Academic Thesis: Declaration Of Authorship

I, Zhen Tang, 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.

Title of thesis: Susceptibility weighted image analysis methods for hypoxic-ischaemic encephalopathy prognosis

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. Either none of this work has been published before submission, or parts of this work have been published as:

   Tang, Z., Mahmoodi, S., Dasmahapatra, S., Darekar, A., & Vollmer, B. (2020, July). Ridge Detection and Analysis of Susceptibility-Weighted Magnetic Resonance Imaging in Neonatal Hypoxic-Ischaemic Encephalopathy. In Annual Conference on Medical Image Understanding and Analysis (pp. 307-318). Springer, Cham.

   Tang, Z., Mahmoodi, S., Darekar, A., & Vollmer, B. (2021). Hypoxic-ischaemic encephalopathy prognosis using susceptibility weighted image analysis based on histogram orientation gradient. In 15th International Joint Conference on Biomedical Engineering Systems and Technologies

Signed: .....

Date: .....

UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

School of Electronics and Computer Science

Susceptibility Weighted Image Analysis Methods for Hypoxic-Ischaemic Encephalopathy Prognosis

by

Zhen Tang

A thesis submitted for the degree of

Doctor of Philosophy

August 2022

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

School of Electronics and Computer Science

Doctor of Philosophy

Susceptibility Weighted Image Analysis Methods for Hypoxic-Ischaemic Encephalopathy Prognosis

by Zhen Tang

Neonatal hypoxic-ischaemic encephalopathy (HIE) is a major cause of newborn deaths and neurodevelopmental abnormalities around the world. Susceptibility weighted imaging can provide assistance in the prognosis of neonatal HIE. The propose of this research is to develop a new automated system to assess neonatal brain injury and developmental outcome by detecting and analysing vessel features in SWI images.

In this research, a dataset of SWI images acquired from 42 infants with neonatal HIE is used for feature extraction. Firstly, the ridges representing the veins in the SWI images are detecting to obtain features including the width, intensity value, length of veins, and Hessian eigenvalues for ridges. The normalized histograms of these features are used as feature vector for classification. Individual or concatenated feature vectors are fed into kNN and random forest classifiers to predict the neurological outcomes of infants with HIE at the age of 24 months. We select the balanced SWI dataset to avoid the bias. The feature vectors containing width, intensity, length and eigenvalue show a promising classification accuracy of 78.67% ± 2.58%. Then we use the feature vectors to train support vector regression and random forest regression models for predicting the motor score and cognitive score of infants with HIE assessed by Barley-III at the age of 24 months. Our mean relative errors for cognitive and motor outcome scores are 0.113±0.13 and 0.109±0.067 respectively. The features derived from the ridges of the veins are good predictors of neurological outcome in infants with neonatal HIE.

Further, we design a supervised classifier for automatic prognosis of automated detection of SWI signs of HIE. This classifier also enables to determination of brain regions which have been affected by hypoxic-ischaemic by extracting appropriate features from SWI images. Our classifier can classify the veins in the SWI images into normal and abnormal group by clinical assessment outcomes. The number and location of abnormal veins in the brain of HIE neonates will predict the neurodevelopmental outcomes of infants with HIE at the age of 24 months. Our classifier proposed in this study demonstrates a superior performance in

HIE prognosis for the dataset associated with cognitive and motor outcomes. The accuracy of early prediction of motor outcome at 2 years of age using SWI images in newborns by our classifier achieves 75% ± 13.9%. We also employed the linear regression, polynomial regression, and support vector regression model to predict outcomes and the lower mean relative errors for motor and cognitive outcomes are 0.088±0.073 and 0.101±0.11 respectively.

Then we extract the feature vectors of global and local brain by the histogram of oriented gradient descriptor. We obtain the brain regions associated with motor and cognitive function by image registration. The histogram of oriented gradient feature vectors of these brain regions are fed into the kNN and random forest classifiers to predict the motor and cognitive outcome. The result shows an effective classification for cognitive outcome, which achieved an accuracy of 76.25±10.9.

In addition, we propose a convolutional neural network model to classify the SWI images with HIE. Due to the lack of a large dataset, transfer learning method with fine-tuning a pre-trained ResNet 50 is introduced. we train a convolutional neural network model to classify the SWI images with HIE. Due to the lack of data, transfer learning method with fine-tuning a pre-trained ResNet 50 network is introduced. The balanced datasets are selected randomly to avoid bias in classification. Then we develop a rule-based system to improve the classification performance, with an accuracy of 0.933 ± 0.086. We also compute heatmaps produced by the Grad-CAM technique to analyze which areas of SWI images contributed more to the classification results.

Our research demonstrates that the features derived from the vascular ridges improve the prognostic value of SWI images in HIE. Furthermore, our findings suggest that it is possible to predict neurological, motor, and cognitive outcomes by numerical analysis of their neonatal SWI images and to identify brain regions on SWI affected by HIE.

Table of Contents

List of Figures

List of Tables

Abbreviations

HIE	hypoxic-ischaemic encephalopathy
HI	hypoxic-ischaemic
SWI	susceptibility weighted imaging
CBF	cerebral blood flow
EEG	electroencephalography
CT	computed tomography
MRI	magnetic resonance imaging
ATP	adenosine triphosphate
T1WI	T1-wieghted imaging
T2WI	T2-weighted imaging
T2-FLAIR	T2 fluid attenuated inversion recovery
DWI	Diffusion weighted imaging
MRS	Magnetic resonance spectroscopy
kNN	k-nearest neighbours
SVR	support vector regression
HOG	histogram of oriented gradients
ResNet	residual neural network
Grad-CAM	gradient-weighted class activation mapping
aEEG	amplitude-integrated electroencephalography
DTI	diffusion tensor imaging
PWI	perfusion-weighted imaging
PLIC	posterior limb of the internal capsule
BGT	basal ganglia and thalamic area
IC	internal capsule
NICHD	national institute of child health and human development
TH	therapeutic hypothermia
Bayley-III	Bayley scales of infant and toddler development
ADC	apparent diffusion coefficient
NICU	neonatal intensive care unit

HRS	hierarchical region splitting
ROI	regions of interest
HRA	health research authority
IRAS	integrated research application system
FLASH	fast low angle shot
IR	inversion recovery
PD	proton density
QSM	quantitative susceptibility maps
CMC	cumulative match characteristic
ROC	receiver operating characteristic
SD	standard deviation
RFR	random forest regression
SVR	support vector regression
MRE	mean relative error
MRI	magnetic resonance imaging
MD	mahalanobis distance
ANTs	advanced normalisation tools
PR	polynomial regression
LR	linear regression
CNNs	convolutional neural networks
VGG-Nets	visual geometry group networks
ReLu	rectified linear unit

Acknowledgements

I would like to express my gratitude to my supervisors Dr. Sasan Mahmoodi and Professor Mark Nixon for their support and guidance, and I would also like to thank my parents and Dr. Di Meng. Thanks all my friends for helping me to my research.

Chapter 1

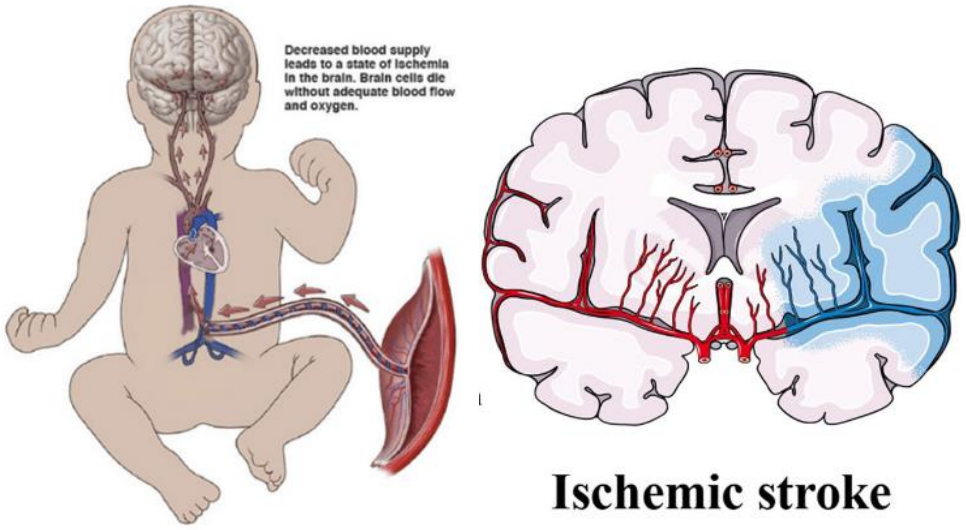
Context and Contributions

1.1. Introduction

Hypoxic-ischaemic (HI) brain injury is a type of neonatal brain damage caused by oxygen deprivation and limited blood flow, and it is an important cause of perinatal death or neurodevelopmental (motor, cognitive, behavioural and speech) impairment in newborns worldwide [1-2]. About 0.2% of infants are affected by HI in developed countries [3]. In addition, 20% to 50% percent of asphyxiated newborns show hypoxic-ischemic encephalopathy (HIE) and die in the neonatal period. As many as 25% of survivors show permanent neuropsychological disorders [4]. HIE carries a high risk for neuro-motor, cognitive and behavioural difficulties, epilepsy, and visual and hearing impairment in survivors. Early diagnosis and assessment of the injury location and its extent is important for counselling and identifying those who may benefit from early intervention [5]. Because of the complexity of HI brain damage, traditional methods of diagnosing HIE results are time-consuming and inefficient [6]. Therefore, the application of an automatic method would be useful to streamline the procedure for specialists to make an early diagnosis.

1.1.1. Aetiology of HIE

Various problems or medical complications may cause HIE, any of which can occur before, during or after the baby is born [7-8]: (1) in the antepartum period, e.g. problems with blood circulation to the placenta, lung malformations; (2) during the intrapartum period, e.g. umbilical cord accidents, prolonged late stages of labour; (3) in the postpartum period, e.g. severe cardiac or pulmonary disease, severe prematurity.



a. b.

Figure 1. 1 a) Decreased blood flow to brain [28] (Reproduced with permission of Ref. 28, Copyright of © 2016 Nucleus Medical Media. All rights reserved.); b) Ischaemic stroke [29] (Reproduced with permission of Ref. 29, Copyright of © 2020 current nanoscience. All rights reserved.)

Perinatal asphyxia is a major cause of HIE in neonates, as it is associated with a lack of blood supply to the brain cells and low oxygen levels, resulting in neonatal brain dysregulation and secondary brain damage [9-10]. Asphyxia can be caused by any factor that impairs blood circulation and gas exchange between the mother and the foetus and causes a decrease in blood oxygen concentration [11]. According to statistics, intrauterine asphyxia occurs in 50% of newborns under normal circumstances, birth asphyxia occurs in about 40% of newborns and only 10% of newborns are born with a congenital condition [12]. The occurrence of perinatal asphyxia is closely related to several factors [11]:

1) umbilical cord prolapse
2) uterine rupture
3) maternal collapse

4) placental abruption

5) maternal comorbidity, etc.

1.1.2. Pathogenesis of HIE

The human brain is supplied with energy by glucose oxidation and its metabolism is so active that it can account for over 50% of the whole body's oxygen consumption [13]. The brain tissue of newborns is very low in glycogen, and therefore cerebral blood circulation is the key to supporting the entire energy supply [7].

Hypoxic leads to damage to the cell membranes of the neonatal brain metabolism, a severe deficiency of adenosine triphosphate, impairment of sodium pump transport channels, lactate build-up and increased carbon dioxide, resulting in metabolic acidosis [10]. The vicious cycle of reduced perfusion and increased permeability of the vascular wall eventually leads to cellular oedema, further exacerbating the ischaemic environment. Excessive leakage of intracellular water molecules, sodium ions and proteins creates neuronal cell death accompanied by intracranial haemorrhage [14]. If prolonged, it often results in irreversible brain damage, which clinically manifests as permanent neurological deficits, such as mental retardation, ataxia, intermittent epilepsy and spasticity, and even cerebral palsy [14]. Figure 1.2 shows that the pathogenesis of hypoxic-ischaemic encephalopathy.

With intensive clinical research, it was found that the main mechanism of neonatal ischaemic-hypoxic encephalopathy lies in the disturbance of cerebral blood flow (CBF). Lactic acid accumulation and local tissue acidosis disrupt the automatic control of CBF, leading to brain cell necrosis [14-15]. The increased venous pressure and damage to the endothelium can cause capillaries or small veins to rupture and bleed due to venous stasis and increased partial pressure of carbon dioxide stimulated by hypoxic [9]. Changes in CBF include chronic ischemia and hypoxic (damage to the brain's selective vulnerability), acute ischemia and hypoxic (damage to the thalamus and brainstem nuclei) and impaired vascular homeostasis [16].

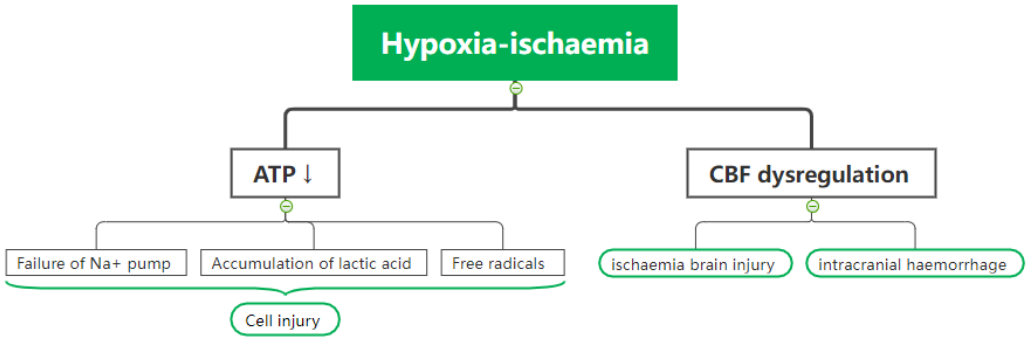


Figure 1. 2 Pathogenesis of hypoxic-ischaemic encephalopathy

1.1.3. Medical image analysis of HIE

Besides the clinical presentation, the diagnosis of HIE can be made by electroencephalography (EEG), brainstem auditory evoked potentials, transcranial Doppler ultrasound, cranial computed tomography (CT) and cranial magnetic resonance imaging (MRI) to detect brain damage in newborns [17].

EEG in neonates could be susceptible to interference. And increased cerebral blood flow will cause a decrease in brain wave amplitude and fast waviness. Transcranial Doppler ultrasound has a limited scanning range as the ultrasound probe probes through the fontanelle. Brainstem auditory evoked potentials require relatively undisturbed shielding. CT has the inherent disadvantage of radiation exposure [17-18].

In contrast, MRI is increasingly used in the diagnosis and prognostic assessment of HIE because of its radiation-free, high soft-tissue resolution and the ability to image multiple sequences and modalities to comprehensively assess the localisation and extent of the lesion [18].

1) Conventional MRI sequences

Conventional MRI sequences, including T1WI, T2WI and T2-FLAIR, can clarify the location and extent of HIE lesions and their relationship to surrounding structures. The water content of the brain is high in the neonatal period. Some of the white matter has not yet been myelinated, and there is little difference between the water content in the cortex and white matter. It is difficult to distinguish grey matter from white matter in conventional T1W1 and T2W1 [19].

2) Diffusion weighted imaging

Diffusion weighted imaging (DWI) is more sensitive to HI diffusion restriction of brain cells, especially for the display of lesions in the white matter areas of the brain, the internal capsule, and parts of the corpus callosum [20].

3) Perfusion weighted imaging

Perfusion weighted imaging assesses local tissue viability and function by quantifying blood perfusion at the capillary level in brain cells. Arterial spin-labeled perfusion MR imaging is used to assess cerebral blood flow in hypoxic ischaemic [21].

4) Diffusion tensor imaging

Diffusion tensor imaging is a functional imaging technique based on the directionality and integrity of in vivo cerebral white matter fibrils presented on DWI, which reflects the diffusion characteristics of water molecules in various directions by diffusion tensor and vector. Diffusion tensor imaging can evaluate HIE-induced white matter damage by showing abnormalities in the white matter fibrils of the brain [22].

5) Magnetic resonance spectroscopy

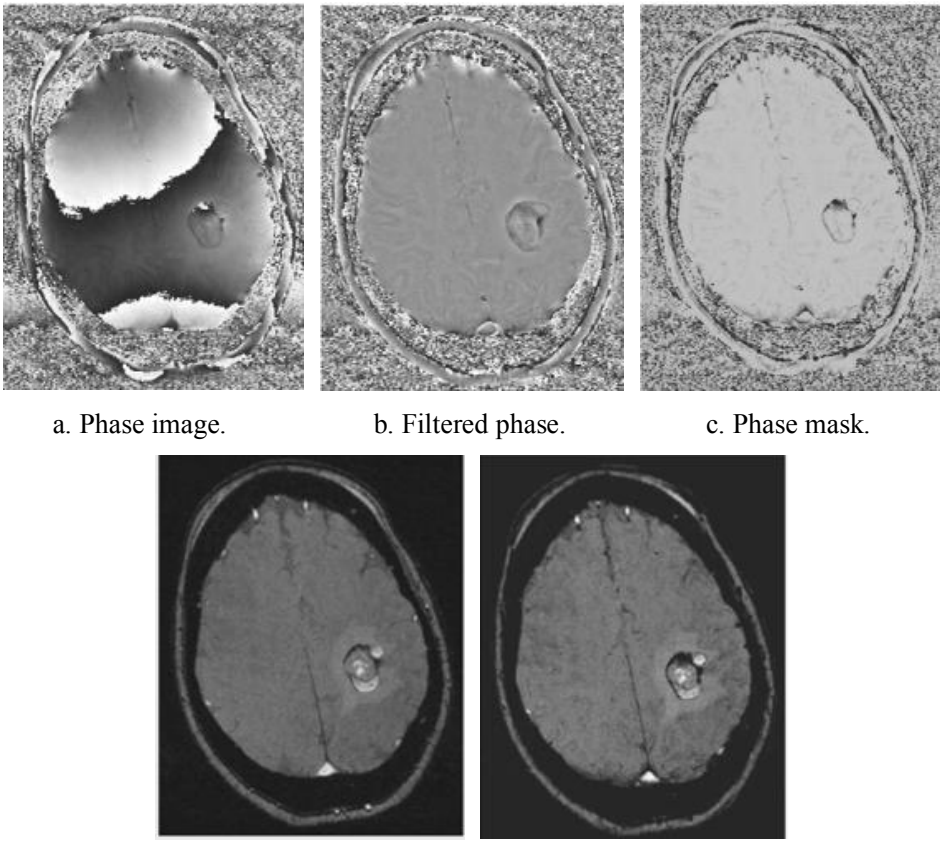
Magnetic resonance spectroscopy (MRS) imaging allows the study of metabolite levels in living cells and quantifies biochemical characteristics. MRS allows quantitative monitoring of changes in lactate, choline, creatine and glutamine concentrations and can provide quantitative evidence of tissue hypoxic in HIE [23].

6) Susceptibility weighted imaging

Susceptibility weighted imaging (SWI) is based on thin high-resolution scans of gradient-echo sequences combined with three-dimensional fully flow-compensated sequence reconstruction. Dynamic-compensated sequence reconstruction is very sensitive to local magnetic field inhomogeneities such as calcifications, haemorrhages, iron deposits; e.g. haemorrhage, and iron deposition, can identify small foci of haemorrhage in the brain parenchyma. SWI has potential diagnostic value in neonatal HIE and confirms that the pathological changes in HIE are closely related to the nature and extent of the injury. SWI sequences are significantly better than conventional sequences in detecting and displaying haemorrhagic foci in children with HIE [24].

1.1.4. Why use SWI in HIE

Magnetic SWI is a new MRI technique that is widely used in clinical practice to highlight paramagnetic material in the blood by exploiting differences in magnetic susceptibility between different tissues [24-25]. Its considerable sensitivity detects intracerebral haemorrhage, venous vasculature and iron deposition. In addition to the magnitude information used by conventional MRI, SWI utilises phase information and undergoes image post-processing to combine the phase image with the magnitude image to create a unique image contrast [24]. As shown in Figure 1.3, due to the large phase difference between the vessel and the surrounding tissue, SWI can obtaine a phase image (see in Fig1.3 (a)) using a high resolution scan. In order to remove the low frequency phase interference caused by the uneven background magnetic field, a low-pass filter (see in Fig1.3 (b)) is applied to the original phase image to obtain a phase mask (see in Fig1.3 (c)). Then the phase mask is multiplied several times with the magnitude image (see in Fig1.3 (d)) provided by the MRI to obtain a clear SWI image (see in Fig1.3 (e)) [25]. It is noted that figure1.3 was provided by the hospital and will therefore be different from the other SWI images in this thesis.



a. Phase image. b. Filtered phase. c. Phase mask.

d. Magnitude image. e. Susceptibility weighted imaging.

Figure 1. 3 Susceptibility weighted image structure [25]

1.2 Contributions

The main contributions in the thesis arising from this research, are as follows:

1 We developed a new automated system to classify SWI images to evaluate neonatal HI injury by detecting and analysing ridges within these images. In this research, a dataset of SWI–MRI results acquired from 42 infants with HIE during the neonatal period was obtained based on ridge analysis of the SWI images. All data is provided by the hospital. The results of each patient's follow-up are also assessed by a team of experts. More details of the data will be discussed in Chapter 3.

Features including the width of the blood vessels, the change in intensity of the veins' pixels in comparison with neighbouring pixels, the length of the blood vessels and Hessian eigenvalues for the ridges were extracted. Normalised histogram parameters in the single or combined features were used to classify SWI images by k-nearest neighbours (*kNN*) and random forest classifiers. The mean and standard deviation of the classification accuracies were derived by randomly selecting 11 datasets ten times from those with a normal neurological outcome (n = 31) at age of 24 months and those with an abnormal neurological outcome (n = 11) to avoid classification biases due to any imbalanced data. We also employed support vector regression (SVR) to predict the scores of motor and cognitive outcomes assessed 24 months after the birth.

This is the first time to use vein's features of neonatal brain with HIE on the SWI images for prognostication. In [76], only hessian eigenvalue of veins in neonatal brain was extracted for classification of patients with HIE and normal infants group. Compared to the work in [76], we extracted more features from vessels, and could predict neurodevelopmental outcomes at two years old by using SWI images of infants with HIE at birth. In the same classifier, the accuracy by using our features with 78.67±2.58% is better than the accuracy obtained using the features in [76], which is 72.27±4.85%.

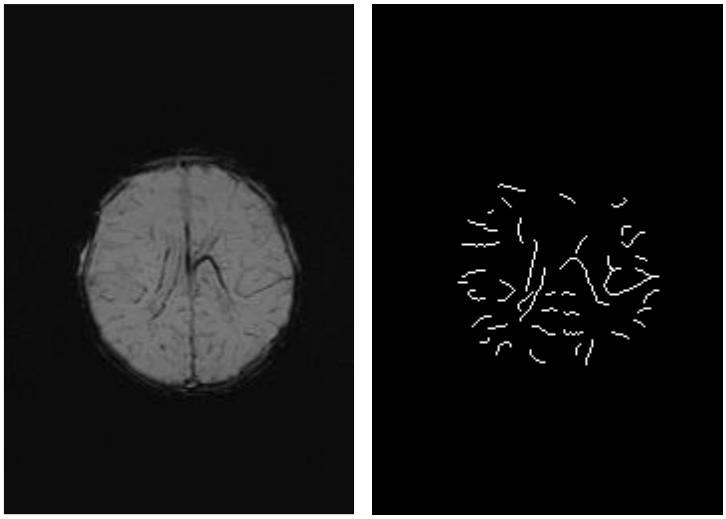


Figure 1. 4 An example of ridge detection [26]

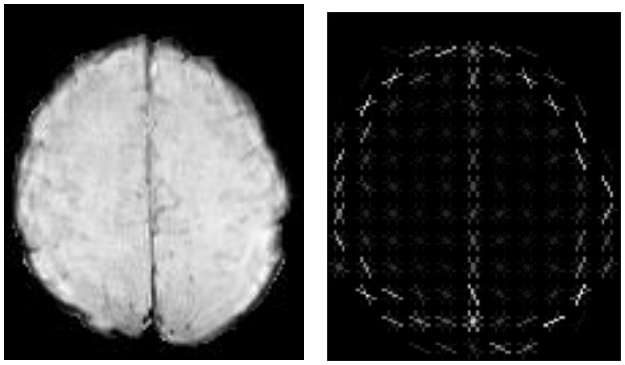
2 We propose an algorithm to separate vessels damaged by HIE from normal vessels. Such an algorithm would then enable us to find the regions of the brain affected by HI injury, and therefore it would help us assess the associations of the affected brain regions with neurodevelopmental outcomes. We have measured the performance of our algorithm in various scenarios. By using our algorithm proposed in this paper in brain regions implemented in the outcomes of interest, we have found some measurements showing strong correlations with outcomes determined by clinical experts who have examined the patients two years after SWI images have been taken.

This is the first study to classify in the appearance of vessels in SW images into two normal and abnormal groups with indirect supervisions. In the previous studies [53-54, 56], the damaged areas of the brain in infants with HIE were usually identified by specialists. Our work avoids the time-consuming and complicated prognosis of neonatal HIE depending on neurologists.

3 We analysed the SWI images by using a histogram of oriented gradients (HOG) as a global feature to identify areas of the neonatal brain affected by HIE. Forty-two infants with neonatal HIE have undergone SWI in the neonatal period and have been investigated through neurodevelopmental assessment at 24 months of age. HOG features are used to represent the SWI images of the whole brain and the region of interest separated from the brain image registration algorithm. We used *kNN* and random forest to classify the SWI images into normal and abnormal groups, and then we compared our results to our previous work. Figure 1.5 shows the HOG of an SWI

image with HIE. In order to clearly demonstrate the results of HOG applied to SWI, we have chosen a SWI image with the background noise removed.

We use the HOG to extract features of neonatal brain with HIE on SWI images, which is more comprehensive than the study in [75]. In [75], only the deep medullary vein of the neonatal brain after HIE was used as the region of interest. However we use the HOG features from different areas in the brain. In [75], the rates of region of interest from HIE group and healthy group were used to compare to distinguish between infants with HIE and healthy infants. Our method predicts the results directly using the HOG features of the region of interest.



a. Original SWI image b. HOG of SWI image

Figure 1. 5 An example of histogram of oriented gradient of SWI image [27]

4 In this research, we used deep learning methods to classify neonatal HIE SWI images by neurology outcome. We used transfer learning of ResNet 50 to train the SWI images with normal and abnormal neurology outcomes at 24 months of age in neonatal HIE. And then we computed heatmaps based on transfer learning to demonstrate which region of the brain contributed most to the classification of neurology outcomes at 24 months of age from SWI images of infants at birth. For abnormal patients, such regions can be interpreted as regions of the brain which have been damaged by HIE.

This is also the first study to use the Grad-CAM technique to visualize the regions in SWI images affected by HIE. The results of the study support that injury to the motor areas can be demonstrated on SWI and that what is detected on SWI is consistent with anatomy and physiology, i.e. early injury to the motor areas leads to cerebral palsy. For recent studies on the application of deep learning to HIE [122], by training a gradient-enhanced deep learning model through using performing feature selection on

MRI, the prediction of poor motor outcomes after HIE achieved an accuracy of 85%. However, the accuracy of our deep networks based on rule-system reached 93.3%.

1.3 Thesis overview

- ***Chapter 2: Review of the Development of Hypoxic-Ischaemic Encephalopathy Diagnostic Methods***

In this chapter, we will review the literatures on the diagnosis and long-term neurodevelopmental outcomes of neonatal HIE. We describe the scoring systems for neonatal HIE. The use of neuroimaging in neonatal HIE is also presented. We also introduce the development of computer vision techniques in detecting neonatal HIE.

- ***Chapter 3: Dataset***

This chapter presents the data we used in our experiments, including ethics, MRI protocols, and information of infants with HIE.

