," Neuroscience informatics 2(4), 100060 (2022).
23. B. Kang and S. Kim, "Analysis of the influence of 3D-CNN on spatial random information in hyperspectral image classification," in *2021 21st International Conference on Control, Automation and Systems (ICCAS)* (2021), pp. 1123–1126.
24. T. Citlak and N. Pillay, "Real-Time Object Classification on an Enamel Paint Coating Conveyor Line Using Mask R-CNN," Automation 5(3), 213–229 (2024).
25. D. Ma, P. Jiang, L. Shu, and S. Geng, "Multi-sensing signals diagnosis and CNN-based detection of porosity defect during Al alloys laser welding," J Manuf Syst 62, 334–346 (2022).
26. J. A. Grant-Jacob and B. Mills, "Deep learning in airborne particulate matter sensing: a review," J Phys Commun 6(12), 122001 (2022).
27. B. Nagy and C. Benedek, "3D CNN-based semantic labeling approach for mobile laser scanning data," IEEE Sens J 19(21), 10034–10045 (2019).

28. Y. Li, H. Yan, and Y. Zhang, "A Deep Learning Method for Material Performance Recognition in Laser Additive Manufacturing," in *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)* (2019), 1, pp. 1735–1740.
29. L. Scime and J. Beuth, "A multi-scale convolutional neural network for autonomous anomaly detection and classification in a laser powder bed fusion additive manufacturing process," Addit Manuf 24, 273–286 (2018).
30. B. Mills and J. Grant-Jacob, "Lasers that learn: the interface of laser machining and machine learning," IET Optoelectronics 15(5), 207–224 (2021).
31. S. Choi and C. Park, "Convolution neural network with laser-induced breakdown spectroscopy as a monitoring tool for laser cleaning process," Sensors 23(1), 83 (2022).
32. D. Chu, H. Zhou, H. Yang, G. Deng, H. Zhao, and S. Zhou, "Real-time acoustic monitoring of laser paint removal based on deep learning," Opt Express 33(1), 1421–1436 (2025).
33. T. Weiss, J. Werner, C. Geiger, and M. F. Zaeh, "Acoustic process monitoring during the laser beam welding of stainless-steel foils using an adjustable ring mode laser beam source," J Laser Appl 36(4), 042030 (2024).
34. S. Luo, X. Ma, J. Xu, M. Li, and L. Cao, "Deep learning based monitoring of spatter behavior by the acoustic signal in selective laser melting," Sensors 21(21), 7179 (2021).
35. C. Seppi, A. Huck, H. N. Kenhagho, E. Schnider, G. Rauter, A. Zam, and P. C. Cattin, "Deep-Learning Approach for Tissue Classification Using Acoustic Waves During Ablation With an Er:YAG Laser," IEEE Access 9, 130543–130553 (2021).
36. Z. Alperovich, G. Yamin, E. Elul, G. Bialolenker, and A. A. Ishaaya, "In situ tissue classification during laser ablation using acoustic signals," J Biophotonics 12(9), e201800405 (2019).
37. J. Grant-Jacob, B. Mills, and M. N. Zervas, "Acoustic and plasma sensing of laser ablation via deep learning," Opt Express 31(17), 28413–28422 (2023).
38. C. Seppi, A. Huck, H. N. Kenhagho, E. Schnider, G. Rauter, A. Zam, and P. C. Cattin, "Deep-Learning Approach for Tissue Classification Using Acoustic Waves During Ablation With an Er:YAG Laser," IEEE Access 9, 130543–130553 (2021).
39. W. Huang and R. Kovacevic, "A neural network and multiple regression method for the characterization of the depth of weld penetration in laser welding based on acoustic signatures," J Intell Manuf 22(2), 131–143 (2011).
40. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 2818–2826.
41. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," Int J Comput Vis 128(2), 336–359 (2020).