- ***Chapter 4: Ridge Detection and Analysis of Susceptibility-Weighted Imaging in Neonatal Hypoxic-Ischaemic Encephalopathy***

In this chapter, we will introduce an automated system for hypoxic-ischemic encephalopathy diagnosis by analysing SWI images. The features of veins in the SWI images such as width, intensity and length will be extracted by ridge detection. Finally, *kNN* and random forest classifiers will be used to classify the SWI images of neonates with HIE. These feature will be employed as regression analysis for motor and cognitive outcome.

- ***Chapter 5: Brain Regions Affected by Hypoxic-Ischaemic***

In this chapter, we will further analyse the features extracted from SWI images of infants with HIE to identify brain regions affected by neonatal HIE. For different neurodevelopmental outcomes of HIE survivors, we design ***algorithm 1*** to detect the location of the veins affected by hypoxic-ischaemic in the brain.

- ***Chapter 6: Hypoxic-Ischaemic Encephalopathy Prognosis Using Susceptibility Weighted Image Analysis Based on Histogram Orientation Gradient***

This chapter will describe the Histogram of Oriented Gradients (HOG) descriptor to extract global and regional features from SWI images. These HOG features are used to classify HIE infants with neurodevelopmental outcomes at the age of 24 months.

- ***Chapter 7: Deep Learning for HIE Prognoses***

This chapter will introduce the application of convolutional neural networks for the classification of SWI images with HIE. In order to consider the issue of SWI image contamination by noise and artefacts in the analysis of the classification of patients in the test group, we introduce two rules (**Rule 1** and **Rule 2**) to avoid these issues. The Grad-CAM is used to analysis which areas of SWI images contributes more to the classification results.

- ***Chapter 8: Conclusions and Future Works***

In this chapter, conclusions are reached and ways of future works are discussed.

1.4 List of Publications

Here is the list of my publications from this research work:

[1] Tang, Z., Mahmoodi, S., Dasmahapatra, S., Darekar, A., & Vollmer, B. (2020, July). Ridge Detection and Analysis of Susceptibility-Weighted Magnetic Resonance Imaging in Neonatal Hypoxic-Ischaemic Encephalopathy. In *Annual Conference on Medical Image Understanding and Analysis* (pp. 307-318). Springer, Cham.

[2] Tang, Z., Mahmoodi, S., Darekar, A., & Vollmer, B. (2021). Hypoxic-ischaemic encephalopathy prognosis using susceptibility weighted image analysis based on histogram orientation gradient. In *15th International Joint Conference on Biomedical Engineering Systems and Technologies*

[3] Tang, Z., Mahmoodi, S., Darekar, A., & Vollmer, B. (2022). Automatic Veins Analysis of Susceptibility Weighted Image in Hypoxic-Ischaemic Encephalopathy. *Magnetic Resonance Imaging* (under review after revisions)

[4] Tang, Z., Mahmoodi, S., Darekar, A., & Vollmer, B. (2022). Application of deep learning for neonatal hypoxic-ischaemic encephalopathy based on susceptibility weighted imaging. *Pattern Recognition Letters* (in submission)

Chapter 2

Review of the Development of Hypoxic-Ischaemic Encephalopathy Diagnostic Methods

Since 1940, experiments have demonstrated that perinatal asphyxia causes brain damage [30], which means failure to establish breathing at birth could cause injury to the brain. As computer technology has developed, medical imaging has become an important tool for assessing hypoxic-ischaemic injury in infants. Tests to confirm HIE after symptoms have been detected include ultrasound, electroencephalogram (EEG) or amplitude-integrated electroencephalography (aEEG), computerised tomography (CT), magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS) [31]. Of these, aEEG is used to monitor seizure activity due to its longer time period than conventional EEG. The introduction of MRI enabled the presence, distribution and severity of acute lesions caused by perinatal cerebral hypoxic-ischaemic to be determined. It also allowed practitioners to detect brain damage and determine its severity. MRI is the most common imaging used for evaluating infants with HIE [17]. After deeper research into neonatal brains injured by hypoxic-ischaemia, different methods of MRI were invented, such as T1- and T2-weighted [19], diffusion-weighted (DWI)

[20], diffusion tensor (DTI) [22], perfusion-weighted (PWI) [21] and susceptibility-weighted (SWI) [24]. Of these, SWI is the newest MRI scan to be used on infants with HIE.

This chapter reviews the neonatal HIE literature, which has been grouped into clinical diagnosis, medical imaging for diagnosis and prognosis and two-dimensional cerebral slices analysis using computer vision technologies. This study focuses on literature concerning image processing techniques when using MRI on infants with HIE.

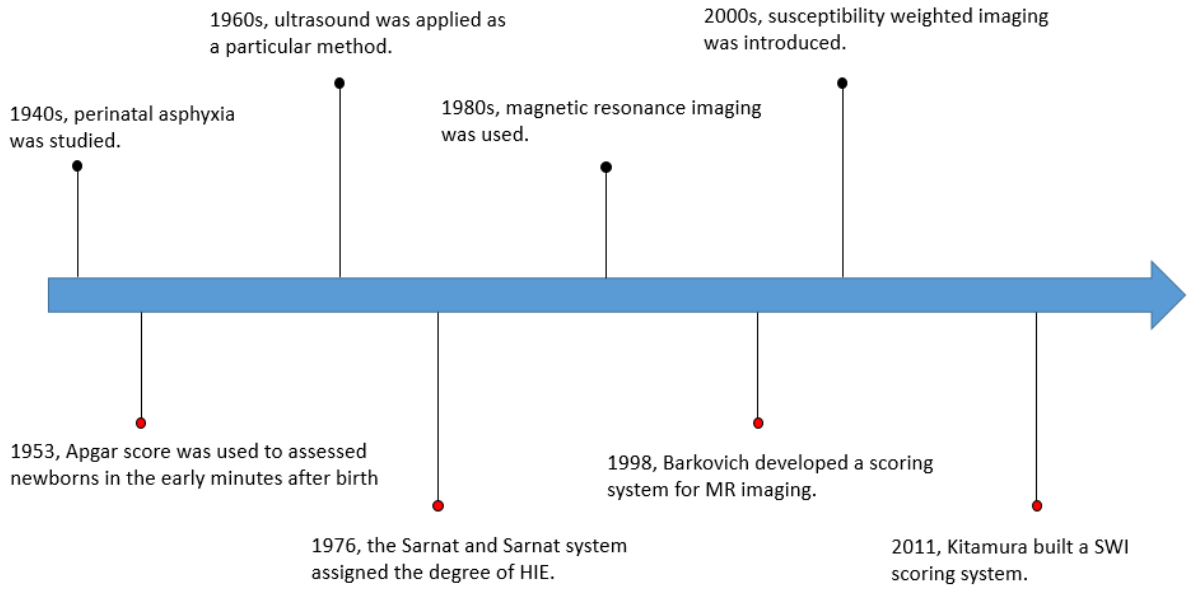


Figure 2. 1 Neonatal HIE diagnostics development

2.1. Diagnosing hypoxic-ischaemic encephalopathy

2.1.1. Clinical diagnosis

After discovering that perinatal asphyxia caused brain damage in newborns, Apgar defined perinatal asphyxia by assessing the physiological condition of an infant soon after birth, resulting in the Apgar score [30]. Sarnat and Sarnat (1976) combined the clinical signs of the newborn with EEG results to create a staged system that classifies neonatal encephalopathy into three categories: mild (Stage 1), moderate (Stage 2) and severe (Stage 3) [32]; Thompson then established an HIE scoring system based on this that included clinical assessment of nine signs: tone, level of consciousness, fits, posture, Moro reflex, grasp, sucking reflex, respiration and fontanel [33].

2.1.2. Neuroimaging methods and scoring systems

Due to the widespread use of neuroimaging in medicine, HIE is diagnosed using clinical signs and neuroimaging (CT, MRI and MRS). Figure 2.1 illustrates the development of the diagnostic method and neuroimaging to diagnose HIE in infants.

Before different neuroimaging methods were introduced, EEG and aEEG were important tools for evaluating neonatal HIE [34]. At the neurodevelopmental follow-up, 84 infants were alive; 64%, 21%, and 15% had normal, moderate or severe relapses, respectively. Their study showed that the EEG results between 36 and 60 h were closely related to prognosis. If the results in the first 48 hours were normal or slightly abnormal, then 95% of prognoses were normal or moderately impaired. Severe EEG abnormalities during the first 48 hours and 96% were considered poor prognoses (severe relapse or death) after two years [35].

Cranial ultrasonography and CT were the first methods used for evaluating HIE outcome in infants [36-37]. Cranial ultrasonography is a simple, non-invasive method that effectively determines the pattern and degree of hypoxic-ischaemic brain injury in newborns, particularly those born prematurely [36]. Although CT can detect haemorrhage well, MRI is better at assessing infants with HIE. However, CT is not recommended for neonates due to the radiation [31].

Since the 1980s, MRI has been used to assess infants with HIE, as it is better than cranial ultrasonography and CT for determining the location and degree of brain damage; it has become the standard tool for diagnosing HIE [38]. Magnetic resonance imaging sequences have evolved from conventional T1- and T2-weighted, DWI, to DTI and SWI. There are different image contrasts in the final brain images of infants with HIE due to the different parameters in different MRI sequences, such as the repetition and echo times [18]. T1-weighted sequences are useful for measuring myelination in neonates' brains with HIE, and T2-weighted sequences can show cortical damage [39]. DWI can detect cerebral oedema caused by ischaemia and hypoxia [20]. DTI assesses structural abnormalities in the brains of infants with HIE using the directionality of water diffusion based on DWI [22]. SWI is very sensitive to venous blood and haemorrhage [40].

The development of different methods of using MRI to diagnose HIE has led to various scoring systems. Barkovich et al. [41] graded infants with HIE using T1- and T2-weighted images. The severity of brain damage was scored from 0 to 4 according to the degree of basal ganglia and watershed. A normal MRI was given a score of 0, and an abnormal signal in the basal ganglia

or thalamus was given a score of 1. A score of 2 indicated an abnormal signal in the cortex, a score of 3 meant abnormal signals in the cortex and basal nuclei, and a score of 4 represented abnormal signals in the whole cortex and basal nuclei. Rutherford et al.[42] also used T1- and T2-weighted images to investigate the relationship between the degree of damage to the posterior limb of the internal capsule (PLIC) and the neurodevelopmental outcome of infants with HIE at one year. Full-term newborns with HIE were graded according to the Sarnat system. All infants with abnormal PLIC signals showed neurodevelopmental impairment, whereas almost all neonates with normal PLIC signals had normal motor outcomes. A scoring system developed by Weeke et al.[43] contained 19 regions, including grey matter, cortex, corpus callosum and haemorrhage, with scores ranging from 0 (normal) to 57 (severe). This system assesses the damage, such as whether it is unilateral or bilateral, using T1- and T2-weighted imaging and DWI. The National Institute of Child Health and Human Development's Neonatal Research Network developed a scoring system that grades brain damage from a score of 0 (normal MRI) to 3 (cerebral hemispheric devastation) [44]. Four categories were used to arrive at a score of 1 or 2: 1A, lesions in the frontal and parietal subcortical areas; 1B, more extensive cerebral lesions in the frontal, parietal and occipital subcortical areas; 2A, lesions in the basal ganglia and thalamic area (BGT) and internal capsule (IC); 2B, lesions in the BGT, IC and cerebral areas.

Recently, Kitamura et al. [45] developed a grading scale system using deoxyhaemoglobin levels in the deep cerebral veins shown by SWI. This scoring system focuses on examining the prominent parts of the deep medullary veins rather than the entire venous system, because they are more consistently present than the cortical veins. The medullary vein is a small cerebral vein located in the deep white matter of the brain, and it is responsible for the flow of blood from the deep white matter of the brain to the cortical venous system. Each SWI sequence was assessed using seven categories: 1, absent deep medullary veins; 2, faint deep medullary veins; 3, minimal deep medullary veins; 4, mildly prominent deep medullary veins; 5, moderately prominent deep medullary veins; 6, very prominent deep medullary veins; 7, extremely prominent deep medullary veins. Depend on the prominence of vein value, patients were divided into two groups: normal (with grading of 2-4) and abnormal (with grading of 1 and 5-7). Variability between interobservers in the susceptibility-weighted imaging classification was analyzed by calculating *k* values. For agreement between the two neuroradiologists on the prominence of vein value ", the *k* value was 0.82. For the differences in clinical variables

between the good and poor outcome groups, the presence of clinical seizures was the only variable associated with worse neurological outcome ($p = 0.02$).

Figure 2.2 shows SWI images with HIE graded from 1 to 7 (without 6) in the Kitamura scoring system. Kitamura scores for all SWI images are assessed by doctors with experience in neurology and neonatal disciplines at Southampton Hospital. Due to the limited dataset, there is no set of SWI images with a grading of 6. Table 2.1. illustrates the five scoring systems. Different scoring systems and MRI images will be used in different cases. In clinical practice, different scoring systems are used depending on the modality of the MRI image. The patient's T1- and T2- weighted images are usually assessed using the Barkovich scoring system, the Rutherford scoring system and the NICHD scoring system. The patient's DWI images are assessed using the Weeke scoring system and the SWI images are assessed using the Kitamura scoring system. In the field of medical image processing, T1- and T2- weighted images were mostly used early to predict the Barkovich score of newborns. With the development of MRI, more studies have used DWI images of patients to predict neurological outcomes. Today, SWI images are a more useful tool for predicting neonatal outcome because of their sensitivity to HIE.

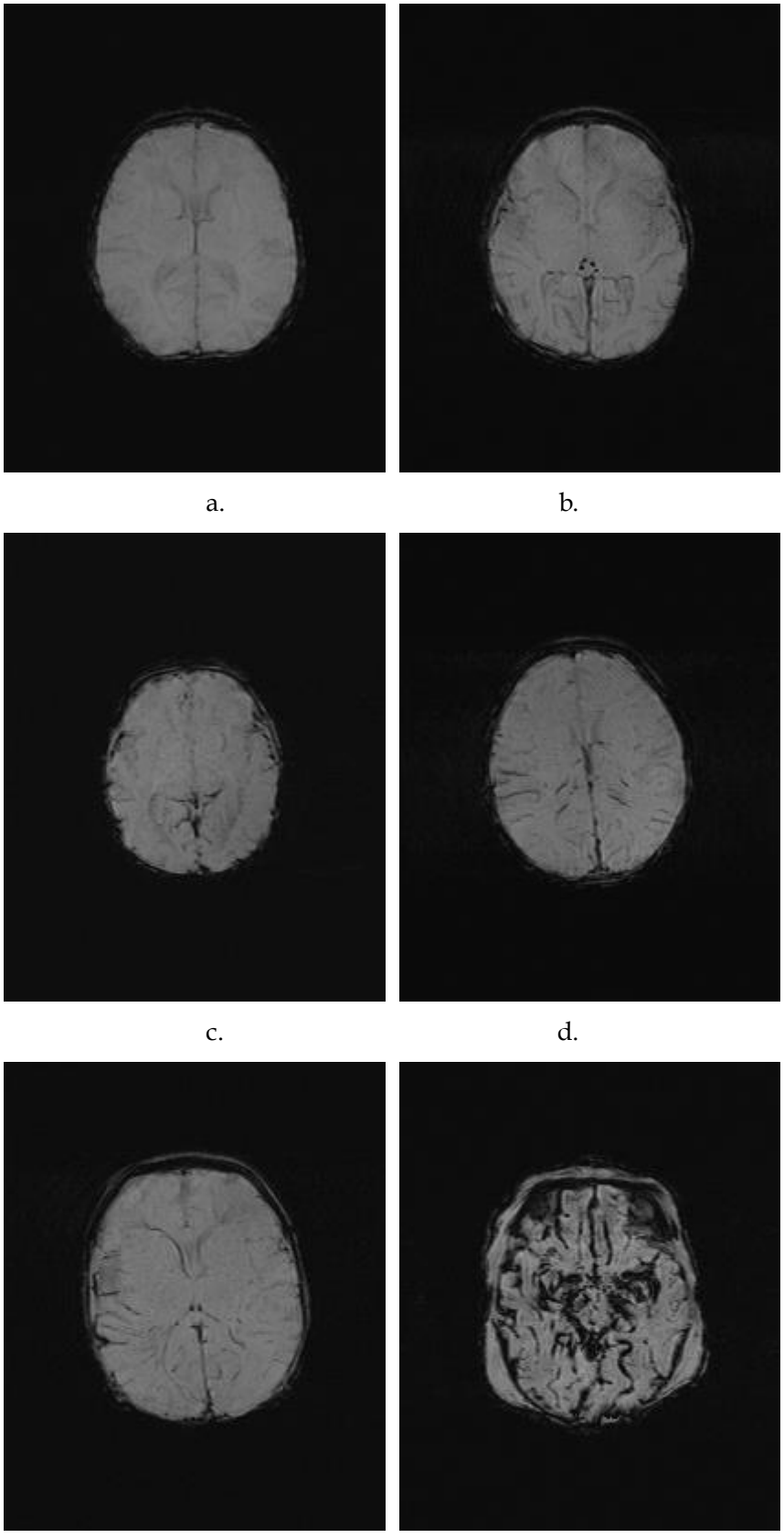


Figure 2. 2 Susceptibility-weighted images showing the categories, except grading 6. a) Grading 1; b) Grading 2; c) Grading 3; d) Grading 4; e) Grading 5; f) Grading 7.

Table 2. 1 The five scoring system

Study	MRI method	Grading
Barkovich et al. [41]	T1- and T2-weighted imaging	0: normal MRI 1: abnormal signals in the basal ganglia or thalamus 2: abnormal signals in the cortex 3: abnormal signals in the cortex and basal nuclei 4: abnormal signals in the whole cortex and basal nuclei
Rutherford et al. [42]	T1- and T2-weighted imaging	PLIC equivocal or abnormal
Weeke et al. [43]	T1- and T2-weighted imaging, DWI	Grey matter Proton MRS White matter Cerebellum Intraventricular haemorrhage
NICHD [44]	T1- and T2-weighted imaging	0: normal 1A: lesions in the frontal and parietal subcortical areas 1B: more extensive cerebral lesions in the frontal, parietal and occipital subcortical areas 2A: lesions in the BGT and IC 2B: lesions in the BGT, IC and cerebral areas 3: cerebral hemispheric devastation.
Kitamura et al. [45]	SWI	1: absent deep medullary veins 2: faint deep medullary veins

		3: minimal deep medullary veins 4: mildly prominent deep medullary veins 5: moderately prominent deep medullary veins 6: very prominent deep medullary veins 7: extremely prominent deep medullary veins

Following the widespread use of neuroimaging, therapeutic hypothermia (TH) has become standard care for infants with HIE and can increase survival [46-48]. Infants who survive HIE are at higher risk of neurodevelopmental impairment due to damaged neural networks, motor impairment and cognitive defects. Follow-up and evaluation of affected infants are necessary to know the long-term effects of HIE and the prognostic value of neuroimaging [49-51].

The British Association of Perinatal Medicine and the National Institute for Clinical Excellence suggest formal neurodevelopmental assessments of surviving infants at around two years of age [54] using methods such as the Bayley Scales of Infant and Toddler Development [52], which is a developmental assessment tool widely used to assess children from 1 to 42 months. Its third edition, Bayley-III, covers cognitive, language, motor, adaptive and social-emotional development. After TH, infants with HIE may still have motor and cognitive impairments [53].

Edmonds et al. [54] assessed cognitive and behavioural outcomes in children, especially those without neuromotor impairment; 87 two-year-old children with HIE were tracked. For children (without cerebral palsy,) tested using the Bayley-III scales (n = 71), 94.4% had normal cognitive scores, 4.2% mild cognitive delay, and 1.4% severe cognitive delay. Furthermore, 98.5% had normal motor scores (>85), 1.5% had a mild motor delay and none had a severe motor delay.

Petra et al. [56] described the motor and behavioural outcome at two years old after perinatal HIE and explored the relationship between outcome and neonatal MRI results. Seventeen infants with HIE were classified as Sarnat Stage I (mild HIE), Stage II (moderate HIE) and Stage III (severe HIE). All infants with HIE had an MRI, including T1- and T2-weighted

imaging. A Barkovich scoring system was used for MRI assessment, and the second version of the Bayley Scales (Bayley-II) was used for assessing motor outcomes. The results showed that four out of six infants with mild HIE at birth had a mild delay at two, while seven of eleven with moderate HIE at birth had a mild delay at the same age. However, the infants mildly delayed at age two had a normal MRI assessment at birth.

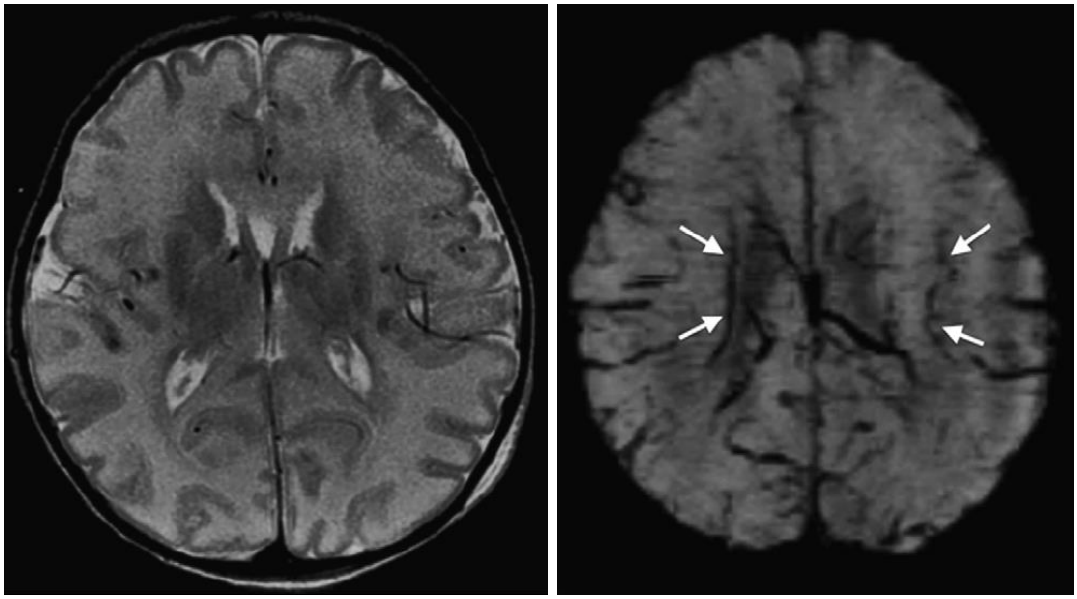
2.1.3. Magnetic resonance imaging as a predictor

Recently, studies have examined using the MRI results of infants with HIE to provide clinicians with diagnostic information and have also explored the relationships between brain injury, neonatal MRI and the outcomes of infants aged 6 to 24 months. It is now common to use neonatal brain MRI results to predict the neurodevelopmental outcomes for infants with HIE to aid prognosis and home care [56].

Haataja et al. [57] collected the MRI T1- and T2-weighted images of 53 tern infants with HIE within 4 week from delivery. The images were assessed using the Rutherford system [42], neurological findings were scored at 9 to 14 months and motor function was assessed at two years. The results showed that the best scores were obtained from infants with normal or mild neonatal MRI findings. The lowest scores were associated with severe basal ganglia and white matter lesions, and infants with low neurological examination scores had severely limited motor function at two years. In [58], 36 infants diagnosed with HIE within 24 hours after birth were scanned using DWI and graded according to the Sarnat system. DWI showed high sensitivity (100%) but very low specificity (20%) in detecting permanent neurological sequelae when comparing the neurological findings from six-month-old infants. In [56], DWI results and apparent diffusion coefficient (ADC) mapping of 26 infants with HIE at an average of 5.4 days scored using the Barkovich system were associated with neurodevelopmental outcomes at two years of age. In addition, the results suggested that MRI performed better than cranial ultrasonography imaging. Goergen et al. [59] analysed the DWI results of 19 infants with HIE at a mean of five days old, and their development at two was assessed using Bayley-III. The results showed a significant association between the DWI results and outcome at two. Francesca et al. [60] utilised EEG and MRI results as predictors for neurodevelopmental outcomes using Bayley-III at 12 and 18 months. Reduced activity of neonates with HIE at 72 hours from birth shown in EEG results and diffuse changes in the basal ganglia registered by MRI were associated with poor neurodevelopmental outcomes at 18 months. Massaro et al. [2] showed that corpus callosum and corticospinal tract in 52 infants with HIE correlated with

neurodevelopmental outcomes assessed by Bayley-II at 15 and 21 months after hypothermia treatment. Nanavati et al. [61] determined the relationship between the MRI images of 17 newborns with HIE graded according to the Barkovich system and the scores of second edition Bayley at 12 and 24 months. Massaro et al. [5] assessed the relationship between brain structure and function in the neonatal brain using DTI and network neurobehavioral scale in 45 neonates with HIE scanned eight days after birth. There was a significant correlation between corticospinal tract integrity and DTI and NICU Network Neurobehavioral Scale neuromotor performance. Trivedi et al. [62] evaluated the MRI results of neonates with moderate-to-severe HIE treated by TH, including white matter, cortical, cerebellar and brainstem. The poorer cognitive and motor outcomes were evaluated using Bayley-III at 18 to 24 months and were associated with higher MRI impairment results. Lakatos et al. [63] demonstrated the prognostic benefit of early MRI combined with MRS in neonates suffering from HIE. Tharmapoopathy et al. [64] evaluated 55 infants at birth and 18 months with HIE using the Rutherford MRI system and Bayley-III. Abnormal MRI patterns (e.g., BGT, white matter and PLIC) were associated with poor neurodevelopmental outcomes (e.g., cognitive, language and motor).

Haack introduced an SWI sequence to MRI, which performed better in diagnosing cerebral vascular pathology and blood products of the neonatal brain [24,65]. SWI can reveal cerebral vascular lesions that other MRI methods may miss [66]. Steven et al. [67] compared the findings of serial MRI, DWI and SWI in an infant with HIE and showed that SWI could reveal areas suffered from hypoxic-ischaemic brain injury before showed on DWI. Early DWI results could be normal in some patients, but SWI would find damage. This suggests that SWI could play a key role in early diagnosis. Meoded et al. [66] found that SWI may be particularly useful in recognising intramedullary veins with low signals in early neonatal HIE. The SWI results of infants with HIE can also emphasise deep subventricular veins beyond T2-weighted image results (see Figure 2.3); this shows that SWI can aid the diagnosis of HIE. Tsui et al. [68] reported a case of a child with cerebral oedema. All intracranial vessels and the micro lesions that may accompany vascular lesions are well depicted on SWI. Li et al. [69] explored the correlation between the degree of brain damage, SWI grading, and MRS grading in neonatal HIE by comparing the detection rate of intracranial haemorrhagic foci and the number of damaged lobes on SWI and MRS. The results of 60 neonatal HIE cases were used for statistical analysis. SWI and MRS images showed that most infants with mild HIE had fewer than two damaged brain lobes, and all infants with severe HIE showed diffuse damage in whole brain.



a. b.

Figure 2. 3 An infant with HIE scanned at five days. a) T2-weighted image; b) SWI image [66]

2.2. Image processing techniques

Diagnosing and treating neonatal HIE by neurologists are time-consuming. Using image processing techniques to automate the diagnosis and predict outcomes is a major research subject [70]. This section describes the application of medical image processing techniques to MRI of infants with HIE in recent years

2.2.1. T1-, T2-, and diffusion-weighted images

As the first MRI sequence to aid the diagnosis of neonatal HIE, T1- and T2-weighted imaging was also the first sequence to be automated.

Ghosh et al. [70] developed a computer imaging method, hierarchical region splitting (HRS), an imaging method that shows ischaemic lesions in neonatal hypoxic-ischaemic injury (HII) on T2-weighted imaging (T2WI). HRS is a region segmentation method that recursively segments the values from T2WI. Since different brain tissues have different contrast levels, a uniform area could represent a single brain tissue type. Ghosh et al. [70] rescaled the signal value of T2WI to a range (0–255) after removing the skull and the background noise. The signal spectrum histogram of T2WI was then calculated and normalised. Finally, the T2WI images were segmented repeatedly based on the threshold set by the histogram and obtained an image from each part to form a tree structure. HRS was stopped when the split region was considered a unified region. Manual segmentation was used to get the same brain tissue in T2WI for

comparison. The HRS method's sensitivity, specificity and similarity combined with manual detection were 0.82, 0.86 and 1.47, respectively. Ghosh et al. [71] extended the functionality of HRS to quantify core–penumbra from T2WI or ADC maps of neonatal HII to animal models. Neonatal HII resulted in irreversibly injured core and salvageable penumbral tissues. Identification and quantification of salvageable tissue is the key to safe intervention. Their data showed that changes to specific MRI signals in damaged tissues reflect subtle differences in the characteristics of core–penumbra tissues, which can be detected and quantified by HRS. The problem with this method is that the threshold for segmenting MRI images cannot be adapted to other datasets and the parameters need to be constantly modified to find the optimum.

Sarioglu et al. [72] defined moderate-to-severe and mild HIE in infants by assessing the texture of basal ganglia and thalami in T1- and T2-weighted images and ADC mappings. Thirty-five infants with HIE (15 mild and 20 moderate-to-severe) were in the experimental group, and 33 infants with normal brains were in the control group. A two-dimensional axial of T1- and T2-weighted images and ADC maps were input to LifeX software to calculate regions of interest. The grey levels were adapted to 128, and all pixel values were automatically scaled with an average ±3 standard deviation. Texture analysis of basal ganglia and thalami obtained from LifeX were calculated into 47 features: seven conventional values, two shapes, six histograms, seven grey-level co-occurrence matrices, eleven grey-level run-lengths, three neighbourhood grey-level difference matrices and eleven grey-level zone length matrices. Statistical analysis calculated the basal ganglia and thalami texture values for mild and moderate-to-severe HIE groups for T1- and T2-weighted images and ADC maps as the *p*-value. All features with *p*-values less than 0.001 served as independent predictors for binary logistic regression analysis to predict moderate-to-severe HIE. Among these features, the histogram entropy values of basal ganglia on the ADC maps showed the best diagnostic performance for infants with moderate-to-severe HIE, with a sensitivity of 95%, a specificity of 93.3% and an accuracy of 94.3%. Manually drawing the region of interest made it time-consuming to calculate the texture analysis for each patient. Another problem with this method is that it only predicts well for infants with moderate-to-severe HIE.

Murphy et al. [73] proposed a new method for DWI to detect HIE. Observers used proprietary software to identify and mark every pixel on the ADC maps that considered representative of ischaemia (cytotoxic oedema). This was the first step in diagnosing, where the 2D slices of DWIs were inspected. Automatic detection of ischaemia was carried out using features of the ADC map and DWI in a supervised learning system using random forest classification. The

first step was to create a brain mask and eliminate the background and non-brain structures. Superpixel and watershed methods were used. The first step in detecting intensity-based homogeneous superpixels was to process the ADC image. A Gaussian kernel of 0.5mm was used to calculate the square sum of the image gradients in each slice's X and Y directions. The non-zero value of the lowest 10% in the gradient image was set to zero, and the watershed transform method determined the final superpixel boundary. Then nine features were calculated in the superpixel to describe the grayscale and diffusion-weighted image of the superpixel in the ADC image and the location of the superpixel in the brain. The ischaemic lesions first were manually segmented in 20 subjects and then developed an automated algorithm using the random forest classifier. Finally, this algorithm's median sensitivity and specificity were 0.72 and 0.99, respectively. The algorithm was then applied to a data set of 54 subjects and used the results of an automated method to rate each subject. It significantly correlated with an MRI score assigned by experienced clinicians ($p < 0.0001$). The problem with this research is that due to reducing voxel results in this method, some valuable information in the HIE image was lost.

In [74], a machine learning tool was developed to detect lesions and predict outcomes for neonatal HIE patients by mining multi-site clinical data. The machine learning method can identify the lesions and extract the patterns of brain injury in MRI to find the best prediction model. Accuracy was measured by comparing predicted results with actual results. Multi-site clinical data was retrospectively collected that could be used to develop MRI analysis tools to address the neonatal HIE data shortage. Clinical data included outcomes of hospital assessments and outcomes after two years. Bayley-III was used to assess developmental functions, including cognitive, language and motor. A framework was proposed to detect HIE lesions and isolate their ADC signals. A normative ADC atlas that quantified the normal range of ADC variation in space and time was developed. Voxels that trained patients to ignore normal changes in their temporal position were entered into new channels (Z_{ADC} maps) to separate HIE-related ADC changes from normal temporal ADC changes in neonatal brain development. The accuracy of using Z_{ADC} maps and machine learning to detect HIE lesions in pilot data was 69%. The limitation of this approach is that it has not been validated in clinical data sets and lacks support for practical applications.

2.2.2. Susceptibility-weighted imaging

Susceptibility-weighted imaging is a new MRI sequence [24]. Given the sensitivity of SWI to venous vessels in the brain, several studies have validated its effectiveness for the diagnosis and prognosis of HIE outcomes [69,75-76]. Several studies have described methods of automatic vein segmentation in SWI images [77-78]. It is noted that there is no public available SWI data. This section introduces some studies for automatic diagnosis of HIE using SWI images.

In [76], feature extraction and classification were used to diagnose HIE patients using SWI images. A ridge line test was used to determine the centreline of the blood vessels. This ridge detection method is based on the retinal vessel segmentation method. First, a local median and Gaussian filters were used to remove noise. The zero crossing of the first derivative of the image was taken in the direction perpendicular to the ridge tangent. The derivative of the image could be calculated using the derivative of the Gaussian kernel. The ridge point was detected by setting the pixel corresponding to the local minimum value (valley ridge) to -1 and the local maximum value (ridge) to +1; otherwise, it was set to 0. However, the vessels corresponding to the ridges were the targets. So the Hessian matrix helped a lot. The absolute maximum eigenvalue of the Hessian matrix corresponding to each pixel on the ridge was calculated, and the corresponding pixel value in the ridge graph was set as the absolute maximum eigenvalue. The pixel value of the image point at the non-ridge position was zero. The direction of the eigenvector at the ridge point was perpendicular to the tangent of the ridge line. As a result, it got a hard ridge image and a soft ridge image for each SWI slice of every patient. Then, the histogram of soft ridges of all infant slices was calculated. These histograms ignored pixels with no grey that corresponded to the background and were unimportant. A soft ridge was fitted a curve with five parameters to avoid overfitting problems. In this method, SWI images were classified by k nearest neighbours based on the parameters of this curve. The accuracy was 91.38%. In [79], an extended three-dimensional local binary pattern was developed to distinguish the oxygenated images of three-dimensional SW images of infants with HIE, thought to be related to the oxygenation levels of the blood – a possible marker of HIE. Seventeen individuals with their SWIs were considered in [79], where seven were affected by HIE. The best classification accuracy reported [79] was 89.9%. However, two issues were related to the classification inaccuracies presented in [76] and [79]. 1) The classification inaccuracies were calculated using Kitamura scores evaluated by radiologists. The data in these

studies were unbalanced. Therefore, the inaccuracies reported in these works are somewhat biased.

Ning et al. [75] compared the small cerebral veins of neonates with HIE with healthy neonates using SWI. The study showed that the deep medullary veins of the brain were more prominent in neonates after HIE. Twenty infants with HIE and 33 healthy infants in the control group were included. Each infant had two axial SWI slices taken from the top of the lateral ventricle. The centrum semiovale in the first slice and white matter from the frontal lobe and temporal-occipital junction in the second slice were selected as regions of interest (ROI) for quantitative analysis. Veins in the ROI were segmented according to signal intensity. The ratio of the segmented vein area to the ROI area was used as a marker to distinguish the HIE infant group from the normal infant group. The results showed that, when comparing the ratios in each ROI, the values in the healthy group were much lower than in the HIE group. The ratios were 0.097 ± 0.087, 0.067 ± 0.048 and 0.108 ± 0.093 in HIE group, and 0.321 ± 0.128, 0.302 ± 0.123 and 0.325±0.125 in normal group. This study had three problems: 1) The data were unbalanced. 2) Too few slices were selected per patient. 3) The ROI was not selected sufficiently.

However, we have solved this problem by randomly selecting balanced data. For each set of SWI images, we have selected more slices than in [75] and also select more brain regions for feature extraction.

2.2.3. Deep learning for hypoxic-ischaemic encephalopathy

In recent years, deep learning has been increasingly used in medical image analysis due to its better performance. Some studies have used deep learning techniques in HIE diagnosis [80-82]. However, these studies used EEG or aEEG of full-term neonates with HIE as a dataset to train deep neural networks, and no published studies have used MRI images for deep learning for diagnosing neonatal HIE.

Wang et al. [81] built an automated network model for diagnosing neonatal HIE using aEEG screen images. In this method, 606 aEEG images from 560 infants with HIE were collected, and each image included the aEEG image and the raw-EEG image; the case report was annotated. Based on the different clinical descriptions, the dataset was reconstructed into three classes with corresponding labels: 1) Dataset A, background pattern (149 images with abnormal tagging criterion and 457 images with normal); 2) Dataset B, sleep–wake cycling (324 images with sleep–wake cycling and 272 images without); 3) Dataset C, seizure (125

with seizure and 481 without); this means that three neural network models need to be trained to predict aEEG images with different labels. The data from the three classes were divided into training data and test data. Five-ford cross-validation was used for the imbalanced problem of the dataset. Every aEEG image with a corresponding label was fed into convolution neural networks as a vector, x. The input space, y, was obtained by adjusting the training model parameters. A pre-trained model of the neural network based on the ImageNet dataset was used to obtain parameters and then fine-tuned on the annotated aEEG images. After the care process, the fine-tuned network model was introduced into a multi-input structure and maximum feature combination operator to reduce interference in aEEG images. The visual geometry group, the inception network and the residual network were treated as pre-trained models. The training data sets with different labels were separately put into pre-trained network models. The result demonstrated that the maximum feature combination operator network model performed well in all three datasets. The accuracy of datasets A, B and C was 96.69%, 97.52% and 88.42%. But the time-consuming annotation of each aEEG image was a limitation of this method, and without the aid of MRI, the reliability of aEGG images is not good enough for diagnosis.

2.3. Summary

Prognosis of HIE have improved with the development of computer technology, neurology and scoring systems. The combined use of multimodal MRI images allows for more accurate assessment of HIE lesions and improves clinical prognosis performance. Initially, studies on neonatal HIE only assessed the clinical manifestations; however, with the development of computer technology, imaging results related to HIE was also assessed. As MRI has become standard in detecting HIE, researchers have tried to determine its predictive value, the relationship of MRI results to HIE and the long-term neurodevelopmental outcomes after hypothermia treatment. The scoring systems based on MRI images have also shifted from assessing basal ganglia or thalamus to evaluating more brain tissue and haemorrhages. Early diagnoses of cerebral ischaemia and hypoxia are difficult due to the poor neurological development of newborns. Cerebral haemodynamic disorder is one of the main mechanisms of damage in HIE. Changes in venous blood flow and blood products sensitive to SWI have been shown to have significant prognostic value. SWI images may eventually become a useful tool for predicting mid- to long-term neurological outcomes in patients with HIE.

Chapter 3

Dataset

3.1. Ethics

Ethical approval for this study and use of anonymized routinely collected clinical outcome data was obtained from the Health Research Authority (HRA), Health and Care Research Wales, (HCRW) (IRAS ID 279072; REC reference 20/HRA/0260) and the National Research Ethics Service (NRES) London, City & East (IRAS ID 143392; REC reference 13/LO/1948).

3.2. Patients

Forty-two infants with neonatal HIE born at gestational age >36+6 weeks are included in this research. After undergoing hypothermia treatment, all newborns have MRI in the neonatal period as part of their clinical care. The participants in this study were scanned at a mean age of 7.8 days (min 1 day max 34 days) after birth. The British Association for Perinatal Medicine and the National Institute for Health and Clinical Excellence (NICE) currently recommend that children who survive neonatal HIE have a formal neurodevelopmental assessment at around 2 years of age. All patients in this study were from the same hospital and were followed up by the same team of specialists.

In the context of a follow-up programme all infants had neurodevelopmental assessments at age 24 months including standardised neurological examination and assessment of cognitive, motor, and language development with the Bayley Scales of Infant and Toddler Development 3, as shown in Figure 3.1. Children are assessed at the follow-up clinic using a standardised protocol that is carried out by a paediatric neurologist or neonatologist with experience in neurological and developmental assessment, together with a physiotherapist. Of the 58 surviving neonates initially included, 16 were lost to follow-up due to moving or death, and 42 were followed up at the age of 2 years.

In this study, hypoxic regions in SWI images have not been annotated by clinical experts. However, we are demonstrating that with our methods, in SWIs we can measure some attributes associated with brain regions to show how much these brain regions have been affected by hypoxic-ischaemic.

3.3. MRI Protocol

MRI was performed on a 1.5 T Siemens Symphony MRI scanner, and included proton density (PD), T1-weighted, T2-weighted, turbo inversion recovery (IR), DWI, and SWI. SWI data was acquired using a, flow-compensated, spoiled gradient echo (FLASH) sequence, with the following pulse sequence parameters: TR/TE/flip angle = 50 ms / 40 ms / 12°, voxel size = 0.9 × 0.9 × 2 mm3, bandwidth = 70 Hz/pixels. Infants were scanned whilst in natural sleep or under general anesthetic.

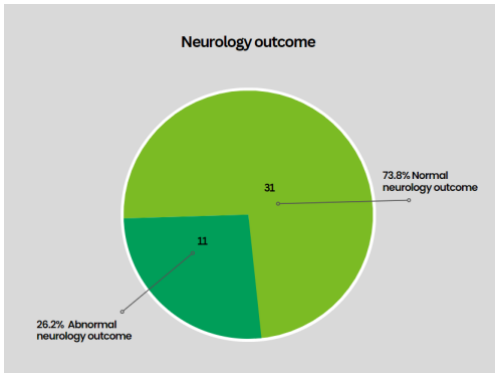
According to a report from Southampton Hospital, SWI scores were more highly correlated with clinical outcomes than Barkovich scores [41]. The NHS does not currently have a standard for SWI in HIE applications. There are approximately 40-60 SWI slices for each newborn with HIE. We automatically sorted the slices for each group of patients and selected seven SWI images from the central slice in both directions, thus avoiding the interference of images containing noisy information (i.e. nose and eyes).

3.4. Outcomes

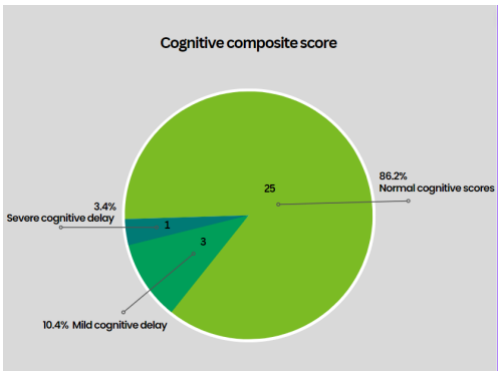
The neurological examination consisted of the assessment of cranial nerve function, movements, posture, reflexes, and muscle tone; and neurological status was then considered as either normal or abnormal (Cerebral Palsy). Of the 42 infants neurologically assessed at the

age of 2 years, 31 (73.8%) have a normal neurological outcome, and 11 (26.2%) have an abnormal neurological outcome.

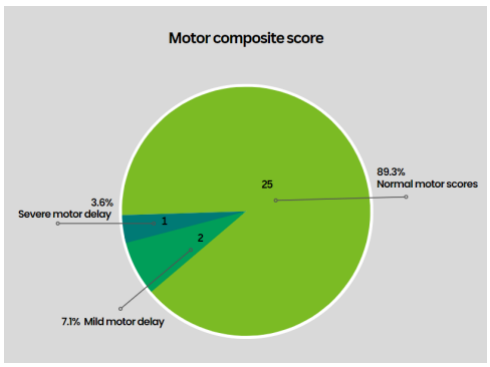
Bayley-3 is a standardised tool that assesses developmental function in infants aged between 1 and 42 months [22]. It includes three sections: cognitive, language and motor; scaled scores and composite scores can be calculated from the raw scores. For this study, composite scores were used. Bayley-3 composite scores have a mean of 100 and a standard deviation (SD) of 15. Development is considered age appropriate if Bayley-3 composite scores are less than one SD of the mean (>85). Mild delay is graded according to a composite score greater than 1–1.5 SD below the mean (77.5–85), and a moderate or severe delay is graded if the score is more than 1.5 SD below the mean (<77.5). In our study, we focus on cognitive and motor development as indicated by Bayley-3 composite scores. Outcome data on neurological status were available for 42 children. Since some of the children were unable to complete the Bayley Scales due to neurological impairment (n=11) or compliance with testing, Bayley-3 outcome data were available for 29 children for the cognitive scales, and for 28 children for the motor scales.



a.



b.



c.

Figure 3.1 a. Count of patients with neurology outcome. b. Count of patients with motor composite score. a. Count of patients with cognitive composite score.

Chapter 4

Ridge Detection and Analysis of Susceptibility Weighted Imaging in Neonatal Hypoxic Ischaemic Encephalopathy

In the previous chapters, we have described the advanced performance of susceptibility-weighted imaging (SWI) in the detection of neonatal hypoxic-ischaemic encephalopathy (HIE). However, there is no research investigating the relationship between the SWI images of infants with HIE at birth and the long-term neurodevelopmental outcomes after neonatal HIE. In this chapter, we develop a new automated system to evaluate venous vascular structure as prognostic indicators of developmental outcomes for infants with HIE by detecting and analysing ridges representing vessels on SWI images.

4.1. Introduction

As described in previous chapters, SWI images are very sensitive to the detection of vascular extraneous blood products and hypoxic-ischaemic veins [67]. The appearance of anoxic vessels is different from the healthy vessels in that the anoxic vessels look darker than the healthy ones [76]. Obviously, blood vessel analysis on the SWI will be useful for diagnosis of neonatal HIE. Automated blood vessel segmentation is an important topic in medical image analysis [86-87,

109]. For various modalities of medical images, there are many methods of vessel segmentation using computer-acid. In [124] and [126], the most recent and innovative blood vessel segmentation algorithms were reviewed, including vessel enhancement, deformable models, tracking and machine learning. As our study was analysing SWI images of neonates with HIE, the researchers' attention focused on the field of vessel segmentation on SWI images.

Silvain et al. [77] proposed an automated method segmenting whole-brain venous blood in SWI by conditional random fields. The conditional random field model with multiple first- and second-order potentials (appearance, shape, location, smoothness potential and data-dependent edge potential) was built to segment cerebral veins on the raw SWI data. However, this method applied the SWI data consisting of 5 echoes. For SWI images acquired from different echoes, the random conditional field model needs to be explored. Ward et al. [78] generated a composite vein image by combining SWI images, quantitative susceptibility maps (QSM) and a vein atlas to improve the accuracy of automated vein segmentation. The vein atlas was created by manual vein tracing with authors. Three different automated vein segmentation methods were employed to generate SWI images, QSM images and composite vein images separately to evaluate the accuracy of segmenting the veins: 1) a Hessian-based vesselness filter followed by an Otsu threshold; 2) a statistical segmentation based on the Markov random field; 3) a ridge-based filter segmentation. The composite vein images showed improved performance for segmenting the veins, and the ridge-based segmentation method presented a good accuracy on SWI images. But this method relied heavily on the combination of SWI and QSM images. Sina et al. [127] proposed a new automated vein segmentation algorithm which combines vessel enhancement filter and local threshold methods with shearlet transform. QSM data and SWI images were processed and input to shearlet scale space. Then by local thresholding, local vein segmentations were computed. Frangi filter and a recursive vessel filter vein segmentation methods were used for comparison. The error metrics in this proposed algorithm was better than other two vein segmentation methods. Xiao et al. [109] suggested a way to segment small veins in SWI images using a deep learning network. The deep network contains 19 convolutional modules. Dense connectivity was used to enhance the network performance. Segmentation results are obtained by dice coefficient, 0.756±0.043. The majority of vein segmentation methods for SWI images aim to obtain the global appearance of the veins. In our study, we expected to obtain more vein features (e.g. width and depth) by using the ridges of the veins.

The approach of vessel segmentation by determining the ridge of vessels has been used different medical tasks. The authors of [123] compared four methods applicable to the extraction of linear structures from mammograms including a method for ridge detection. The authors of [125] used an iterative region-growing method to acquire the centreline pixels of the blood vessels. This method outperformed on the retinal blood vessel dataset to other method [128]. Most vessel segmentation studies extract the vessel features and classify the medical images. However they ignore the properties of the vessel features. Our aim is not only to classify SWI images using the vessel features, but also to further analyze the relationship between vessel features and poor outcomes.

We use a ridge detection method for vessel segmentation on a balanced SWI dataset and employ the neurological outcomes of patients at the age of 24 months as the outcome measure. We present an automated HIE prognosis system by using the following four feature vectors: histograms of the width, intensity, length, and the largest eigenvalue of Hessian of vessels detected as ridges to predict neurological outcomes. Finally, we use support vector regression (SVR) to predict the cognitive and motor scores of infants assessed by Bayley-III Scales at age 24 months from four feature vectors.

4.2. Methods

4.2.1. Image pre-processing

Before processing SWI images, a brain mask is extracted to remove the background signal from the images to precisely identify vessels in the brain slices. We use an active contour model [83] for each SWI slice to obtain a binary brain mask that eliminates noise and background from the calvarium, as shown in Figure 4.2(b).

The active contour is expressed as an energy minimization process. The target feature is a minimum of a suitably formulated energy functional. This energy functional includes more than just edge information: it includes properties that control the way the contour can stretch and curve. The minimum value of the energy function composed of the addition of the contour's internal energy (E_{int}), the constraint energy (E_{con}) and the image energy (E_{image}) is used to obtain the target feature. Representing the position of a snake parametrically by $v\,(\mathrm{s}) = \big(x(s), y(s)\big)$, we can write energy function as

$$E_{snake} = \int_{s=0}^{1} E_{int}\big(v(s)\big) + E_{image}\big(v(s)\big) + E_{con}\big(v(s)\big)\, ds$$

In SWI images with neonatal HIE, this works by minimising the energy that is in part defined by the image and in part by the spline's shape: length and smoothness. The minimization is done implicitly in the shape energy and explicitly in the image energy. E_{int} controls the natural behaviour of the snake and hence the arrangement of the snake points, the image energy; E_{image} attracts the snake to chosen low-level features such as edge points; E_{con} allows higher-level information to control the snake's evolution.

We initialise a circle around the brain and use the default boundary condition to fit a closed curve. The closed curve is then transformed into a binary mask to remove the background of the brain. Figure 4.2(a) shows the original SWI of HIE and Figure 4.2(b) shows the result after using the active contour model.

4.2.2. Ridge detection

There are multiple methods for ridge detection, e.g. vessels tracking [129-130], model-based methods [132], morphology-based techniques [133], and image filtering technique [131]. In our opinion, as the vessel structure is shown clearly enough on the SWI image, we directly compute the zero-crossing point of the two-dimensional grey SWI images as the ridge line point. We have also used morphology-based [133] and method in [76] techniques to compare our approach. As shown in Figure 4.3, we can clearly observe that the morphology-based method does not detect all the vessels, and the method in [76] contains a lot of noise. However, our method extracted the fullest ridges from the SWI images. Now convolutional neural networks (CNN) method is also used to detect ridges [134]. However, lacking a manually annotated dataset and sufficient SWI images, we are unable to use deep learning methods for ridge detection.

In SWI images, blood vessels appear to be ridge-like objects [87]. Zero-crossings of the image derivatives are therefore used to detect the vessels represented by ridges in two-dimensional SWI slices. Generally, using a convolution operator, image $I(x,y;\sigma)$ can be generated by convolving the image $I(x,y)$ with a Gaussian kernel $G(x,y;\sigma)$ of variance σ, which could also reduce noise existing in SWI images.

$$I(x, y; \sigma) = I(x, y) * G(x, y; \sigma)$$

We denote *I(x,y)* to be the SWI image in which we view the brain using the active contour model in section 4.2.1. The change in the gradient of intensity near a pixel is ascertained by the first derivative of the image *I(x,y;σ)*. A ridge point is located where the intensity gradient vanishes. Such a point corresponds to the zero-crossing point. Finally, a connected pixel chain with one-pixel width is obtained for every ridge in the SWI images. In fact, the noise will affect the ridge detection to some extent. We set a threshold of image greyscale value to remove the influence of noise. Figure 4.2(b) shows the result.

4.2.3. Hough transform for centre line removal

The Hough transform (HT) [84] implementation defines a mapping from the image points into an accumulator space (Hough space). The mapping is achieved in a computationally efficient manner, based on the function that describes the target shape. HT counts the potential solutions in an accumulator array that stores the evidence or votes. This strategy is robust and able to handle the noise and occlusion. The point where this line intersects the line in an image is given by:

$$\rho = x\cos(\theta) + y\sin(\theta)$$

where θ is the angle of the line normal to the line in an image and ρ is the length between the origin and the point where the line intersects. Figure 4.2(c) shows that there is a centre line in the brain. Figure 4.2(d) shows the removed line via the Hough transform.

4.2.4. Ridge segmentation

Vessels form ridges in SWI images. To further analyse our SWI images, we need to initially detect and then segment each individual ridge. This segmentation later helps to extract features from each ridge individually. After ridge detection, all of the initial SWI images for infants with HIE are transformed into binary images where 1s (white pixels) represent ridges. A component labelling scheme for connected (ridge) pixels is used here for ridge segmentation. This algorithm is applied to a ridge map (a binary image) to assign a white pixel (ridge pixel) to an individual connected region belonging to a ridge with a certain label. Here, a local 3 x 3 neighbouring window is used to visit each pixel in the ridge map to check eight connectivities surrounding each ridge pixel.

1	1	1
1	10	1
1	1	1

(a)

0	0	0
0	1	0
0	0	1

(b)

Figure 4. 1 (a) Bifurcation point template, where convolution with the ridge point is equal to or greater than 13; (b) End point template

However, before applying this process, it is vital to determine pixels representing the termination and bifurcation points as a part of the ridge segmentation. If a central point is a ridge pixel and there is only one neighbour, it becomes a termination point. A bifurcation point has at least two neighbours. The termination and bifurcation points are detected by looking for maxima in an image computed by convolving a ridge image with the template as shown in Figure 4.1(a) and searching for values of unity in an image obtained by convolving a ridge image with all rotated variants of the termination template as depicted in Figure 4.1(b). Two termination points in each connected ridge are stored in a list representing a segmented ridge. Bifurcation points are eliminated to separate crossing vessels. As shown in Figure 4.2(f), each blood vessel (ridge) in the binary ridge map is segmented and for illustration purposes is depicted in a certain colour. Additionally, 10 raw SWI slices belonging to the same neonate with HIE are selected and displayed as superimposed on top of each other. For entire set of slices, the minimum value of the pixels passed along the slice direction is taken as the pixel value in the projection plane. We perform a projection reconstruction of the 10 SWI slices to obtain a low-density image, minimum intensity projection (MinIP). Figures 4.2(f-g) show the SWI MinIP image along the slice direction and the projection reconstruction of ridge segmentation corresponding 10 SWI slices. It can be observed that the SWI MinIP image demonstrates a better continuity of the veins' paths and extends visualization of small veins.

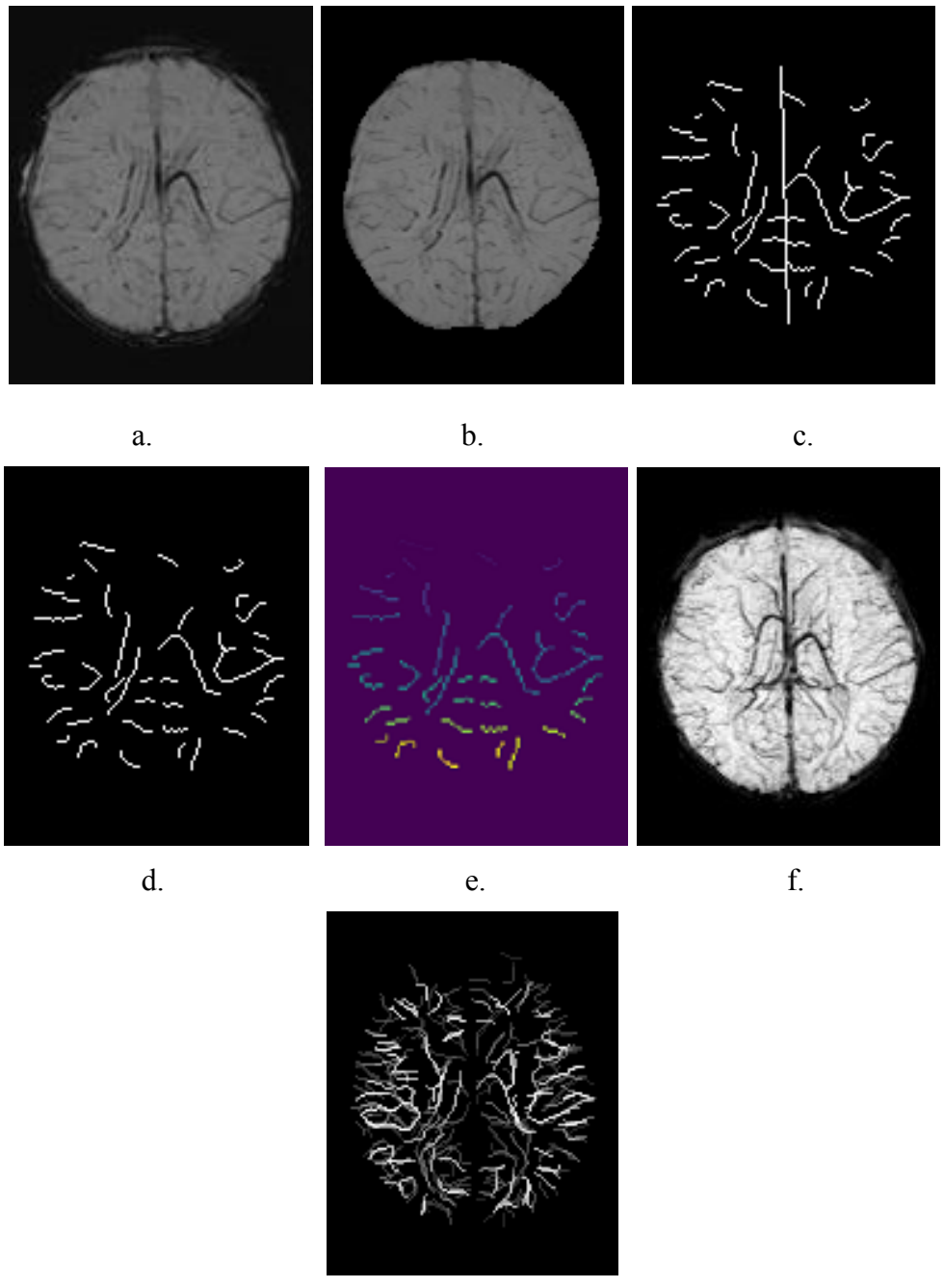
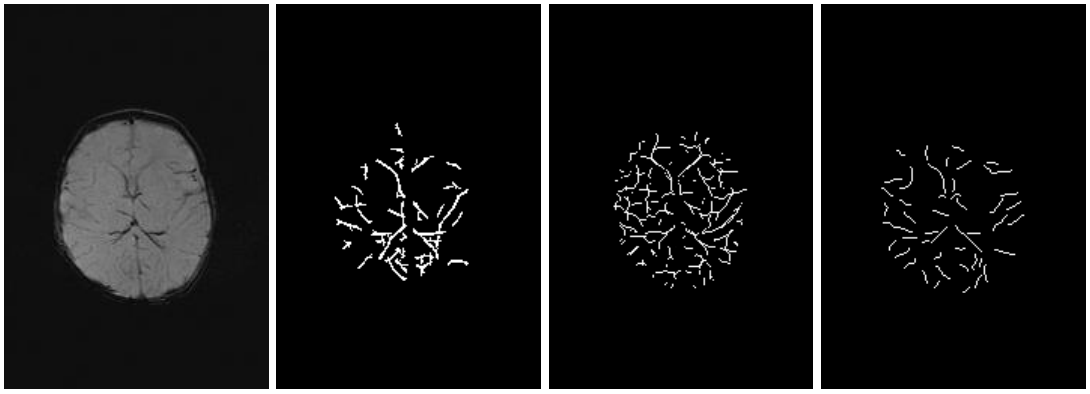


Figure 4. 2 a) Axial SWI image for a 8-day-old term neonate with HIE and category 3 in Kitamura scoring system, normal neurology outcome at age of 2 years; b) The result from the active contour model; c) Ridge detection with the grey value threshold = 40, and the threshold of difference between the zero-crossing point and closest point is 270; d) The centreline of the brain removed with pixel = 2; e) Labelled veins are coloured,; f) Axial SWI MinIP image by 10 slcies; g) Ridge segmentation corresponding to the SWI MinIP image.



a. b. c. d.

Figure 4.3 a) Axial SWI image; b) Results of morphological operator; c) Result of the method in [76]; d) Result of our methods

4.2.5. Feature extraction

As we known, HIE will make the blood vessels in the brain of infants briefly lack blood. When the blood is reoerfused, the vessels will be in a state of expansion. There will be changes in the width of the blood vessels between the infants with HIE and normal infants. Meanwhile, blood products from the blood vessels going out will also be displayed in the SWI images. The appearance of hypoxic-ischaemic vessels is different from the healthy vessels in that the hypoxic-ischaemic vessels look darker than the healthy ones. In the course of HIE, the blood vessel length of normal infants and abnormal infants also changes. Therefore, the intensity and length of the vessels will be changed in the brain of infants with HIE. The eigenvalues of the Hessian matrix corresponding to the pixels of the ridges also become one of the prognostic indicators for infants with HIE. Four features, as described below, are extracted from the segmented vessels for vessel classification:

1) Vessel Width Measurements

Vessel width is calculated by measuring the distance between points on the vessel boundaries perpendicular to the ridge orientation as shown in Figure 4.3(a). For each labelled (segmented) ridge (blood vessel), a Canny edge detector is also applied to find the blood vessel edges. Then, the boundary points on the Canny edges are assigned to two sets on both sides of the ridge, as shown in Figure 4.3(a). Some junctions are derived from both sides of the ridge along a line normal to BC in Figure 4.3(a) by connecting three consecutive points on the ridge at point A, as shown in Figure 4.3(a). We can obtain the coordinates of ridge point A and the corresponding two edge points (x_1, y_1) and (x_2, y_2). The vessel width is defined as the shortest Euclidean

distance between these two points from boundaries in both sides of the ridge. Finally, the histogram of all vessel widths is calculated as a feature vector known here as *Width* .

2) Vessel Intensity Measurements

Depending on the level of blood oxygenation, SWI images for neonatal HIE will demonstrate veins with varying signal intensities. We utilise the difference between the intensity values of ridge points and those of boundary points of vessels as a marker of hemodynamic or vascular changes in the brain due to HIE. By tracking the ridge pixel coordinates of the vessel and the corresponding edge point pairs of the vessel boundary in the original SWI images, the maximum pixel value difference between the ridge point and the two shortest edge points on either side of the ridge point (Figure 4.4(a)) is measured and known here as vessel intensity as a feature for HIE outcome prognosis. Similar to the Width feature, the histogram of all vessel intensities is computed as a feature vector referred to as *Intensity* in this thesis. Figure 4.4(b-c) shows the vessel of an infant with HIE and the value of vessel pixels.

3) Vessel Length Measurements

We also measure the vessel length as a feature for HIE outcome prognosis. For each segmented ridge, we count the number of pixels from start point to end point as vessel length [107]. Ultimately, a histogram of all vessel lengths within an SWI dataset constitutes a feature vector known as *Length* here.

4) Ridge Eigenvalue

In SWI images, a ridge point is the local minimum point in the direction of the largest gradient change. The local intensity of the pixel on the ridge is derived from the Hessian matrix consisting of the second-order partial derivative of the SWI image. After calculating a 2 × 2 Hessian matrix for every pixel on ridges in the two-dimensional SW image, for each ridge point, two eigenvalues (λ) of the Hessian matrix are computed, where the eigenvalue with the maximal absolute value is treated as a feature for HIE outcome prognosis.

$$H = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix}$$

where $I_{xy} = \frac{\partial^2 I}{\partial_x \partial_y}$. A histogram of ridge eigenvalues is also considered as a feature vector known as *Eigenvalue* here.

(a) (b) (c)

Figure 4. 4 (a) The pixel pairs according to perpendicular line; (b) Greyscale SWI; (c) The value in a grey-scale SW image.

4.2.6. Error of vessel segmentation

In some cases, real vessels may not be precisely the same as the ridges detected by our algorithm. The performance of the ridge detection method employed here is measured by computing the error in the vessels segmented by our method from the manually segmented vessels (ground truth). Five SWI slices, with a size of 290 × 256 pixels, covering different brain anatomy areas, from five infants are manually annotated as our ground truth for segmented vessels. Manual ridge segmentation is performed by experts to provide a ground truth for error measurements for our ridge detection algorithm. For time-consuming and other reasons, only five SWI images were annotated and then validated by clinicians. This sets up five mask images where pixels on vessels are set to one, and non-vessel pixels are considered to be zero. We manually segment 2,455 pixels as vessels in these five slices. Next, ridges of the same five slices are segmented by the segmentation algorithm used here. Figure 4.4 (a-c) shows segmentation of our method on SWI image with HIE and compare the ridges with manual segmentation. In addition, the SWI MinIP image created by 10 raw SWI slices with neonatal HIE and corresponding segmentation are shown in Figure 4.4 (d-f). We compute error pixels n as the difference in the number of pixels between automatically segmented ridges and manually segmented ridges of a vessel. The error is normalised as:

$$p = \frac{\varepsilon}{m}$$

where m is the total number of ground truth pixels for the vessel. There are 187 error pixels in the five SWI images; these are pixels present in automated vessel segmentation but absent in the manual image or vice versa. There is good consistency between the manual annotators ($p < 0.05$). The error accuracy in each of the five images is calculated using error pixels. We calculate the mean and standard deviation of the error as $p = 7.75\% \pm 1.97\%$ for our vessel segmentation algorithm.

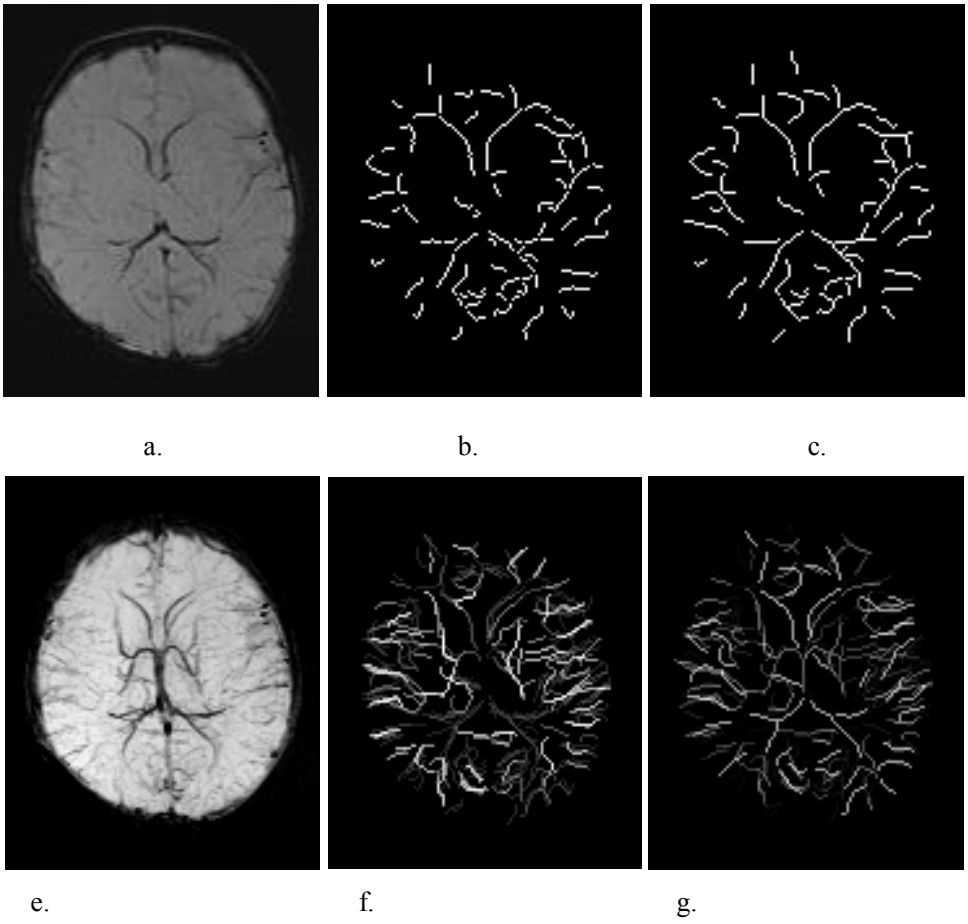


Figure 4. 3 a) Axial SWI image for a 15-day-old term newborn with HIE and category 3 in Kitamura scoring system, normal neurology outcome at age of 2 years; b) Ridge segmentation for SWI image using our method; c) Manual segmentation. d) Axial SWI MinIP image through 10 slices; e) Automatic ridge segmentation corresponding to the SWI MinIP image; f) Manual segmentation corresponding to the SWI MinIP image.

4.3. Classification for neurological outcomes

4.3.1. Classification algorithms

1) *k*-nearest neighbors (*kNN*)

kNN is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. As we used Euclidean distance defined as:

$$D = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2}$$

where x, y are the coordinates of image pixels and N is number of pixels in the image.

2) Random forest classification

Random forest [85] is a supervised learning algorithm. It can be used both for classification and regression. A forest is comprised of trees. Random forest creates decision trees on randomly selected data samples to obtain prediction from each tree and to select the best solution by means of voting. A random forest can fit a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The sub-sample size is always the same as the original input sample size.

4.3.2. Results

Having calculated the histogram of the width, intensity, length and Hessian eigenvalue of the vessels in the SWI images, 42 infants with information on neurology outcomes are classified into two groups: 31 with normal neurological outcomes and 11 with abnormal neurological outcomes (cerebral palsy). The SWI dataset and the approval are introduced in Chapter 3. To balance the data distribution, 11 out of the 31 patients with normal neurology outcomes are randomly selected for the classification process. This selection is performed 10 times in each experiment so that classification accuracies are reported as mean and standard deviation of the classification accuracies over these 10 experiments. The selection of data can solve the problem of unbalanced data used for classification in [76] and [79]. Histograms of the four aforementioned features are normalised to the area below each histogram. These normalised histograms merge into a single feature vector by concatenating feature vectors and are then fed into a *kNN* and a random forest classifier to classify the SWI images into normal or abnormal groups. Meanwhile, we apply an unnormalized histogram to classify the SWI images.

Table 4. 1 Classification accuracy with unnormalized histograms

Classifier	Unnormalized Features	Accuracy Result
kNN	Width	70.45±5.39%
	Intensity	72.23±4.03%
	Length	73.39±3.11%
	Eigenvalue	70.45±3.21%
	Width + Intensity + Length	69.86±4.79%
	Width + Intensity + Length + Eigenvalue	69.72±3.98%
Random Forest	Width	70.17±4.51%
	Intensity	72.37±3.03%
	Length	73.93±5.79%
	Eigenvalue	71.91±7.43%
	Width + Intensity + Length	76.44±5.79%
	Width + Intensity + Length + Eigenvalue	76.72±5.43%

Table 4. 2 Classification accuracy with normalized histogram

Classifier	Normalized Features	Accuracy Result
kNN	$\frac{Width}{\lvert Width\rvert}$	69.69±4.23%
	$\frac{Intensity}{\lvert Intensity\rvert}$	71.37±4.31%
	$\frac{Length}{\lvert Length\rvert}$	72.93±3.83%
	$\frac{Eigenvalue}{\lvert Eigenvalue\rvert}$	72.27±4.85%
	$\frac{Width}{\lvert Width\rvert}+\frac{Intensity}{\lvert Intensity\rvert}+\frac{Length}{\lvert Length\rvert}$	70.86±3.03%
	$\frac{Width}{\lvert Width\rvert}+\frac{Intensity}{\lvert Intensity\rvert}+\frac{Length}{\lvert Length\rvert}+\frac{Eigenvalue}{\lvert Eigenvalue\rvert}$	75.45±5.81%
Random Forest	$\frac{Width}{\lvert Width\rvert}$	71.17±7.45%
	$\frac{Intensity}{\lvert Intensity\rvert}$	72.83±5.67%
	$\frac{Length}{\lvert Length\rvert}$	74.33±8.68%
	$\frac{Eigenvalue}{\lvert Eigenvalue\rvert}$	71.17±6.38%
	$\frac{Width}{\lvert Width\rvert}+\frac{Intensity}{\lvert Intensity\rvert}+\frac{Length}{\lvert Length\rvert}$	78.33±4.43%
	$\frac{Width}{\lvert Width\rvert}+\frac{Intensity}{\lvert Intensity\rvert}+\frac{Length}{\lvert Length\rvert}+\frac{Eigenvalue}{\lvert Eigenvalue\rvert}$	78.67±2.58%

Table 4. 3 The Pearson's Correlation Coefficient between the features

	Width	Intensity	Length	Eigenvalue
Width	1	0.4831	0.1667	-0.1714
Intensity	0.4831	1	0.4789	-0.3122
Length	0.1667	0.4789	1	-0.1973
Eigenvalue	-0.1714	-0.3122	-0.1973	1

To assess the difference between the two features, a paired t-test was used. A p-value of less than 0.05 was considered statistically significant. In Table 4.4, the classification results of concatenated features from the random forest are significantly better than *eigenvalue* ($p < 0.003$) for normalized histogram. In the pairs of *width* and concatenated features, intensity and concatenated features (*width, length*, intensity, and *eigenvalue*), the random forest classification results show significant difference ($p < 0.05$).

Table 4. 4 The *p*-value between the different features

Pairs	Unnormalized histogram		Normalized histogram	
	Knn	RF	*Knn*	RF
Width vs Intensity	0.413	0.233	0.462	0.677
Width vs Length	0.181	0.174	0.121	0.402
Width vs Eigenvalue	0.992	0.679	0.272	0.891
Width vs WIL	0.817	**0.017**	0.584	**0.036**
Width vs WILE	0.781	**0.016**	**0.021**	**0.013**
Intensity vs Length	0.622	0.561	0.438	0.597
Intensity vs Eigenvalue	0.212	0.733	0.704	0.518
Intensity vs WIL	0.273	0.069	0.735	0.06
Intensity vs WILE	0.248	0.063	**0.077**	**0.018**
Length vs Eigenvalue	0.076	0.493	0.725	0.291
Length vs WIL	0.154	0.307	0.163	0.321
Length vs WILE	0.084	0.266	0.175	0.218
Eigenvalue vs WIL	0.778	0.123	0.426	**0.0086**
Eigenvalue vs WILE	0.726	0.109	0.141	**0.0009**
WIL vs WILE	0.998	0.885	**0.023**	0.827

WIL: Width + Intensity + Length. WILE: Width + Intensity + Length + Eigenvalue

The classification accuracies of unnormalized and normalized histograms are presented in Table 4.1 and Table 4.2, respectively. In Table 4.1, accuracy results with a single feature vector

show higher performance than the concatenated features. In both unnormalized and normalized histograms, the random forest classification performs better than the *kNN* classification while using the *length* and concatenated features. There are significant differences between two classifier with unnormalized histogram ($p < 0.05$). No significant differences were provided between the classification results for the normalized histogram. We can observe that unnormalized width, intensity, length and eigenvalue features get higher accuracy than the same normalized features when a single feature is used for *kNN* classification. As tabulated in Table 4.2, the highest accuracy 78.67% ± 2.58% is achieved by the random forest classifier with four feature vectors (width, intensity, length and Hessian eigenvalue). However, by using only length as a feature vector, a slightly better classification accuracy than the Hessian eigenvalue feature proposed in [76] is achieved. An interesting finding is that for unnormalized histogram, the classification performance decreases by using concatenated features. This is because when unnormalised histograms of multiple features are combined together, some redundant information is included. It can affect the classification results. In Table 4.1 and Table 4.2, the classification strategy for both classifiers is a leave-one-out cross-validation with balanced data (11 patients with normal neurology outcome and 11 patients with abnormal neurology outcome).

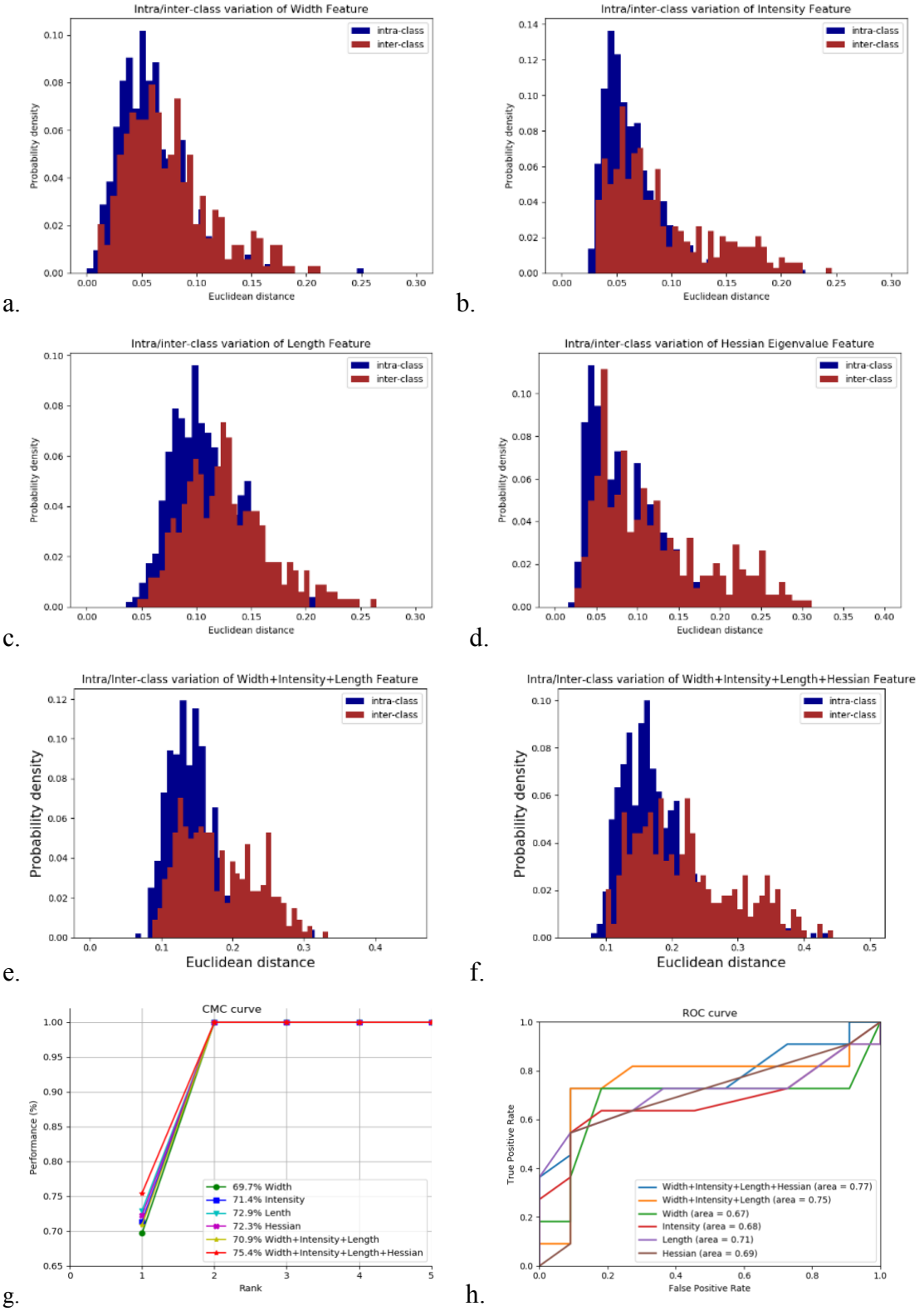


Figure 4.4 Intra/inter-class variations for (a) Width, (b) Intensity, (c) Length, (d) Eigenvalue, (e) Width+Intensity+Length, (f) Width+Intensity+Length+Eigenvalue, (g) CMC associated with the kNN classifier, (h) ROC curve of our method with the random forest classifier

Figure 4.5 (a to f) illustrates the inter- and intra-class variation histograms for various combinations of the four aforementioned features for both normal and abnormal groups in our dataset. The blue bars in Figure 4.5 depict the intra-class variations for feature vectors in the same group, and the brown bars represent inter-class variations for feature vectors from different groups. As shown in Figure 4.5(f), the overlap between the two histograms is less than the overlaps of the histograms in Figure 4.5(a to d). This is consistent with the results presented in Table 4.2, indicating that the classification accuracy of combined features is higher than that of the single features. Figure 4.5(g) also shows the cumulative match characteristic (CMC) curve for the *kNN* classifier. As shown in figure 4.5(g), the highest accuracy is a concatenated vector of the four features, with 75.4% in rank 1. It is observed that all of the *kNN* classifiers achieve 100% accuracy in rank 2. Figure 4.5(h) displays the receiver operating characteristic (ROC) curves for the various features and their combinations when the *kNN* classifier is used for classification. As shown in figure 4.5(h) for each feature or its combination, the area under the ROC curve is also displayed. Figure 4.5(g and h) show the combination of four features mentioned above, i.e. width, intensity, length and Hessian feature vectors (blue curves) have performed best for this classification task. We also calculate Pearson's correlation coefficient r for all possible pairs of feature vectors, as demonstrated in Table 4.3. The observation in Table 4.3 is that some features exhibit a certain dependency. This explains why the results of *kNN* classification for the combination of width, intensity and length are less accurate than the accuracy obtained using only one of the features. To increase the classification accuracy, we apply the random forest revealing a better performance as presented in Table 4.2.

4.4. Regression analysis for motor and cognitive development

HIE is associated with a high risk of cognitive and motor impairments in children [17]. Such impairments are evaluated by using the clinical outcomes of patients through a scoring system two years after birth. It is therefore important to predict such clinical outcomes immediately after the birth. In our database, motor and cognitive development of infants with neonatal HIE are assessed by Bayley Scales at age 24 months, which results in a scaled score [see Chapter 3]. The scores can either be used as a continuous variable or one for which cut-offs of 'normal' or 'delayed' can be applied. In normal populations, both motor and cognitive scores have a

mean of 100 and a standard deviation (SD) of 15. A patient is considered to have a normal score if the score is within 1 standard deviation of the mean. Mild delay is a score equal or more than 1 SD but less than 1.5 SD below the mean, and moderate/severe delay is a score more than 1.5 SD below the mean. We have 29 patients with cognitive scores and 28 with motor scores. All groups include patients with both normal and abnormal neurological outcomes. There is no direct correlation between neurological outcomes and cognitive and motor scores.

For children tested with the Bayley Scales, mean Bayley composite scores are broadly average [see Table 4.5]. When considering the presence of delay, 86.2% of tested children have normal cognitive scores (>85), 10.4% mild cognitive delay (77.5–85) and 3.4% severe cognitive delay (<77.5); and 89.3% have normal motor scores (>85), 7.1% mild motor delay (77.5–85), 3.6% severe motor delay (<77.5).

Table 4. 5 Results from testing with the Bayley Scales

	All tested children	Number with Normal Scores	Number with Mild Delay	Number with Severe Delay
Cognitive composite score, mean (SD), min-max	101.2 (17.76) min 60–max 145	25	3	1
Motor composite score, mean (SD), min-max	100.7 (16.35) min 55–max 133	25	2	1

Support vector regression (SVR) and random forest regression (RFR) are applied to find a regression model to predict the cognitive and motor scores, respectively, by using the four feature vectors (*width*, *intensity*, *length* and *Hessian eigenvalue*) that we have extracted, and a leave-one-out strategy is employed to evaluate the performance of the regression methods. These four feature vectors show promising results. Similarly, we use the unnormalized and normalized histograms of four feature vectors. By training the SVR and RFR models with the four feature vectors (as input, x_n) and cognitive or motor scores (as output, y_n), the mean absolute and mean relative score errors in our score calculations for all patients in our dataset can be predicted as shown in Table 4.6 and Table 4.7. Meanwhile, we obtain the prediction score of these infants from SVR and RFR models as shown in Table 4.8 to Table 4.11.

To measure the performance of these regression models, the error for each test patient is calculated as the absolute difference between the true score for cognitive or motor scores and the corresponding predicted score provided by using trained different regression models with these ratios and the mean relative error (MRE) as defined below:

$$\text{MRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|p_i - t_i|}{t_i}$$

where n is the number of patients in our database, p_i is the regression model prediction value, t_i is the true value (ground truth). A mean and standard deviation are calculated for both absolute errors and MRE.

Table 4. 6 Regression models with unnormalized parameters

Regressions	Groups	Mean Error	Mean Relative Error
SVR	29 Patients with cognitive scores	12.34±9.45	0.116±0.09
	28 Patients with motor scores	12.37±8.67	0.116±0.10
RFR	29 Patients with cognitive scores	13.32±7.95	0.133±0.081
	28 Patients with motor scores	12.58±8.43	0.116±0.09

Table 4. 7 Regression models with normalized parameters

Regressions	Groups	Mean Error	Mean Relative Error
SVR	29 Patients with cognitive scores	11.40±13.24	0.113±0.13
	28 Patients with motor scores	10.98±7.67	0.109±0.067
RFR	29 Patients with cognitive scores	15.14±11.65	0.131±0.10
	28 Patients with motor scores	12.69±9.11	0.117±0.09

In Table 4.6 and Table 4.7, each patient is given the predicted score by using the trained SVR and RFR. Errors are calculated for each test patient as absolute differences between the true scores of cognitive or motor scores and the corresponding predicted scores, and we obtain a mean and a standard deviation for the errors. The error divided by the average value of scores across the data is also used to compute the relative value. From Table 4.7, normalized histograms show better prediction than unnormalized histograms in Table 4.6 through the mean relative error. There are significant difference in all regression cases (p <0.001). As shown in Table 4.7, the SVR regression method achieves better performance for predicting motor scores than predicting cognitive scores, i.e. the relative value for patients with motor scores is 0.109 ±0.067. However, the mean relative error for the cognitive scores is 0.113±0.13, which is

slightly higher than that for motor scores. It is expected that our regression errors decrease if we get more data with scores for clinical outcomes.

Table 4. 8 Prediction Scores from SVR with normalized parameters

	All tested children	Number with Normal Scores	Number with Mild Delay	Number with Severe Delay
Cognitive composite score, mean (SD), min-max	99.3 (4.7) min 91–max 110	29	0	0
Motor composite score, mean (SD), min-max	101.3 (7.74) min 87–max 114	28	0	0

Table 4. 9 Prediction Scores from RFR with normalized parameters

	All tested children	Number with Normal Scores	Number with Mild Delay	Number with Severe Delay
Cognitive composite score, mean (SD), min-max	103.2 (9.5) min 85–max 127	27	2 (1 correct 1 incorrect)	0
Motor composite score, mean (SD), min-max	102.7 (6.1) min 89–max 115	28	0	0

Table 4. 10 Prediction Scores from SVR with unnormalized parameters

	All tested children	Number with Normal	Number with Mild Delay	Number with Severe Delay
Cognitive composite score, mean (SD), min-max	99.4 (0.46) min 98–max 100	29	0	0

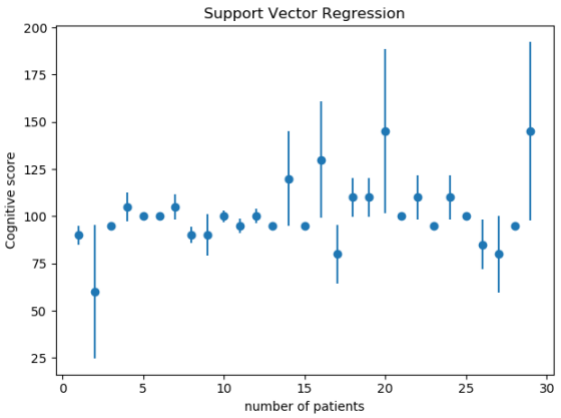
Motor composite score, mean (SD), min-max	100.7 (0.6) min 99–max 102	28	0	0

Table 4. 1 Prediction Scores from RFR with unnormalized parameters

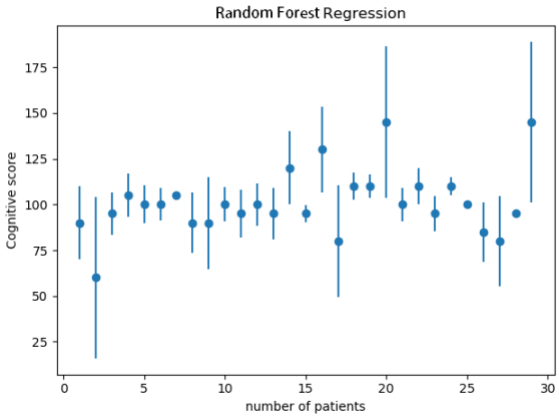
	All tested children	Number with Normal Scores	Number with Mild Delay	Number with Severe Delay
Cognitive composite score, mean (SD), min-max	102.8 (5.9) min 83–max 113	28	1 (1 incorrect)	0
Motor composite score, mean (SD), min-max	101.4 (4.9) min 85–max 109	27	1 (1 incorrect)	0

In Table 4.8 to Table 4.11, we applied normalized and unnormalized parameters to fit the regression model, respectively. The predicted scores fluctuated around the median, which avoided overfitting each patient's original score. For the overall model, with a minimum amount of overfitting shows that the regression model performs well. We divide the patients with cognitive and motor scores into three groups, ‘normal’, ‘mild delay’ and ‘severe delay’. In 29 patients with cognitive scores, there are 25 infants in the normal group, 3 infants in the mild delay group and 1 in the severe delay group. In 28 patients with motor scores, there are 25 infants in the normal group, 2 infants in the mild delay group and 1 in the severe delay group. However, in our prediction scores, almost all patients are concentrated in the normal group. We need to know which scores form the right and wrong groups and how many patients are grouped incorrectly and how many patients are grouped correctly. From the above table, the prediction results of the regression model with normalized histograms as feature vectors are better than those with unnormalized histograms. Figure 4.6 shows the errorbars between the true values and the prediction values of the regression model with normalized histogram and unnormalized histogram for cognitive scores (figure 4.6 (a, b, e and f), and for motor scores (figure 4.6(c, d, g and h).In Figure 4.6, the blue dots are the true composite scores, and the bar is an indication of the variation between the predicted value and the true value. As observed

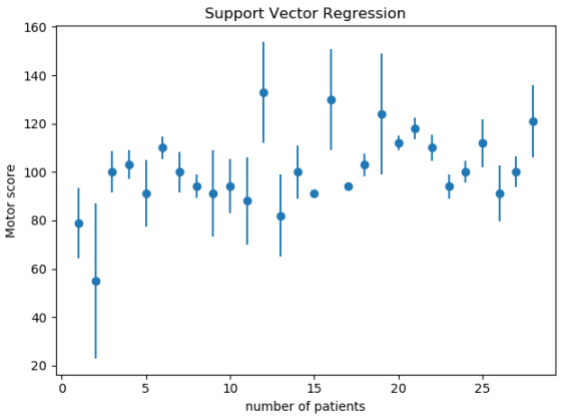
from figure 4.6 (a), by using the SVR models with normalized histogram, the differences between predicted values and the true values are lowest for cognitive scores. Also for motor scores, the lowest differences between the predicted scores and the true scores are shown in figure 4.6 (c) obtained by SVR models with normalized histogram. The SVR models for motor scores perform the best, as the error range between the predicted scores and true scores wavered in a small range.



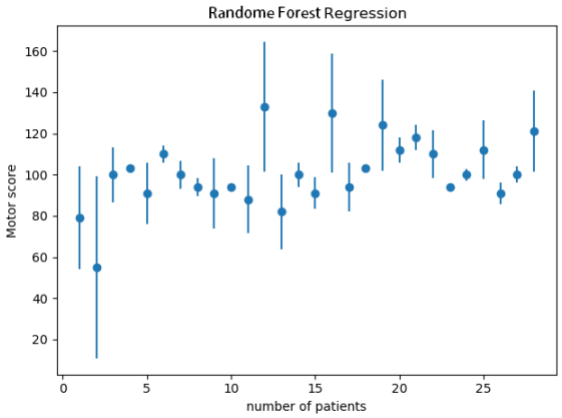
a.



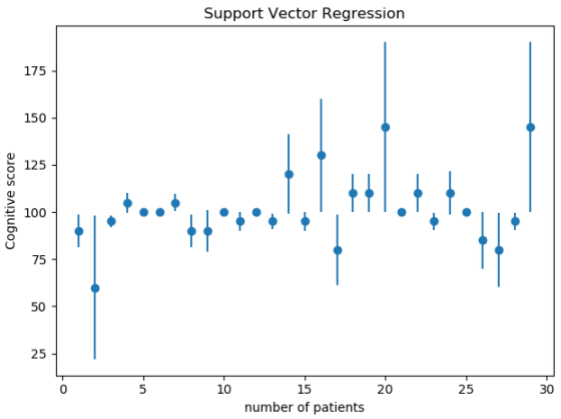
b.

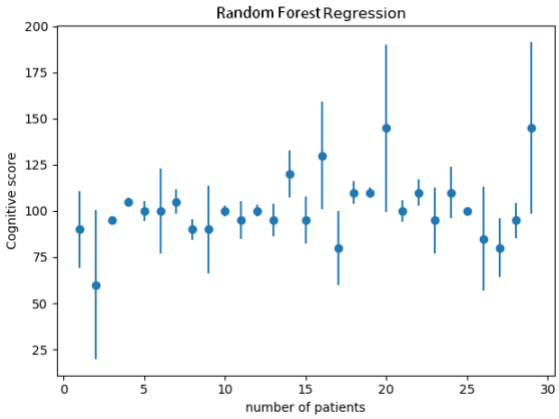


c.

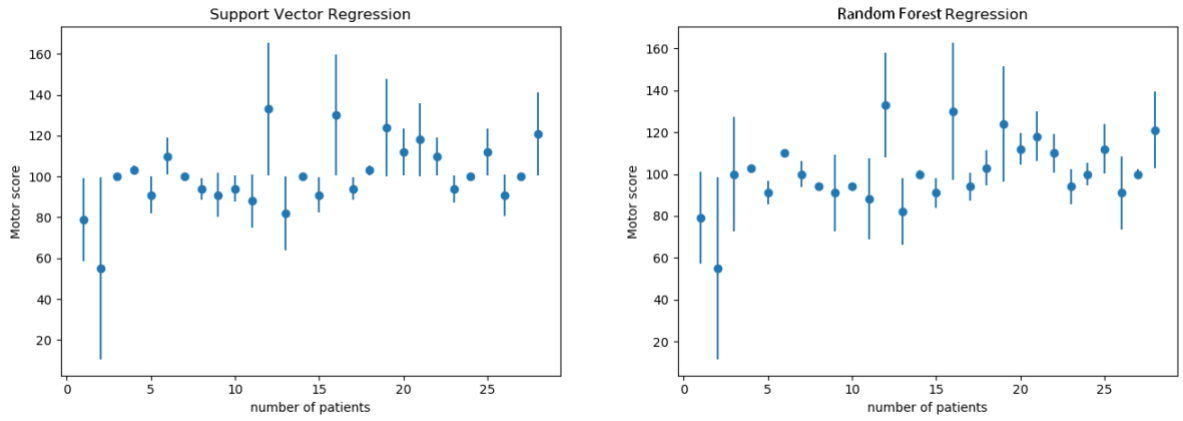


d.





e. f.



g. h.

Figure 4. 5 (a) SVR with normalized features for cognitive score, (b) RFR with normalized features for cognitive score, (c) SVR with normalized features for motor score, (d) RFR with normalized features for motor score, (e) SVR with unnormalized features for cognitive score, (f) RFR with unnormalized features for cognitive score, (g) SVR with unnormalized features for motor score, (h) RFR with unnormalized features for motor score

4.5. Summary

In this chapter, we examined the structure and signal intensities of venous vessels to extract some features from SWI images of infants for HIE outcome prognosis. We used width, intensity and length to classify an infant into one of the neurologically normal or abnormal groups. In our approach, we used the Active Contour model to remove the background, which is outside the brain tissue. Then we applied the ridge detection method to obtain the ridges representing thevessels in order to enable us to compute histograms of *width*, *intensity* and *length* of vessels as our feature vectors. To this end, we applied a vessel segmentation method to segment each vessel in each SWI image with a separate label. We computed 1) the minimum Euclidean distance of the pixels on the vessel edge as width value, 2) the difference between the ridge point pixels intensity and edge pixels intensity as intensity value of the ridge and 3) the length of segmented ridges by counting the number of pixels labelled as a ridge. Meanwhile, we also measured the eigenvalues of the Hessian matrix for all ridge pixels and their neighbourhood and then normalised as well as concatenated all these features for classification. We balanced our training dataset to avoid any classification bias. All four features and their combinations were fed to *kNN* and random forest classifiers with a leave-one-out cross-validation strategy. The concatenated features consisting of *width*, *intensity*, *length* and *Hessian eigenvalue* in the random forest classifier present the best accuracy rate of 78.67±2.58%. Our result is better than the classification accuracy reported in [76] which is 72.27±4.85%. It is also

noted that the data used in [76] is imbalanced and the classification in [76] has been achieved using the scale proposed by Kitamura [45]. However, in this chapter, our HIE classification is carried out by using a balanced dataset and classification accuracies have been evaluated based on two neurologically normal and abnormal outcomes for infant patients. Finally, we trained the SVR model to predict cognitive and motor scores and use mean relative errors to measure the performance of the regression. Due to the lack of data, the lowest relative error of our SVR model is associated with motor scores, with a mean relative error of 0.109±0.067. We expect these results to improve if we obtain more data on the scores for clinical outcomes. As a result, our method improves the prognostic value of SWI images in HIE.

Chapter 5

Brain Regions Affected by Hypoxic-Ischaemic

Newborns diagnosed with hypoxic-ischaemic encephalopathy (HIE) at birth will have long-term neurodevelopmental impairment [89]. In the previous chapter, we developed an automatic system to predict the developmental outcome of infants with HIE at the age of 24 months from susceptibility-weighted imaging (SWI) images scanned at birth by detecting and analysing veins on neo-natal SWI images. However, it would be interesting to find the brain regions which cause the long-term neurodevelopment impairment of infants with HIE. In this chapter, the feature vectors extracted from the SWI images of infants with HIE are further numerically analysed to determine the brain regions which have been affected by neonatal HIE.

5.1. Introduction

Perinatal asphyxia will cause hypoxic-ischaemic brain injury in infants at birth. Generally, newborns with a clinical diagnosis of HIE will be treated with therapeutic hypothermia (TH), which is an effective neuroprotective therapy [47]. Therapeutic hypothermia (TH) can reduce the risk of neurodevelopment deficits in infants with moderate to severe HIE, but survivors still show widespread neurodevelopment impairment [49]. Recent research illustrates that infants

with mild and moderate HIE suffer from long-term motor and cognitive deficits [89]. Following the use of neuroimaging as a diagnostic and prognostic tool for HIE, studies using neuroimaging of infants with HIE to analyse the relationship between areas of the brain damaged by hypoxic-ischaemic (HI) and long-term neurodevelopment outcomes have become increasingly popular.

Tusor et al. [90] revealed that the fractional anisotropy values in the corpora callosa and posterior limbs of the internal capsule of neonates with unfavourable outcomes were still lower. This was done by analysing the relationship between the fractional anisotropy values exported from diffusion tensor imaging in different regions of the neonatal brain with HIE and the neurodevelopmental performance. A work [91] showed that the functional connectivity in prefrontal-hippocampal networks was disrupted in neonatal brains suffering from HIE, which was related to cognitive function. In another study [92], a high correlation was found between cerebral glucose metabolism and the severity of HIE, with the most active brain regions being the thalamus, the basal ganglia and the sensorimotor cortex.

Jiang et al. [93] evaluated resting-state functional connectivity in the motor network of 16 infants with HIE and 11 neurologically intact infants using T2-weighted images and analysed the differences in the functional connectivity between the two groups. The T2-weighted images were aligned to the neonatal template and segmented into white matter, grey matter and cerebrospinal fluid. For each infant in the two groups, spherical regions of interest (ROI) were placed in the motor cortex and motor association regions of each hemisphere on the T2 images and then these ROIs were extracted for analysis. The results showed that infants with HIE had reduced resting-state functional connectivity between primary motor areas and between hemispheres within the primary motor cortex compared to neurologically intact infants. This study reveals potential deficits in primary and co-motor areas in newborns with HIE but lacks the support of follow-up results. Megan et al. [94] collected 66 newborns with HIE classified according to the Sarnat system (mild, moderate and severe). The magnetic resonance imaging (MRI) scanning (including T1- and T2-weighted imaging, diffusion-weighted imaging and apparent diffusion coefficient mapping) were performed on 66 newborns with HIE after TH within 14 days of birth. MRI images of infants with HIE were scored using three scoring systems: the Barkovich score, the National Institute of Child Health and Human Development score and the Weeke score, and neurodevelopmental outcomes were assessed by Bayley Scales (Bayley-III) at two years of age. By analysing the relationship between the different scoring systems and neurodevelopmental outcomes, the results showed that these three scoring systems

were associated with the Bayley-III cognitive and motor composite scores at age two. In the Barkovich scoring system, a damaged basal ganglia/thalamus was associated with severe motor impairment, and the watershed injury pattern was related to cognitive impairment. However, this study was limited by the experts who scored the MRI images. The reliability was inconsistent between different scorers.

Given the correlation of cerebral blood flow with HI injury and the clearly visible borders and morphological features of micro-haemorrhages in SWI images after HIE [69], we propose an algorithm to separate vessels damaged by HIE from normal vessels. Such an algorithm would then enable us to find the regions of the brain affected by HI injuries, and it would help us to assess the associations of the affected brain regions with the neurodevelopmental outcomes.

5.2. Methods

5.2.1. Image pre-processing

The data and ethics approvals used in this chapter are taken from Chapter 3. We used the methods in Chapter 4 for image pre-processing and feature extraction of blood vessels on SWI images. While measuring the length of blood vessels in the neonatal brain, the coordinates of the start point and the identification number of the SWI slice where the vessel is found, and the identification number for each patient are stored row by row in a location vector to enable us to find the patient who this vessel is associated with as well as to locate the vessel inside the brain (x, y, and slice number).

5.2.2. Experimental design and approach

Based on our results in Chapter 4, our hypothesis is that in SWI images, there are some changes in the appearances of some vessels affected by HIE. Therefore we can assume that in a brain affected by hypoxia-ischaemia, there are two groups of vessels: normal and abnormal. Here we have used the term “abnormal” for vessels that are affected by HIE and therefore appear differently in SWI images from normal vessels. We are therefore aiming to locate the damaged brain regions by detecting these abnormal vessels and investigating if it is possible to predict the outcomes of clinical assessments at birth (24 months before the infants are clinically assessed) by analysing the patients’ SWI images. However, it is tedious and time-consuming for medical experts to annotate such abnormal vessels. In the absence of ground truth data

(vessels annotated/labelled by medical experts as abnormal), it is not trivial to separate normal and abnormal vessels in any given SWI.

We, therefore, propose ***Algorithm 1*** to be able to separate normal and abnormal vessels. ***Algorithm 1*** is a supervised classification method to classify vessels into normal and abnormal groups without using any ground truth for normal and abnormal vessels by using the outcomes of clinical assessments as indirect supervision. In order to supervise the classification process in our algorithm proposed here, the neurological, motor and cognitive outcomes obtained from clinical assessments are used as ground truths to train ***Algorithm 1***.

***Algorithm 1*: Training**

TV= total number of vessels in each patient

MD = Mahalanobis Distance

-Extract width, length, intensity and eigenvalue histograms (features) of all vessels

-Calculate the Mean and Covariance Matrix of vessels in feature space for Normal patients

// Finding optimal value

-Set interval ranges r_t for threshold T, r_q for percentage q

-FOR each vessel belonging to a patient, $T \in r_t$ do

-Count the number of vessels (c) with MD >T

FOR $q \in r_q$ do

IF $^{c}/_{TV} \geq q\%$

Classify the patient as abnormal

ELSE

Classify the patient as normal

END // for IF

-Calculate the accuracy using the number of correctly Classified patients

-Increase q

END // for q

-Increase T

END // for T

Find T and q with highest classification accuracy.

As can be seen from the pseudo-code of the training of ***Algorithm 1***, we initially consider all vessels in all with normal outcomes. Each vessel is represented by a feature vector which consists of concatenated feature vectors extracted from SWI images as described in Chapter 4. Therefore, in our feature space, there are vessels represented by their feature vectors extracted

from slices of all patients with normal outcomes in our dataset. We then compute the centre and covariance matrix for these feature points (vessels) representing normal patients in the feature space. Each vessel is defined as a vector with three concatenated features (*width, intensity,* and *eigenvalue histograms*) as well as the value of the vessel *length*. Next, all vessels for all patients with abnormal outcomes are compared with our original feature points in the feature space. Then every vessel in the feature space is classified into one of two groups: normal or abnormal vessels. The Mahalanobis Distance (MD) [95] of a feature point from the original group of feature points in our feature space is computed. If the MD of this feature point is less than a threshold *T*, this feature point is considered to be a member of a group of normal vessels. However, if this MD is more than threshold *T*, then the feature point is considered to be a member of a group of abnormal vessels. MD is calculated as:

$$D = \sqrt{(x-\mu)^T S^{-1}(x-\mu)}$$

μ is the mean vector, S is the covariance matrix for the original group (associated with normal patients) of feature points in our feature space, and x is the feature vector associated with a vessel in the feature space. The above operation is performed for all vessels of all SWI slices of a patient.

The MD values of each vessel for each patient are calculated and saved row by row in two lists (normal and abnormal groups). MD values belonging to the same patient will have the same patient label. Each MD value is compared with a set threshold to determine if it is an abnormal value (above the threshold), which is considered to be a vessel affected by HIE. Raising the threshold means that most of the vessels in the abnormal group of patients are classified as normal, and vice versa. In finding the optimal threshold, we follow two rules: 1) if the MD of one or more vessels is judged to be abnormal in all SWI sections of an infant with HIE, then the infant is abnormal; 2) if all vessels in all SWI sections are less than the threshold, then the infant is normal. In general, the optimal threshold should be that all patients in the normal group are classified as normal and all patients in the abnormal group are classified as abnormal with at least one vessel settled in outlier. Therefore an increase or decrease in the T value will directly affect the classification performance of the different groups of patients.

However, SWI images can be contaminated with noise and artefacts. This means that SWI images from patients with normal neurological findings can show abnormal signals, resulting in vessels with large MD values being classified as abnormal. To eliminate the effects of noise and artefacts, we introduced percentages of vessels to ensure accuracy of classification, i.e. a

patient with a certain percentage of total vessels with an MD value greater than a threshold is classified as abnormal. Therefore, we employ the following rule:

> *If q% or more vessels of a patient are considered to be abnormal, then the patient is considered to have an abnormal outcome. Otherwise, the patient is considered to have a normal outcome.*

Therefore, it becomes important to determine the T value and percentage. With certain values for q and T, ***Algorithm 1*** predicts some patients with normal outcomes and some other patients with abnormal outcomes. This is because each change in T or q value will affect the classification of patients in the normal and abnormal groups. We then measure the accuracy of our algorithm for these values of q and T by comparing the outcome of our algorithm with the outcome of the clinical assessments (ground truth). In order to examine the contributions of different T values including small changes on the classification of patients, we use the exhaustive method to measure the variation in classification accuracy for all T values in steps of 0.01 from 0 to the maximum MD value and for all q values in steps of 0.1% from 0. Table 5.1, 5.2, and 5.3 show the different accuracies as T and q are changed at the steps for neurology outcome, motor outcome, and cognitive outcome. Due to space constraints, we have only shown part of the tables.

Table 5. 1 The accuracies of classification with all T and q values for neurology outcome

Threshold T	... q=1.9%			q=2%			q=2.1% ...		
	Normal group	Abnormal group	Total patients	Normal group	Abnormal group	Total patients	Normal group	Abnormal group	Total patients
…	…	…	…	…	…	…	…	…	…
7.6	0.73	0.78	0.76	0.63	0.875	0.81	0.54	0.875	0.76
7.61	0.73	0.81	0.79	0.63	0.875	0.81	0.54	0.91	0.81
7.62	0.73	0.84	0.81	0.63	0.91	0.84	0.54	0.93	0.84
7.63	0.73	0.88	0.83	0.63	0.91	0.84	0.54	0.93	0.84
7.64	0.73	0.91	0.86	0.54	0.94	0.84	0.45	0.96	0.84
7.65	0.73	0.91	0.86	0.54	0.94	0.84	0.45	0.96	0.84
7.66	**0.73**	**0.94**	**0.88**	0.54	0.96	0.86	0.45	0.96	0.84
7.67	0.64	0.94	0.86	0.54	0.96	0.86	0.45	0.96	0.84
7.68	0.64	0.96	0.88	0.54	0.96	0.86	0.45	0.96	0.84
7.69	0.64	0.96	0.88	0.54	0.96	0.86	0.45	0.96	0.84
7.7	0.64	0.96	0.88	0.54	0.96	0.86	0.45	0.96	0.84
…	…	…	…	…	…	…	…	…	…

Table 5. 2 The accuracies of classification with all T and q values for cognitive outcome

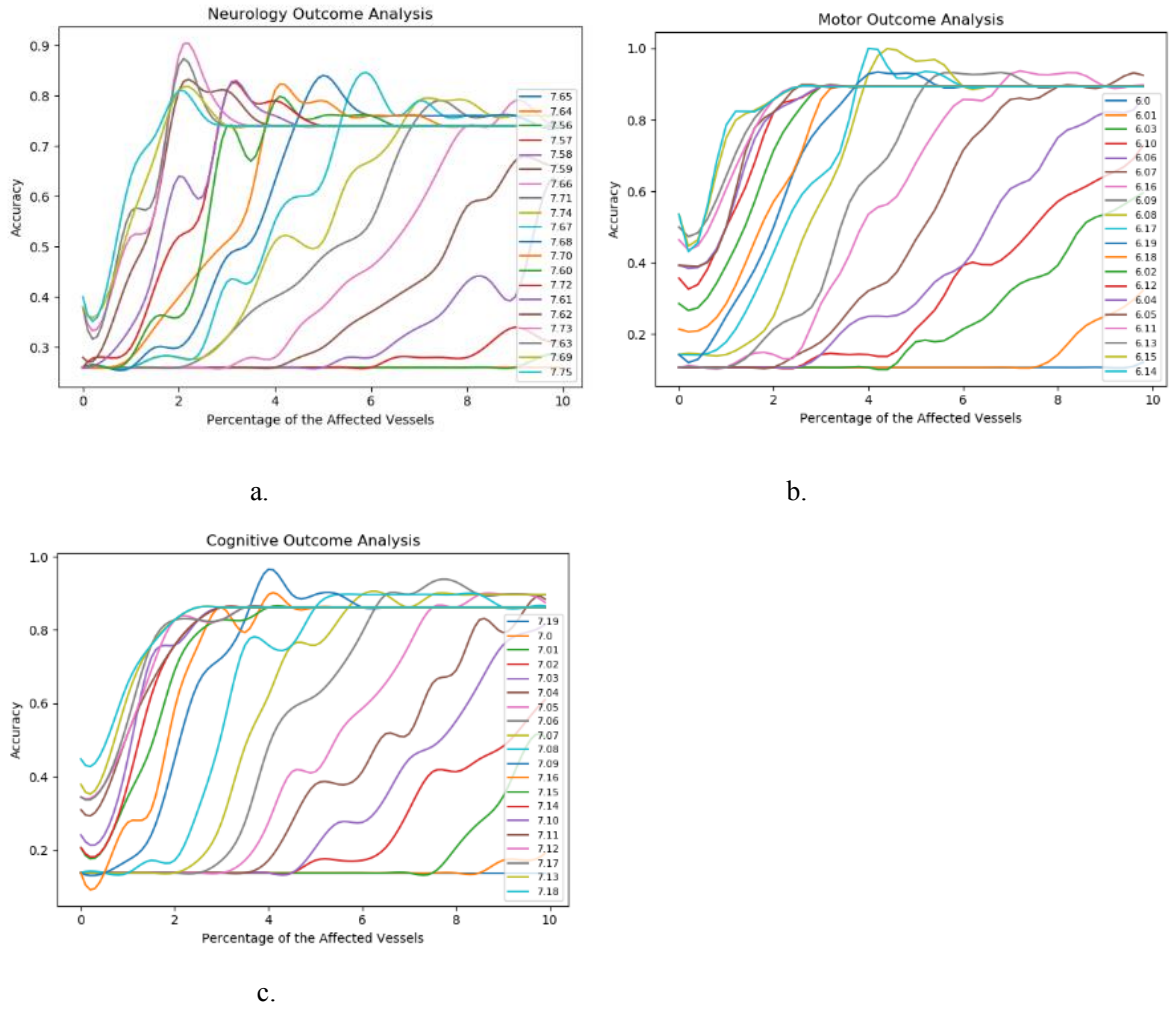
Threshold T	... q=3.8%			q=3.9%			q=4% ...		
	Normal group	Abnormal group	Total patients	Normal group	Abnormal group	Total patients	Normal group	Abnormal group	Total patients
…	…	…	…	…	…	…	…	…	…
7.03	0.75	0.72	0.72	0.75	0.76	0.75	0.75	0.76	0.75
7.04	0.75	0.72	0.72	0.75	0.76	0.75	0.75	0.76	0.75
7.05	0.75	0.76	0.75	0.75	0.8	0.78	0.75	0.8	0.78
7.06	0.75	0.8	0.78	0.75	0.88	0.84	0.75	0.88	0.84
7.07	0.75	0.8	0.78	0.75	0.88	0.84	0.75	0.88	0.84
7.08	0.75	0.84	0.81	0.75	0.96	0.93	0.75	0.96	0.93
7.09	0.75	0.88	0.84	**0.75**	**1**	**0.96**	0.75	0.96	0.93
7.1	0.75	0.88	0.84	0.75	1	0.96	0.5	1	0.93
7.11	0.75	0.88	0.84	0.75	1	0.96	0.5	1	0.93
7.12	0.75	0.88	0.84	0.75	1	0.96	0.5	1	0.93
…	…	…	…	…	…	…	…	…	…

Table 5. 3 The accuracies of classification with all *T* and *q* values for motor outcome

Threshold *T*	... *q*=4%			*q*=4.1%			*q*=4.7% ...		
	Normal group	Abnormal group	Total patients	Normal group	Abnormal group	Total patients	Normal group	Abnormal group	Total patients
…	…	…	…	…	…	…	…	…	…
6.05	0.66	0.36	0.39	1	0.6	0.57	1	0.92	0.92
6.06	0.66	0.439	0.46	1	0.76	0.78	1	0.92	0.92
6.07	0.33	0.76	0.71	1	0.8	0.82	1	0.96	0.96
6.08	0.66	0.76	0.75	1	0.84	0.86	**1**	**1**	**1**
…	…	…	…	…	…	…	…	…	…
6.14	0.33	0.8	0.75	1	0.92	0.92	0.33	1	0.92
6.15	0.33	0.96	0.89	1	0.96	0.96	0.33	1	0.92
6.16	0	1	0.89	1	0.96	0.96	0.33	1	0.92
6.17	0	1	0.89	**1**	**1**	**1**	0	1	0.89
6.18	0	1	0.89	0.66	1	0.96	0	1	0.89
…	…	…	…	…	…	…	…	…	…

By changing *q* and *T*, we measure the accuracies of ***Algorithm 1*** for all values of *q* and *T* as shown in table 5.1, 5.2 and 5.3. Finally, we choose *q* and *T* values that correspond to the maximum accuracy. Figure 5.1(a) shows how ***Algorithm 1*** accuracies change with respect to *q* by changing *T*. In figure 5.2(a), different accuracy curves with respect to *q* for the neurological assessment outcomes are plotted in different colours for the different values of *T*. Ideally, the maximum for ***Algorithm 1*** accuracies is desired to be 100%. However, due to the

distribution of vessel feature vectors in the feature space, this may not be achievable. In such cases, we choose the maximum accuracy over all of the T and q values for an assessment outcome. Figures 5.1(b) and 5.1(c) also depict the accuracy curves of our algorithm over T and q for motor and cognitive developmental outcomes. Having found the optimal values for T and q for each assessment outcome, we then manage to separate normal and abnormal vessels in our feature space.



a. b.

c.

Figure 5. 1 Choosing threshold T and percentage q where the accuracy in the training stage is at maximum for a) Neurology outcome analysis, b) Motor outcome analysis and c) Cognitive outcome analysis

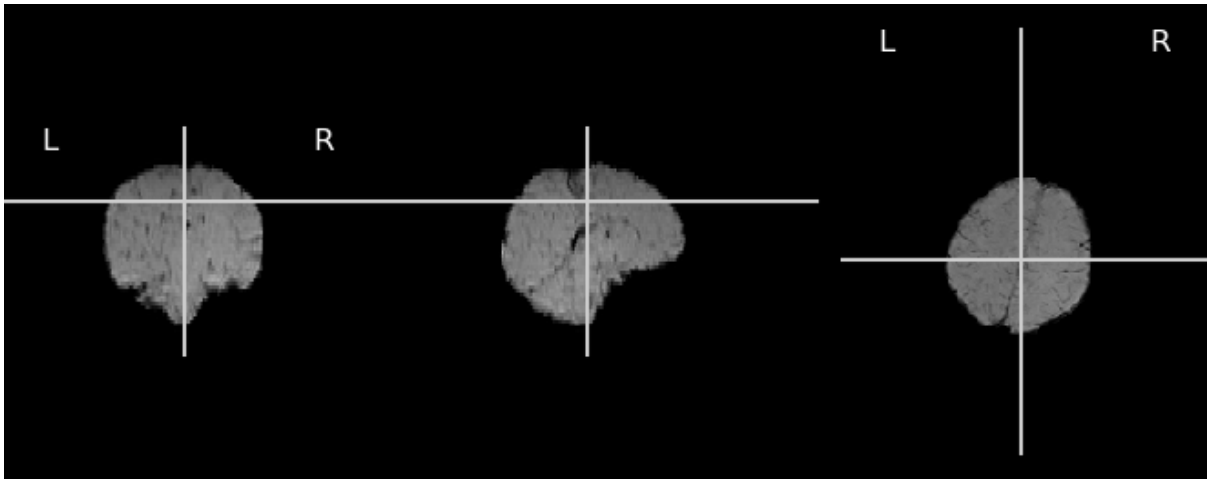
5.2.3. Image registration

The neurological assessment involves assessing various motor phenomena [30, 36]. We believe an abnormal neurological outcome is caused by damages in the posterior section of the frontal lobe in the brain of infants with HIE that are within the primary motor area, pre-motor area and supplementary motor area related to motor functions. We explore the relationship between the percentage of the affected vessels in these regions and the neurology outcomes. The motor

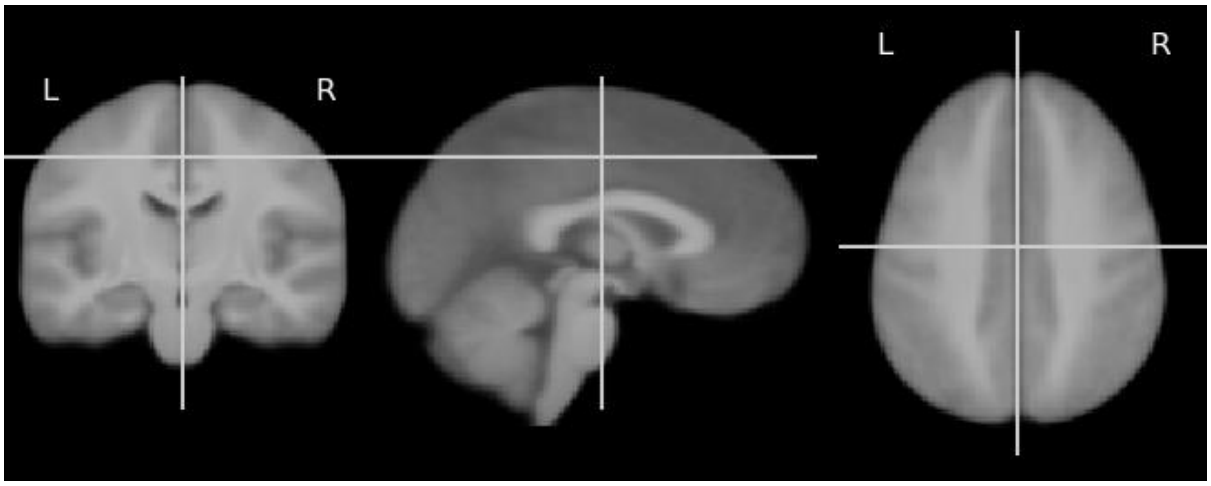
cortex is extensively engaged while using the Bayley-III assessment on two-year-old infants with HIE [54, 96]. We predict that motor scores are related to the percentage of abnormal vessels in the motor areas of the brain, as in the primary motor area, pre-motor area and supplementary motor area implicated in the motor outcomes. Cognitive functions are always interrelated using different brain areas covering attention, memory, language and executive functions. As measuring the number of abnormal vessels is related to cognitive functions, three areas of the brain are considered in our analysis: the frontal lobe, the temporal lobe and the parietal lobe.

In order to look for the brain regions affected by HIE, we register the atlas, including the average intensity image, the tissue density maps, the structure density maps, and the maximum probabilistic maps and labels, as a reference template with the SWIs to identify individual lobes in the SWIs. The brain template LPBA40/AIR (https://resource.loni.usc.edu/resources/atlases-downloads/) and Brodmann template are selected for this study. The image registration of SWI datasets is carried out using Advanced Normalisation Tools (ANTs) [97-98], which have the best quality for the registration of brain magnetic resonance images. We convert all SWI images of each infant in our dataset into 3D brain images to be registered. The strategy for the ANTs registration programme, for which we opt, is to map the SWI images onto the template brain images using similarity transform and obtain the registered SWI images.

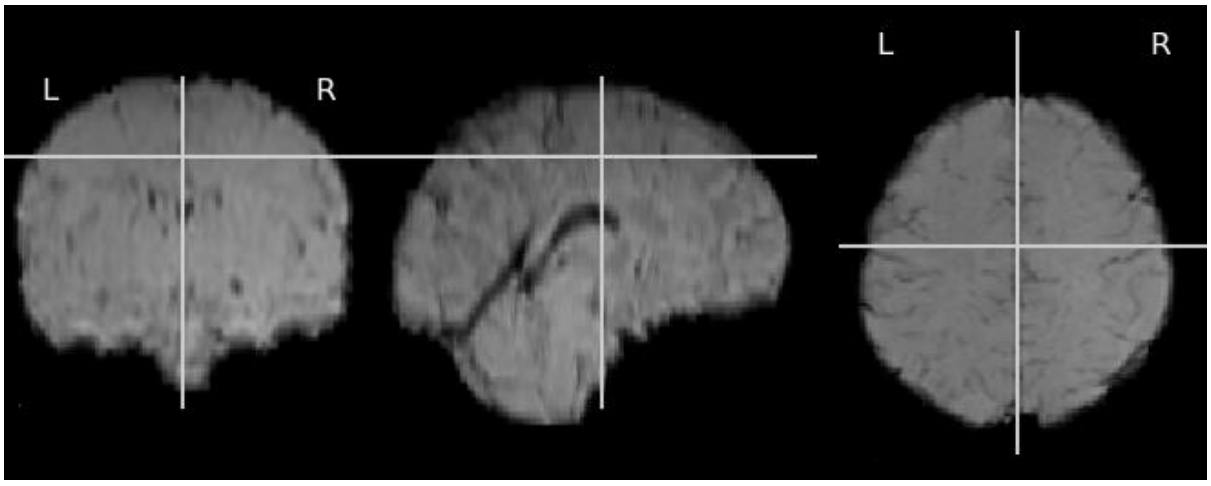
The LPBA40/AIR template provides a standard normalised space containing 56 brain structures and partition labels, such as the frontal lobe and parietal lobe [99]. Since the SWI images of each infant with neonatal HIE are transformed/registered into a template brain of an atlas-based on an image registration method, we map the 56 labels of the maximum probability maps onto the new registered SWI images for analysis. We eventually consider the primary motor area, pre-motor area, and supplementary motor area of the 3D images and explore the relationship between the vessel features of SWI in these areas and neurology and motor outcomes at 24 months of age. To explore the relationship between vessel features in SWI images and cognitive outcomes, the frontal lobe, parietal lobe and temporal lobe of the brain as cognitive areas were examined. Figure 5.2(c) displays the SWI images after registration. As shown in Figure 5.2(c), the area covered in blue represents the motor area. The motor areas of the brain in SWI images are therefore selected by registering SWI images to the template brain.



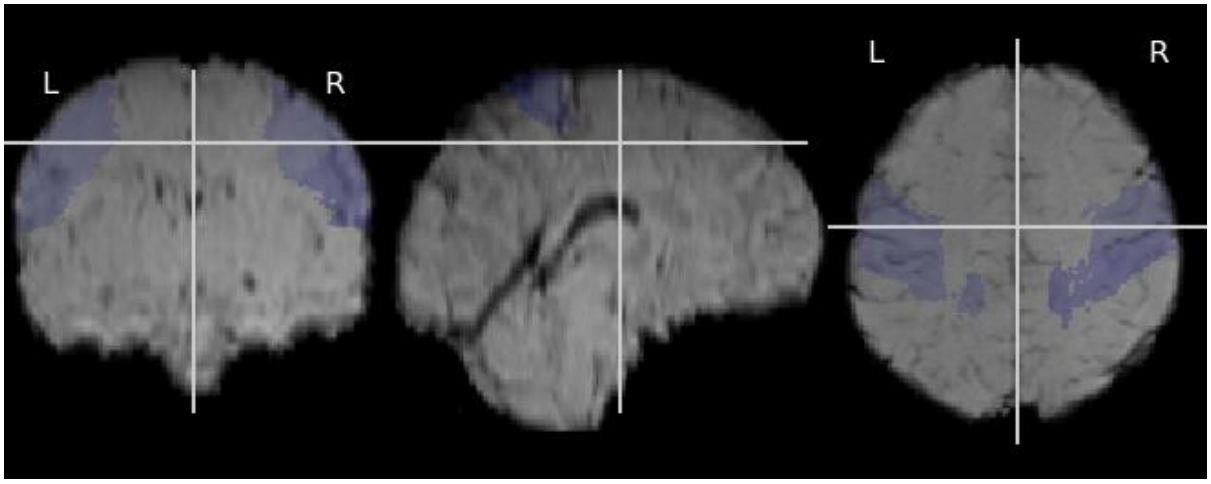
a.



b.



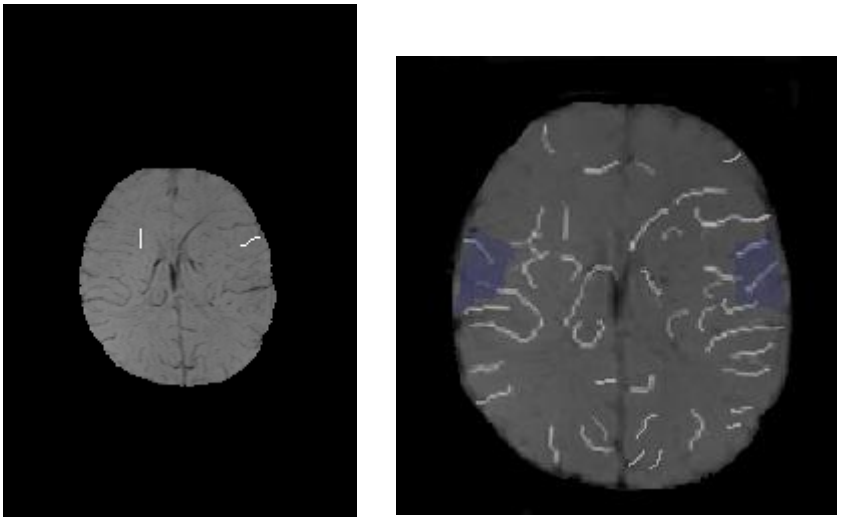
c.



d.

Figure 5. 2 a) Raw SWI images of infants with HIE; b) template brain; c) SWI images after registration; d. motor area covered by blue colour.

Finally, for the SWI images of neonates with HIE in the neurology outcome group, the motor outcome group and the cognitive outcome group, all of the affected vessels detected through ***Algorithm 1*** are recorded. Based on the experience of clinical experts and previous literature findings, we analyse the relationship between the percentage of abnormal vessels in different regions in SWI images and the developmental outcomes of infants. Two statistical methods are used to determine the location of the affected vessels: one is to observe where the affected vessels belonged on the SWI images according to previous empirical knowledge. The second is to obtain the location of the affected vessels with the help of SWI image registration. In reference to the list established during the measurement of vessel length, the affected vessel is mapped onto a ridge within the labelled motor or cognitive region. If more than half the length of the abnormal vessel is present in these regions, the affected vessel belongs to these regions. As shown in Figure 5.3, two vessels are detected as 'abnormal' by ***Algorithm 1*** from an SWI image of infants with HIE who had an abnormal neurology outcome at the age of two years. Comparing the ridges in the motor areas on the registered SWI images, one 'abnormal' vessel is located in the motor area.



a. b.

Figure 5. 3 a. Two vessels are detected as 'abnormal' by ***Algorithm 1***; b. Ridges on the registered SWI image.

5.3. Results

5.3.1. Relationship between brain regions with affected vessels and neurological outcome

With 42 infants assessed clinically for neurological outcomes, a threshold of T=7.66 and a percentage of q=1.9% correspond with the best accuracy for our algorithm to enable us to separate the normal and abnormal vessels as depicted in figure 5.4 (a). Table 5 shows the effect of small changes in T and q values with respect to the results. The best total accuracy for ***Algorithm 1*** is 88.1%, with 29 out of 31 patients with a normal neurology outcome being correctly grouped (93.55%) and eight out of 11 patients with an abnormal neurology outcome being correctly grouped too (8/11: 93.75%). The fact that the maximum total accuracy here is less than 100% (88.1%) implies that some normal vessels are grouped as abnormal, and some abnormal vessels are grouped as normal. We expect that ***Algorithm 1*** will classify the patients with normal neurology outcome and patients with abnormal neurology outcome into normal and abnormal groups 100% correctly. This is for the training data. We, therefore, propose a fine-tuning method in the following paragraph to find a more accurate boundary, separating normal and abnormal vessels by using a *kNN* classifier.

Fine-tuning: The aim of fine-tuning is to find a more accurate boundary between normal and abnormal vessels in the feature space. Such a fine-tuning technique can improve our algorithm performance to up to 100% in the training stage. To this end, we choose two thresholds: T_i, T_o

in the neighbourhood of the threshold T_{max}, corresponding to the maximum accuracy to ensure that $T_i < T_{max} < T_o$. These thresholds divide the feature space into three regions: The first region is the inner feature subspace considered to contain normal vessels (i.e. MD $<T_i$) and the outer feature subspace (and second region) is considered to accommodate abnormal vessels (i.e. MD $>T_o$). The third region is the middle subspace between these two above regions (i.e., $T_i <$ MD$<T_o$). By setting these two thresholds, some feature points are placed in the middle subspace between the inner and outer subspaces in the feature space. The feature points in the middle subspace are considered as a test set, and the rest of the feature points inside the inner and outer regions are used to train a *kNN* classifier. This *kNN* classifier classifies the feature points in the middle subspace (the test set) into the normal group (in the inner subspace) and the abnormal group (in the outer subspace) to determine a more accurate boundary between normal and abnormal groups in the feature space. We keep decreasing T_i and increasing T_o step by step and measure the classification accuracy in each step. We stop the fine-tuning process if either the classification accuracy reaches 100% or the maximum classification accuracy decreases with respect to the previous iteration. In our experiments, with the first step, we have always reached 100% accuracy and therefore terminate the fine-tuning process. It is noted that q and T_{max} are kept constant during the fine-tuning process.

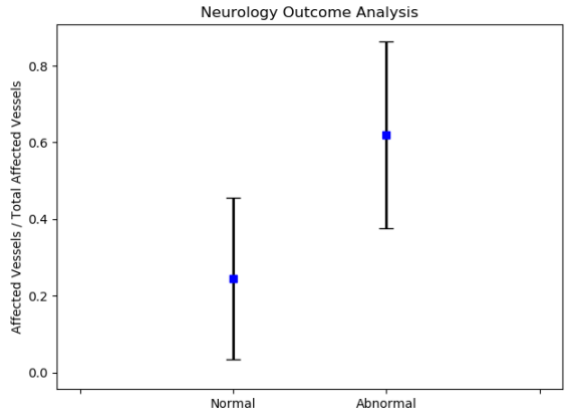
Once each patient has met our rule (q=1.9%) and been correctly grouped, the fine-tuning is complete at 100% accuracy. Here, the two thresholds are 7.6 and 7.71. Before analysing the motor regions damaged by the HI injury, based on neurological outcomes, we perform a validation experiment with balanced data (11 patients with abnormal neurology and 11 patients with normal neurology randomly selected from 31 patients in the group with normal neurology as the training dataset) to measure the performance of ***Algorithm 1***. A leave-one-out strategy is employed here, which means that ***Algorithm 1*** is trained on all balanced data except for one test patient. Having trained our algorithm by computing T and q, we then measure the number of cases (classification accuracy) where the test patient is classified correctly. The random selection of 11 out of 31 patients from the normal group is repeated ten times, and the classification accuracy is measured each time. The final accuracy is calculated as 0.727±0.056 (the mean ± standard deviation over these ten experiments). Table 5.4 shows the classification results of these ten experiments.

Table 5. 4 Performance of classification over ten times

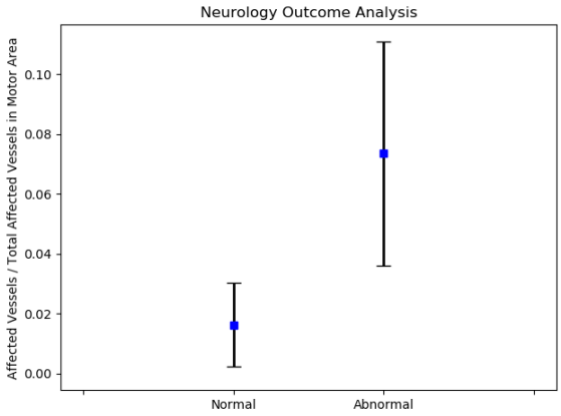
Actual patients / Test patients		Neurology outcome (Normal)	Neurology outcome (Abnormal)
1st experiment	Neurology outcome (Normal)	7	2
	Neurology outcome (Abnormal)	4	9
2nd experiment	Neurology outcome (Normal)	9	4
	Neurology outcome (Abnormal)	2	7
3rd experiment	Neurology outcome (Normal)	8	3
	Neurology outcome (Abnormal)	3	8
4th experiment	Neurology outcome (Normal)	8	4
	Neurology outcome (Abnormal)	3	7
5th experiment	Neurology outcome (Normal)	9	5
	Neurology outcome (Abnormal)	2	6
6th experiment	Neurology outcome (Normal)	10	2
	Neurology outcome (Abnormal)	1	9
7th experiment	Neurology outcome (Normal)	9	4
	Neurology outcome (Abnormal)	2	7
8th experiment	Neurology outcome (Normal)	7	3
	Neurology outcome (Abnormal)	4	8

9th experiment	Neurology outcome (Normal)	8	2
	Neurology outcome (Abnormal)	3	9
10th experiment	Neurology outcome (Normal)	9	5
	Neurology outcome (Abnormal)	2	6

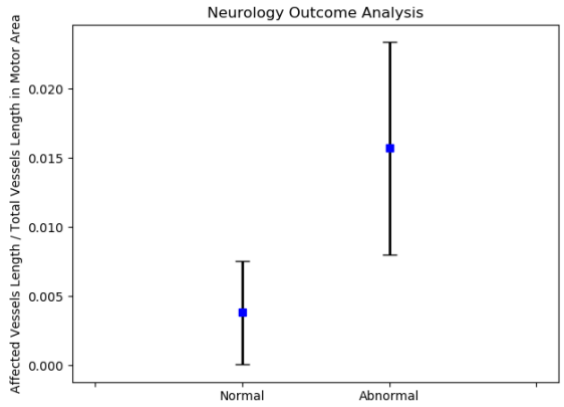
After the fine-tuning process, the abnormal vessel percentage for each patient is recalculated. We explore the relationship between the percentage of the affected vessels in the cortical motor regions of the brain (the primary motor area, the pre-motor area and the supplementary motor area) and the neurology outcomes. In these 42 patients, all of the vessels grouped as abnormal by ***Algorithm 1*** with fine-tuning are counted. Finally, the ratio of the number of abnormal vessels in the motor areas to the total number of abnormal vessels, the ratio of the number of abnormal vessels in motor areas to the total number of vessels in the motor areas, and the ratio of the length of abnormal vessels in motor areas to the total length of vessels in motor areas are measured separately for each patient in the normal and abnormal groups to evaluate the relationship between the neurological outcomes and motor cortex damage for neurological outcomes at the age of 2 years. For the number of abnormal vessels detected by ***Algorithm 1*** in the motor area, the abnormal vessel counts and the ratios measurements are achieved by the observer's experience and the registering SWI image with labelled motor areas from a brain template. As demonstrated in Figure 5.4, the number and percentage of the affected vessels in the first six figures (i-vi) were counted and calculated by the observer, and the number and percentage of the affected vessels in the last six (a-f) were counted automatically and calculated by comparison with the labelled areas. The mean and standard deviation of the ratios in infants with a normal neurology outcome and infants with an abnormal neurology outcome is shown in Figure 5.4, which shows the individual ratio as blue dots for infants with a normal neurology outcome and red dots for infants with an abnormal neurology outcome in scatterplots.



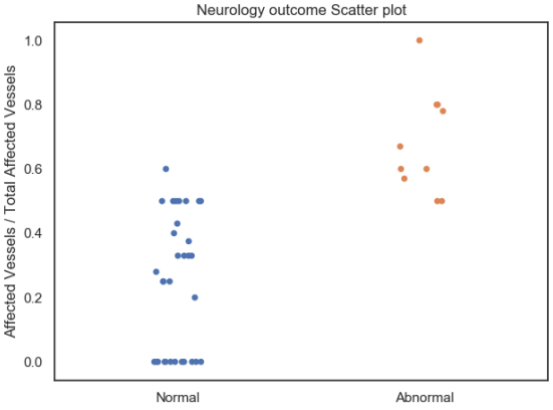
i.



ii.

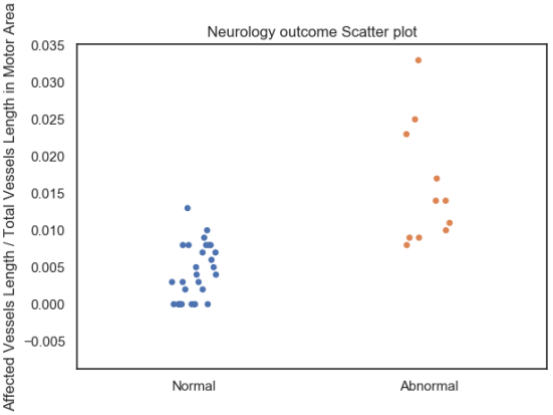


iii.

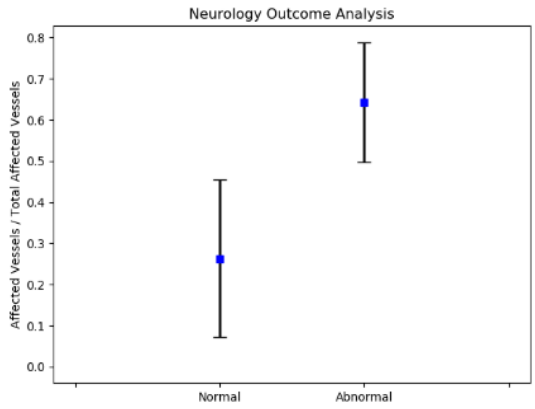


iv.

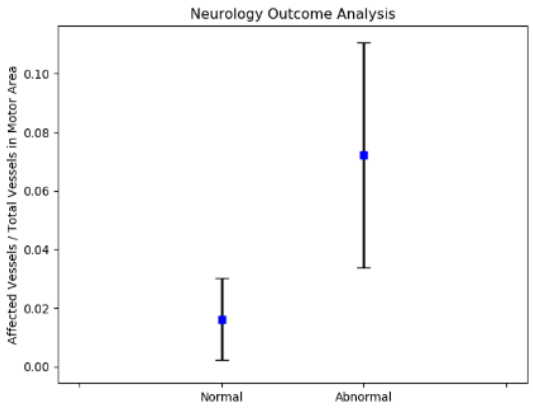
v.



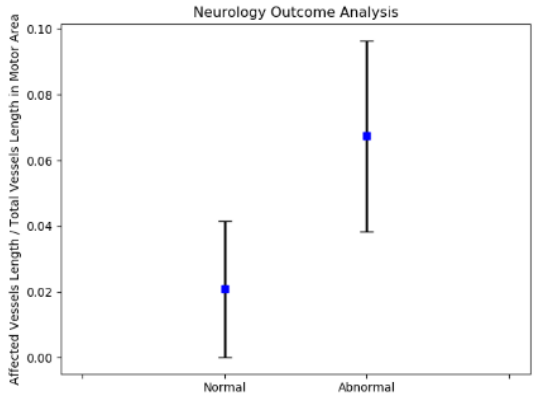
vi.



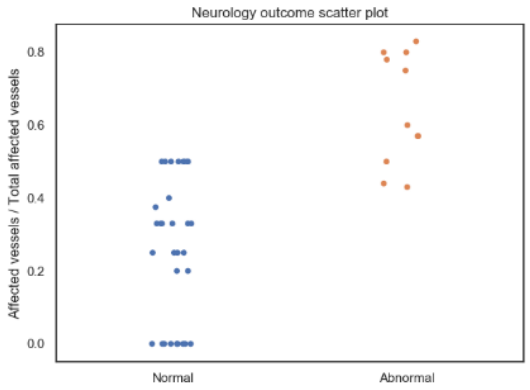
a.



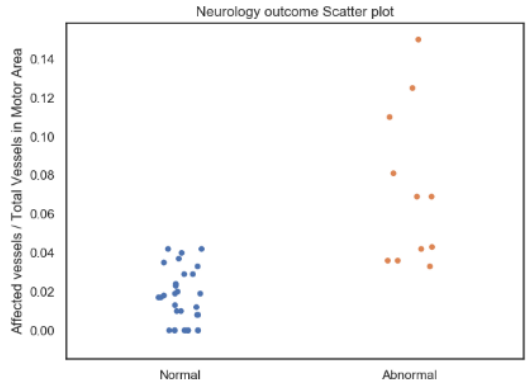
b.



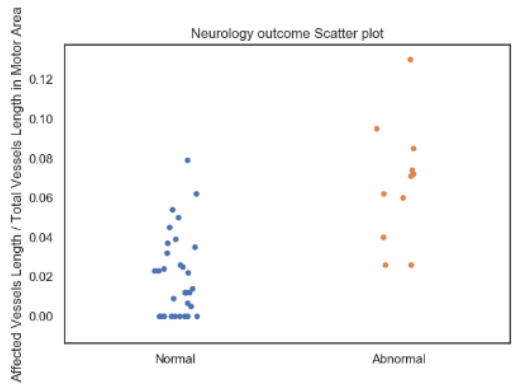
c.



d.



e.



f.

Figure 5. 4 Ratio of the number of affected vessels in the motor area to the total number of affected vessels: i) and iv) are by the observer, and a) and d) are from the labelled motor areas; Ratio of the number of affected vessels in the motor area to the total number of vessels in the motor area: ii) and v) by the observer, and b) and e) are from the labelled motor areas; Ratio of the length of affected vessels in the motor areas (primary, pre- and supplementary) to the total length of vessels in the motor area: iii) and vi) are by the observer, and c) and f) are from the labelled motor areas.

As observed in Figure 5.4 (b and e), there is a significant difference in the ratios between the normal and abnormal groups. And after the SWI image registration, there is less overlap in ratios between the normal and abnormal groups. Our results indicate that the ratio of affected vessels in the motor regions by SWI images to the total number of vessels in the motor areas is correlated to neurological outcomes after HIE. Consequently, the neurological outcome at the age of two years could almost always be predicted by the ratios of the affected vessels described above in the motor areas of cerebral SWI in newborns with HIE. Due to the initial misclassification where two patients with normal neurology are marked as having abnormal motor outcomes and three patients with abnormal neurology are classified as patients with normal neurology, even after fine-tuning, there is some overlap between the range of ratios for the percentages of affected vessels in patients with normal and abnormal neurology.

5.3.2. Affected brain regions associated with delayed motor development

Of the 28 infants with HIE who were assessed with Bayley-3 scales, 25 infants have normal motor development (>85), two infants have mild motor delay (77.5-85), and one has severe motor delay (<77.5). As observed in Figure 5.1(b), the highest total accuracy of 100% is achieved with a threshold T of 6.17 and a percentage q of 4.1%, and a threshold T of 6.08 and a percentage q of 4.7%. These are the ideal values for an optimal threshold T and percentage q, where patients in both the normal group with the normal motor outcome and the abnormal group with mild motor delay and severe motor delay are 100% correctly grouped by ***Algorithm 1***.

Before measuring the number of affected vessels in the motor regions of the brain, we perform an experiment to measure the classification accuracy of ***Algorithm 1*** with unseen test data for validation. We use balanced data (three patients with delayed motor scores and three patients with normal motor scores randomly selected from 25 from the normal motor score group). A leave-one-out strategy is then used here to measure the accuracy. In this balanced data (3 patients with delayed motor scores and 3 patients with normal motor scores), only one patient is used as the test data and the rest are all put into ***Algorithm 1*** as the training data. The final accuracy of classification using our algorithm to separate patients with delayed motor scores from normal motor scores is computed by repeating the above process ten times with random selections from the patients' dataset with normal motor outcomes to achieve an accuracy of 0.75±0.139 (mean ± standard deviation). Table 5.5 shows the classification results of these ten experiments.

Table 5. 5 Performance of classification over ten times

Actual patients / Test patients		Normal motor score	Abnormal motor score
1st experiment	Normal motor score	2	0
	Abnormal motor score	1	3
2nd experiment	Normal motor score	2	0
	Abnormal motor score	1	3
3rd experiment	Normal motor score	3	1
	Abnormal motor score	0	2
4th experiment	Normal motor score	1	1
	Abnormal motor score	2	2
5th experiment	Normal motor score	1	0
	Abnormal motor score	2	3
6th experiment	Normal motor score	2	1
	Abnormal motor score	1	2
7th experiment	Normal motor score	2	1
	Abnormal motor score	1	2
8th experiment	Normal motor score	2	0
	Abnormal motor score	1	3
9th experiment	Normal motor score	3	0
	Abnormal motor score	0	3
10th experiment	Normal motor score	2	1
	Abnormal motor score	1	2

Based on previous knowledge, we firstly predict that motor scores are related to the percentage of abnormal vessels in the motor areas of the brain. In these 28 patients, all of the vessels are grouped as abnormal with the most optimal thresholds T of 6.17, a percentage q of 4.1%, thresholds T of 6.08 and a percentage q of 4.7% are considered. Having trained our ***Algorithm 1*** with fine-tuning, the following terms are measured:

R_{m1}: The ratio of the number of abnormal vessels in the motor area to the total number of abnormal vessels in the brain.

R_{m2}: The ratio of the number of abnormal vessels in the motor area to the total number of vessels in the motor area.

R_{m3}: The ratio of the length of abnormal vessels in the motor area to the length of total vessels in the motor areas.

Similarly, two counting methods were used for the number of abnormal vessels detected by ***Algorithm 1*** in the motor areas. One is generated by the observer and the other by registering the SWI images. Figure 5.5 shows the ratios under the two counting methods and Table 5.6 shows the correlation between the motor score and R_{m1}, R_{m2} and R_{m3} with T of 6.17 and q of 4.1%.

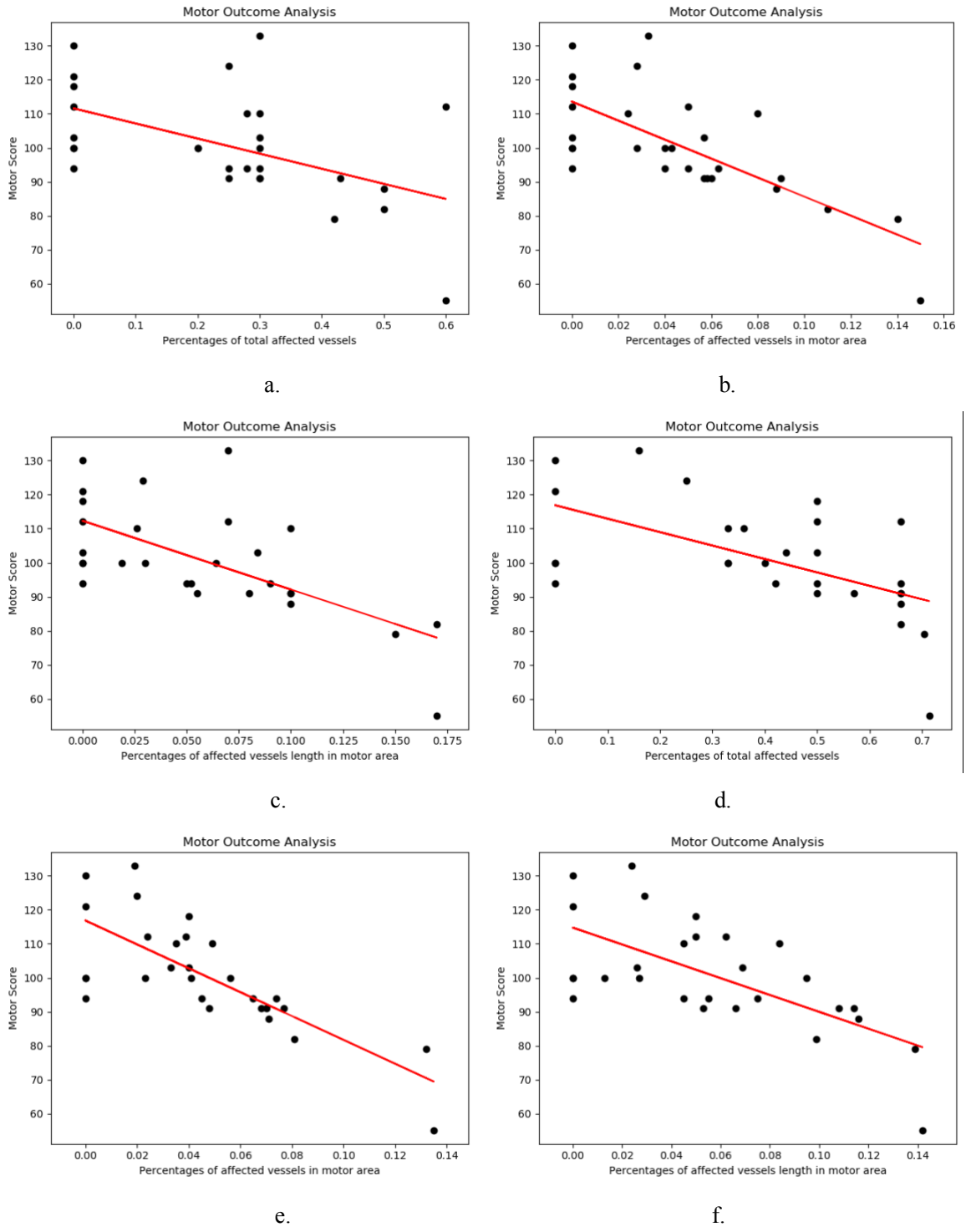


Figure 5. 5 The ratio of the number of abnormal vessels in the motor area to the total number of affected vessels in the brain, a) is by the observer and d) is from the labelled motor areas; The ratio of the number of abnormal vessels in the motor area to the total number of vessels in the motor area, b) is by the observer and e) is from the labelled motor areas; The ratio of the length of abnormal vessels in the motor area to the total length of vessels in the motor area, c) is by the observer and f) is from the labelled motor areas. All plots in this figure are with T of 6.17 and q of 4.1%.

Table 5. 6 Correlation of ratios with motor scores (T of 6.17 and q of 4.1%)

By observer		
Ratio measurement	Motor score	p-value
R_{m1}	-0.51	0.004
R_{m2}	-0.71	0.000016
R_{m3}	-0.65	0.00019
By registering SWI image		
Ratio measurement	Motor score	
R_{m1}	-0.58	0.00116
R_{m2}	**-0.76**	0.000003
R_{m3}	-0.65	0.000164

As shown in Figure 5.5, scatterplots of these ratios with respect to motor scores in the first three figures (a-c) were counted and calculated by the observer and the rest three figures (c-f) by the registering SWI image. Table 5.3 respectively shows the correlation between the motor score and R_{m1}, R_{m2} and R_{m3} with T of 6.17 and q of 4.1%. As observed from this figure, there are negative correlations between motor scores in one side with R_{m1}, R_{m2} and R_{m3} in the other side. It is noted from table 5.6 that the correlation between the motor score and the percentage of the abnormal vessels in motor areas to the total number of vessels in the motor area is -0.76 (p-value < 0.0001) which is the highest among other measurements. The difference between all ratios and motor scores is significant (p-value < 0.005).

Figure 5.6 illustrates the ratios under two counting methods and Table 5.7 illustrates the correlation between the motor score and R_{m1}, R_{m2} and R_{m3} with T of 6.08 and q of 4.7%.

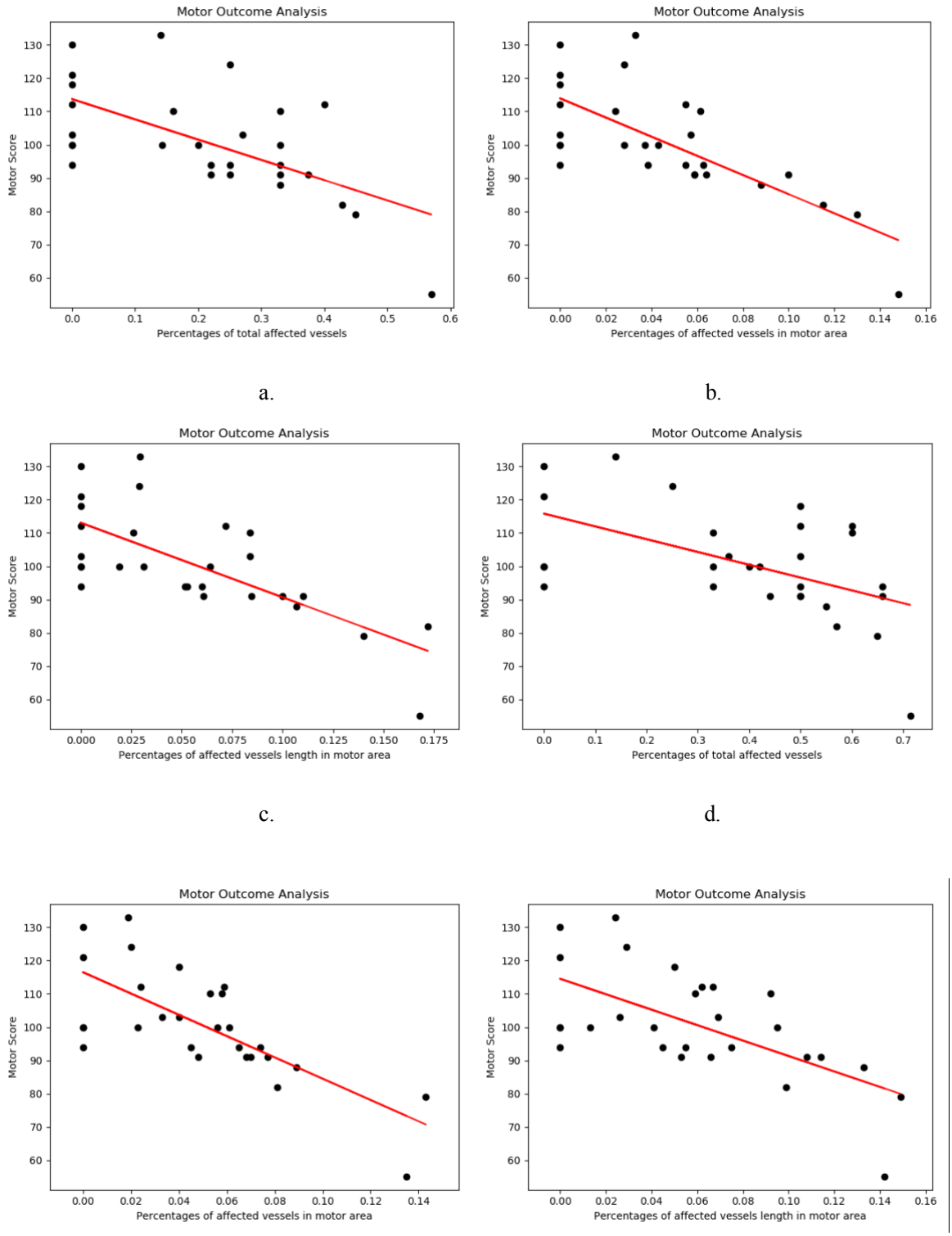


Figure 5. 6 The ratio of the number of abnormal vessels in the motor area to the total number of affected vessels in brain, a) is by the observer and d) is from the labelled motor areas; The ratio of the number of abnormal vessels in motor area to the total number of vessels in motor area, b) is by the observer and e) is from the labelled motor areas; The ratio of the length of abnormal vessels in motor area to the total length of vessels in motor area, c) is by the observer and f) is from the labelled motor areas. All plots in this figure are with T of 6.08 and q of 4.7%.

Table 5. 7 Correlation of ratios with motor scores (T of 6.08 and q of 4.7%)

By observer		
Ratios measurement	Motor score	p-value
R_{m1}	-0.62	0.00045
R_{m2}	**-0.73**	0.000009
R_{m3}	-0.71	0.000022
By registering SWI image		
Ratios measurement	Motor score	
R_{m1}	-0.53	0.0033
R_{m2}	**-0.73**	0.000012
R_{m3}	-0.63	0.00029

As shown in Figure 5.6, scatterplots of these ratios with respect to the motor scores in the first three figures (a-c) were counted and calculated by the observer and the rest of the three figures (d-f) were counted and calculated automatically by the registering SWI image. Table 5.7 shows the correlation between the motor score and R_{m1}, R_{m2} and R_{m3} with T of 6.08 and q of 4.7%. As observed from this figure, there are negative correlations between the motor scores on one side with R_{m1} and R_{m2} and R_{m3} on the other side. As seen in Table 5.7, the correlation between the motor score and the percentage of abnormal vessels in the motor areas to the total number of vessels in the motor area is -0.73, indicating the highest correlation of any other measure. The difference between all ratios and motor scores is significant (p-value < 0.005).

As shown in both Figures 5.5 and 5.6, as the percentage of abnormal vessels increases in the motor cortex, the motor score decreases. The results presented in Figures 5.5 and 5.6 demonstrate that the percentage of abnormal vessels in the motor areas is correlated with motor development for patients with normal and abnormal neurology. For patients whose motor scores are considered delayed (motor score <85), these abnormal vessels are clustered in areas associated with the motor function of the brain. It has been clarified in [135] that HIE patients with poor blood perfusion to several important brain structures (precentral gyrus) had Bayley-III scores below 80. However, the precentral gyrus is the anatomical location of the primary

motor cortex. As shown in the figure 5.7, in patients with a motor score of less than 85, the abnormal vessels detected by ***Algorithm 1*** are almost concentrated in the motor cortex. White ridges represent abnormal blood vessels. The presence of these vessels in the motor cortex of patients with a normal motor score is lower than those in patients with abnormal motor outcomes. Therefore, the number of abnormal vessels found in SW images of the motor cortex of the neonatal brain affected by hypoxia-ischaemia is related to motor development at the age of two years.

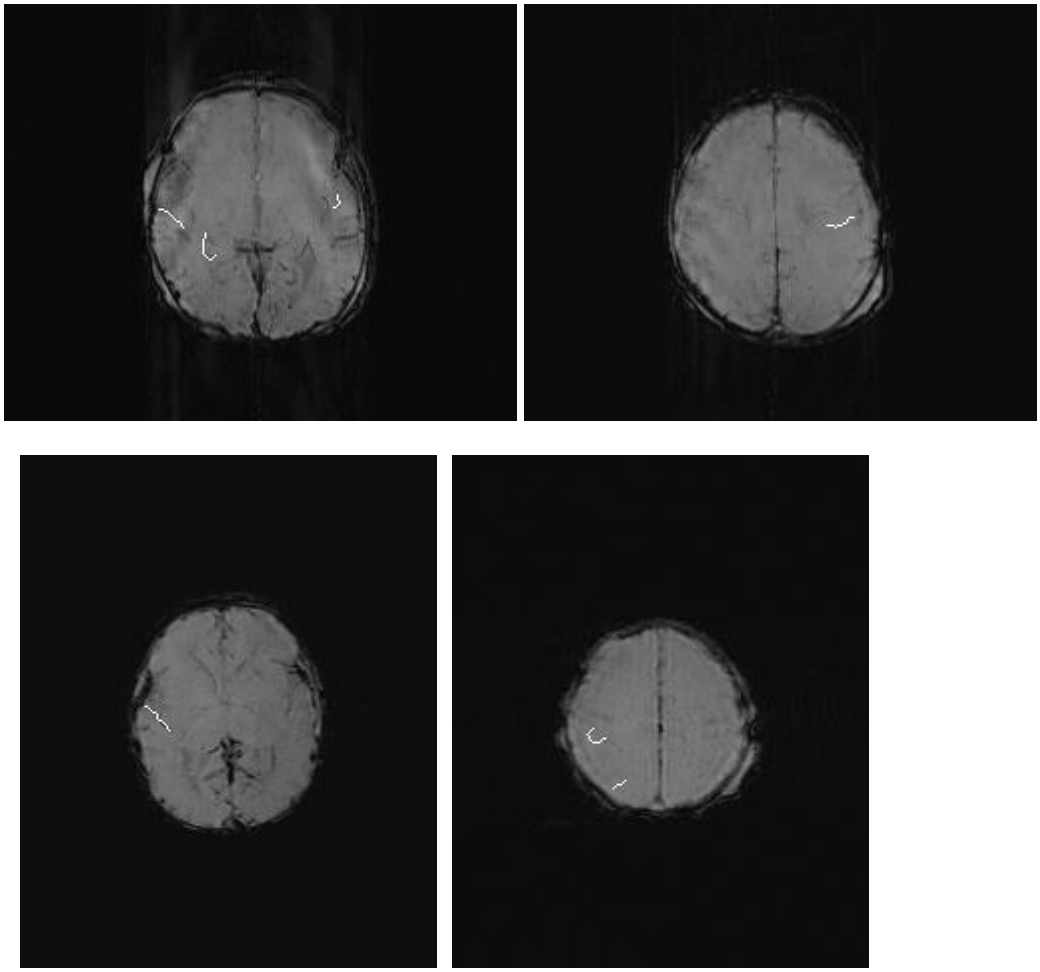


Figure 5. 7 The location of abnormal veins in patients with delayed motor scores (<85).

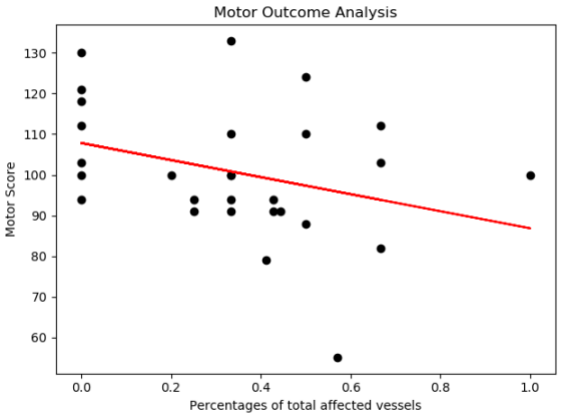
In addition to the motor areas, we measure the ratios of abnormal vessels detected by ***Algorithm 1*** in other regions with T of 6.17 and q of 4.1% and T of 6.08 and q of 4.7%, respectively, such as the frontal lobe, temporal lobe, parietal lobe and occipital lobe. The ratios of abnormal vessels in the temporal lobe cannot be counted and plotted due to only two abnormal vessels being detected in this region. For each lobe, the same three ratios of abnormal vessels are calculated as follows:

R_{mf1}, R_{mp1} or R_{mo1}: The ratio of the number of abnormal vessels in the frontal, parietal or occipital lobe to the total number of abnormal vessels in the brain.

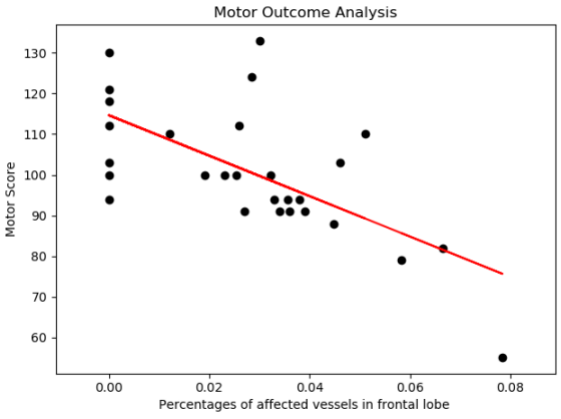
R_{mf2}, R_{mp2} or R_{mo2}: The ratio of the number of abnormal vessels in the frontal, parietal or occipital lobe to the total number of vessels in the frontal, parietal or occipital lobe.

R_{mf3}, R_{mp3} or R_{mo3}: The ratio of the length of abnormal vessels in the frontal, parietal or occipital lobe to the length of total vessels in the frontal, parietal or occipital lobe.

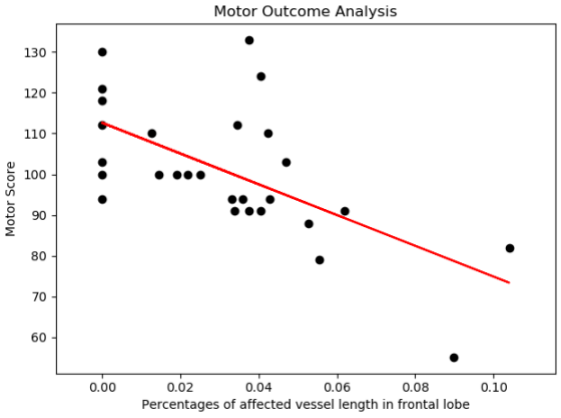
Figure 5.8 shows the ratios of abnormal vessels in the frontal, parietal and occipital lobe, and Table 5.8 shows the correlation between the motor score and $R_{mf1-mf3}$, $R_{mp1-mp3}$ and $R_{mo1-mo3}$ with T of 6.17 and q of 4.1%. Here, all abnormal vessel counts and the ratio measurements are achieved automatically by registering the SWI images, including the detected abnormal vessels, with a brain template containing all brain lobes.


Motor Outcome Analysis
Motor Score
Percentages of total affected vessels

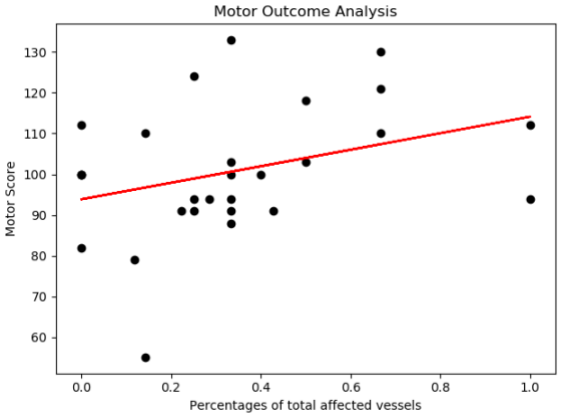
a.


Motor Outcome Analysis
Motor Score
Percentages of affected vessels in frontal lobe

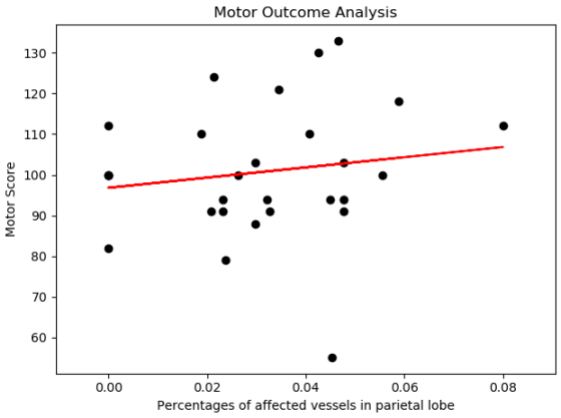
b.


Motor Outcome Analysis
Motor Score
Percentages of affected vessel length in frontal lobe

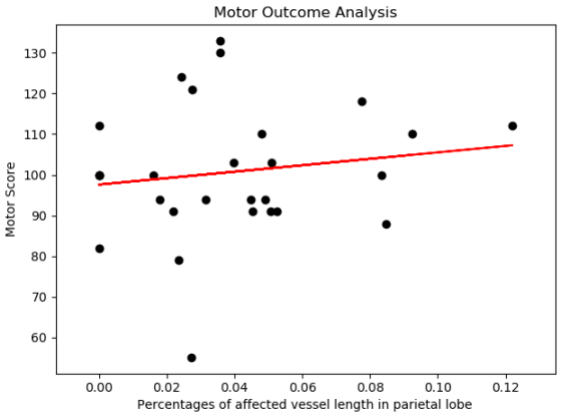
c.


Motor Outcome Analysis
Motor Score
Percentages of total affected vessels

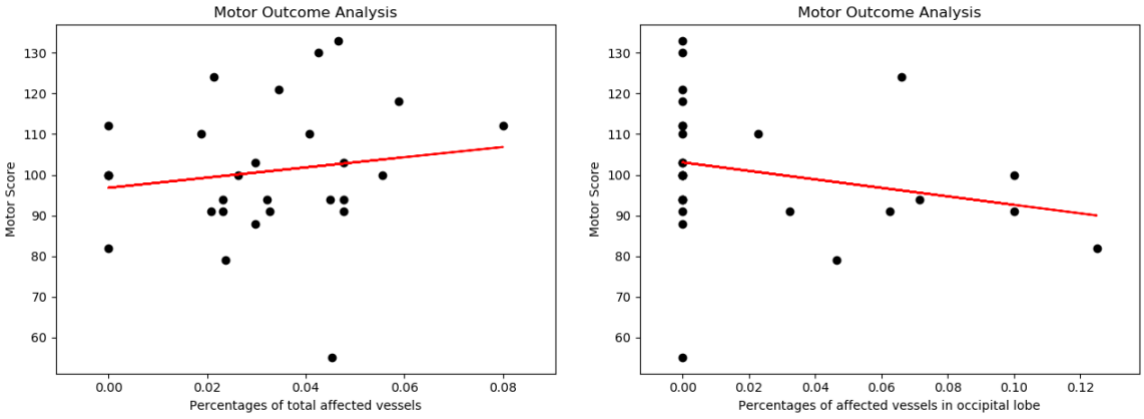
d.


Motor Outcome Analysis
Motor Score
Percentages of affected vessels in parietal lobe

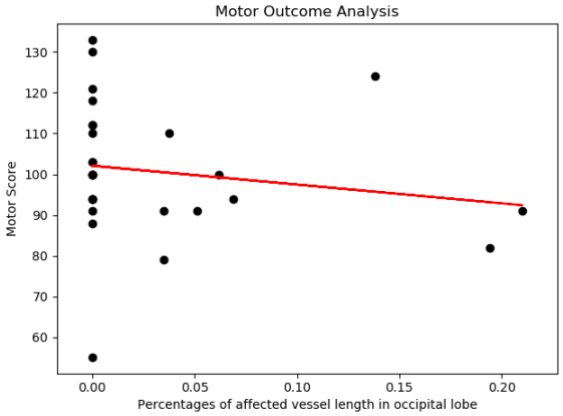
e.


Motor Outcome Analysis
Motor Score
Percentages of affected vessel length in parietal lobe

f.



g. h.



i.

Figure 5. 8 The ratio of the number of abnormal vessels in the frontal (a), parietal (d) and occipital (g) lobe to the total number of affected; The ratio of the number of abnormal vessels in the frontal (b), parietal (e) and occipital (h) lobe to the total number of vessels in the frontal, parietal and occipital lobe; The ratio of the length of abnormal vessels in the frontal (c), parietal (f) and occipital (i) lobe to the total length of vessels in the frontal, parietal and occipital lobe. All plots in this figure are with T of 6.17 and q of 4.1%.

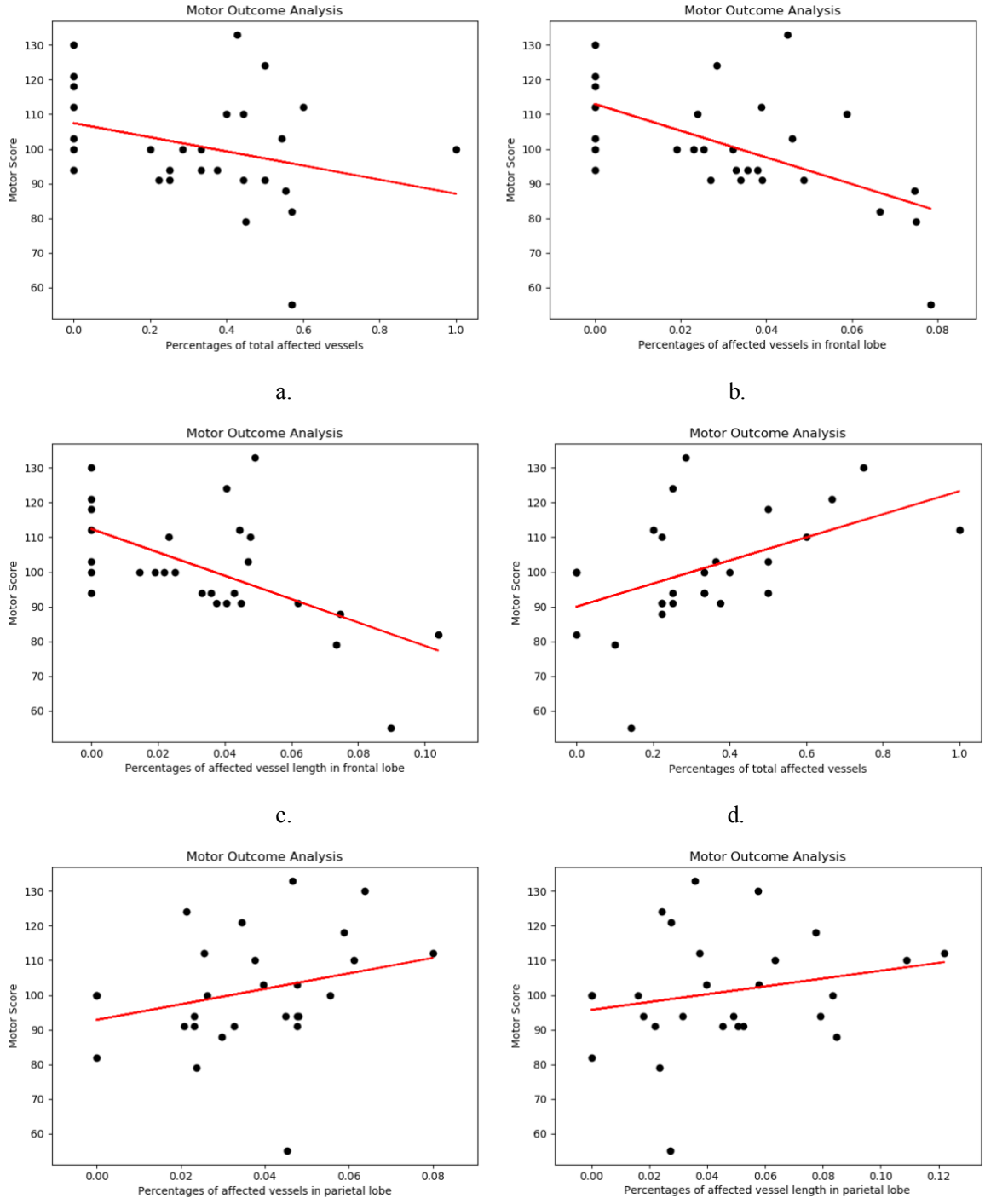
Table 5. 8 Correlation of ratios with motor scores (T of 6.17 and q of 4.1%)

Frontal lobe		
Ratio measurement	Motor score	*p*-value
R_{mf1}	-0.33	0.088
R_{mf2}	**-0.65**	0.00017
R_{mf3}	-0.62	0.0004
Parietal lobe		
Ratio measurement	Motor score	
R_{mp1}	0.34	0.079
R_{mp2}	0.15	0.4347
R_{mp3}	0.15	0.4482
Occipital lobe		
Ratio measurement	Motor score	
R_{mo1}	-0.22	0.2670
R_{mo2}	-0.24	0.2149
R_{mo3}	-0.16	0.4039

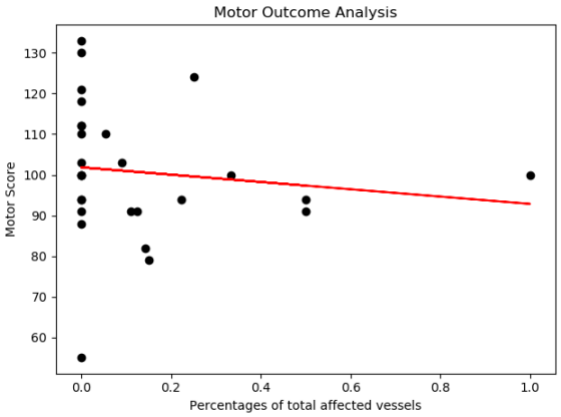
The scatterplots of these ratios of the frontal, parietal and occipital lobe with respect to motor scores are shown in Figure 5.8. Table 5.8 shows the correlation between the motor score and ratios of the frontal, parietal and occipital lobe with T of 6.17 and q of 4.1%. It is noted from Table 5.8 that the correlation between the motor score and the percentage of abnormal vessels in the frontal lobe to the total number of vessels in the frontal lobe is -0.65 (p-value<0.0002), which is the highest among other measurements but is still lower than the correlation in motor area. As observed from this figure, the ratio of abnormal vessels in the frontal lobe has a strong

correlation with the motor score (Figure 5.8(b) and (c)), while the correlation between the motor score and the ratio of abnormal vessels in the parietal and occipital lobe is not significant, as shown in Figure 5.8 (d – i).

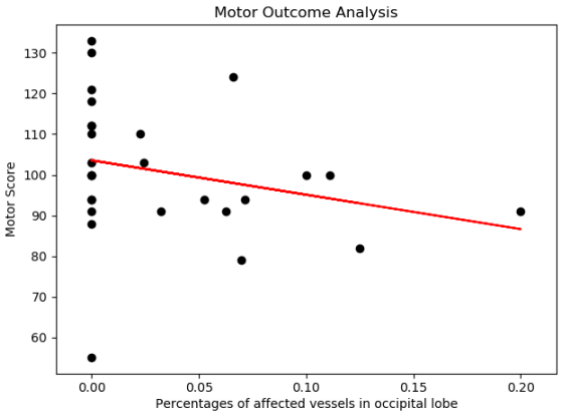
Figure 5.9 shows the ratios of abnormal vessels in the frontal, parietal and occipital lobe and Table 5.9 shows the correlation between the motor score and $R_{mf1-mf3}$, R_{mp1-p3} and $R_{mo1-mo3}$ with T of 6.08 and q of 4.7%.



e.

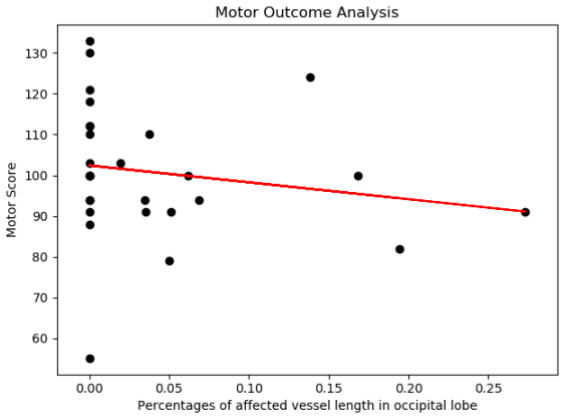


f.



g.

h.



i.

Figure 5. 9 The ratio of the number of abnormal vessels in the frontal (a), parietal (d) and occipital (g) lobe to the total number of affected; The ratio of the number of abnormal vessels in the frontal (b), parietal (e) and occipital (h) lobe to the total number of vessels in the frontal, parietal and occipital lobe; The ratio of the length of abnormal vessels in the frontal (c), parietal (f) and occipital (i) lobe to the total length of vessels in the frontal, parietal and occipital lobe. All plots in this figure are with T of 6.08 and q of 4.7%.

Table 5. 9 Correlation of ratios with motor scores (T of 6.08 and q of 4.7%)

Frontal lobe		
Ratio measurement	Motor score	p-value
R_{mf1}	-0.31	0.1096
R_{mf2}	-0.56	0.00133
R_{mf3}	-0.59	0.0010
Parietal lobe		
Ratio measurement	Motor score	
R_{mp1}	0.48	0.0094
R_{mp2}	0.28	0.1472
R_{mp3}	0.22	0.2542
Occipital lobe		
Ratio measurement	Motor score	
R_{mo1}	-0.12	0.5296
R_{mo2}	-0.26	0.1767
R_{mo3}	-0.17	0.3681

Figure 5.9 shows scatter plots of the ratios of the frontal, parietal and occipital lobe relative to motor scores with T of 6.08 and q of 4.7%. Table 5.5 shows the correlations between motor scores and the ratios of the frontal, parietal and occipital lobes, respectively. As can be seen in Table 5.9, the correlation between motor score and the ratio of the length of abnormal vessels in the frontal lobe to the length of total vessels in the frontal lobe was -0.59 (p-value<0.001), which was the highest among other measurement. Comparing with the ratios of the motor areas,

the percentage of the abnormal vessels in motor areas to the total number of vessels in the motor area possesses the best correlation with the motor score, which is -0.76 (p-value < 0.0002).

Figures 5.8 and 5.9 show the results of the percentage of abnormal vessels in the same cortex for HIE patients with motor outcome by using different optimal thresholds (T=6.17 and T=6.08). We explore the ratios of abnormal vessels in all brain lobes. Since there is almost no number of abnormal vessels detected from ***Algorithm 1*** in the temporal lobes of patients with motor outcome, we only calculate the ratios of the number of abnormal vessels in frontal lobe, parietal lobe, and occipital lobe. In all scatter plots, patients with a ratio of zero are shown due to the absence of abnormal vessels detected in the frontal lobe, parietal lobe, or occipital lobe. The comparison shows that at a threshold T=6.17, the correlation between motor scores and the percentage of abnormal frontal vessels and the total number of frontal vessels is the highest of all measurements, which is -0.65 (p-value<0.0002). The scatter plots show that the correlation between motor scores and the ratios of the number of abnormal vessels in the frontal lobe at a threshold of 6.17 is stronger than that at a threshold of 6.08. The ratios of abnormal vessels in the parietal and occipital lobes using both thresholds (T=6.17 and T=6.08) are not correlate significantly correlated with motor scores (p-value >0.005).

5.3.3. Affected brain regions associated with cognitive development

The vessels of 29 infants of which 25 infants had normal cognitive outcome (cognitive scores>85), 3 infants with mild cognitive delay (cognitive scores between 77.5 and 85) and one severe cognitive delay (cognitive scores <77.5), are optimally partitioned by our aforementioned algorithm proposed in this chapter with a threshold T of 7.09 and a percentage q of 3.9%. In the training stage, ***Algorithm 1*** achieves 100% accuracy for patients that have normal cognitive score and 75% accuracy for patients that have mild and severe cognitive delay, and the total accuracy is therefore 96.55%. We also use the *kNN* classifier for fine-tuning as proposed in section 5.3.1 and improve the accuracy of 100% in the training stage as observed in Figure 5.1(c). As observed from this figure, there are two thresholds: 7.06 and 7.13 for which the classification accuracy reaches 100%.

Before measuring the correlations between our measurements with cognitive scores, we undertake an experiment to measure the performance of our algorithm using a leave-one-out strategy with unseen test data by balancing our dataset with cognitive outcomes i.e., 4 patients with delayed cognitive scores and 4 patients with normal cognitive score randomly selected

from 25 normal cognitive score group. In our dataset, there are four patients with delayed cognitive scores. For balanced data, we randomly selected four datasets ten times from the patients with normal cognitive. The final accuracy for classifying patients with delayed cognitive score and patients with normal cognitive score is computed as, 0.65 ± 0.098 (mean ± standard deviation) by repeating the experiment 10 times the aforementioned experiment in the above paragraph. The results are summarised in table 5.10 and are compared with the classification accuracies for Random Forest and *kNN*. As can be observed from the table, ***Algorithm 1*** has superior performance for the datasets associated with cognitive and motor outcomes.

Table 5. 10. Performance of classification in ten times

Actual patients / Test patients		cognitive score >85	cognitive score ≤ 85
1st experiment	cognitive score >85	3	1
	cognitive score ≤ 85	1	3
2nd experiment	cognitive score >85	2	1
	cognitive score ≤ 85	2	3
3rd experiment	cognitive score >85	3	1
	cognitive score ≤ 85	1	3
4th experiment	cognitive score >85	2	2
	cognitive score ≤ 85	2	2
5th experiment	cognitive score >85	2	1
	cognitive score ≤ 85	2	3
6th experiment	cognitive score >85	1	1
	cognitive score ≤ 85	3	3
7th experiment	cognitive score >85	2	1
	cognitive score ≤ 85	2	3

8th experiment	cognitive score >85	2	1
	cognitive score ≤ 85	2	3
9th experiment	cognitive score >85	2	0
	cognitive score ≤ 85	2	4
10th experiment	cognitive score >85	3	1
	cognitive score ≤ 85	1	3

Cognitive functions are interrelated, and overlap using different brain areas covering, for example, attention, memory, language, and executive functions. Here, to measure the number of abnormal vessels related to cognitive functions, three areas of the brain are considered in our analysis: the frontal lobe, temporal lobe and parietal lobe. The vessels grouped as abnormal after fine-tuning with a threshold of T=7.09 in each of the three lobes are counted (q=3.9%), and the following terms are calculated:

R_{c1}: The ratio of the number of abnormal vessels in the three lobes to the total number of abnormal vessels in the brain.

R_{c2}: The ratio of the number of abnormal vessels in the three lobes to the total number of vessels in the three lobes.

R_{c3}: The ratio of the length of abnormal vessels in the three lobes to the length of total vessels in the three lobes.

In the same way, two counting methods were used for the number of abnormal vessels detected by ***Algorithm 1*** in the three lobes. One is obtained by an observer and the other automatically by registering the SWI images. Figure 5.10 shows the scatterplots of the ratios from the two counting methods and Table 5.11 shows the correlation between the cognitive score and R_{c1}, R_{c2} and R_{c3} with T of 7.09 and q of 3.9%.

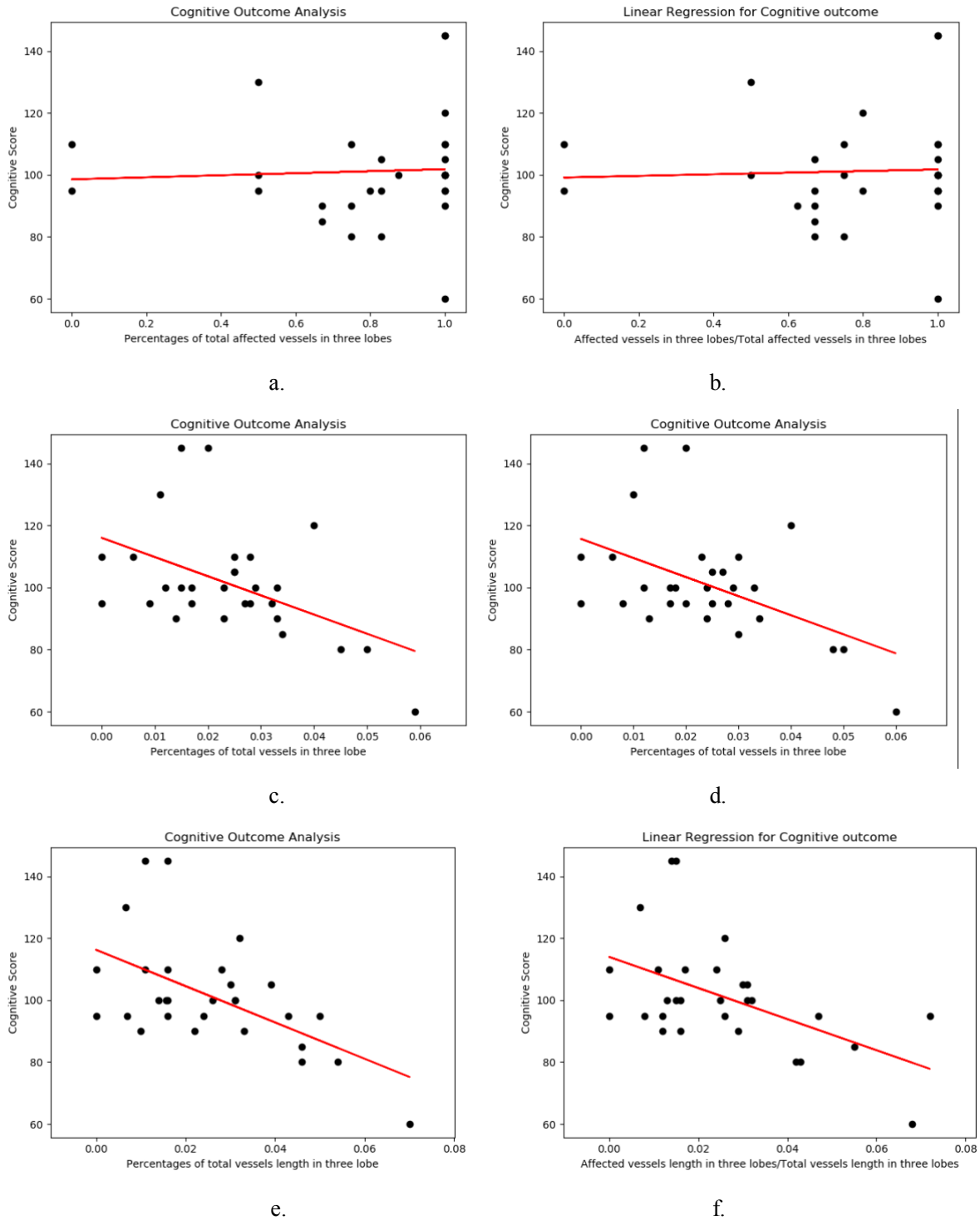


Figure 5. 10 The ratio of the number of abnormal vessels in three lobes to the total number of affected vessels in brain, a) is by the observer, b) is from the labelled three lobes; The ratio of the number of abnormal vessels in the three lobes to the total number of vessels in the three lobes, c) is by the observer, d) is from the labelled three lobes; The ratio of the length of abnormal vessels in three lobes to the total length of vessels in three lobes, e) is by the observer, f) is from the labelled three lobes. All plots in this figure are with T of 7.09 and q of 3.9%.

Table 5. 11 Correlation of ratios with cognitive scores (T of 7.09 and q of 3.9%)

By observer		
Ratio measurement	Cognitive score	*p*-value
R_{c1}	0.05	0.7916
R_{c2}	-0.49	0.007
R_{c3}	-0.56	0.0014
By registering SWI image		
Ratio measurement	Cognitive score	
R_{c1}	0.04	0.8351
R_{c2}	-0.48	0.008
R_{c3}	-0.52	0.0042

The scatter plots in Figure 5.10 depict the various relationships between the ratios of abnormal vessels in the three lobes with respect to cognitive development. The scatterplots of these ratios in the first three figures (a-c) are counted and calculated by the observer and the rest three figures (d-f) are counted and calculated by the registering SWI image. Table 5.11 shows the correlation between the cognitive score and R_{c1}, R_{c2} and R_{c3} with T of 7.09 and q of 3.9%. As observed from this figure, the measurements R_{c2} and R_{c3} demonstrate stronger correlations with the cognitive scores, but the measurement R_{c1} does not show any significant correlation with the cognitive scores at all. The correlation coefficient is -0.52 for abnormal vessels length as tabulated in Table 5.11.

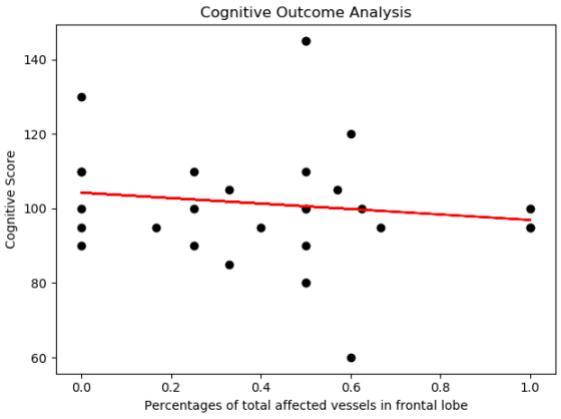
Except for the combined three lobes, we measure the ratio of abnormal vessels detected by ***Algorithm 1*** in each individual lobe, such as the frontal, temporal, parietal and occipital lobes. Here, all abnormal vessel counts and ratio measurements depend on registering SWI image with a brain template containing all brain lobes. For each lobe, same three ratios of abnormal vessels are calculated as follow:

R_{cf1}, R_{ct1}, R_{cp1} or R_{co1}: The ratio of the number of abnormal vessels in the only frontal, temporal, parietal or occipital lobe to the total number of abnormal vessels in the brain.

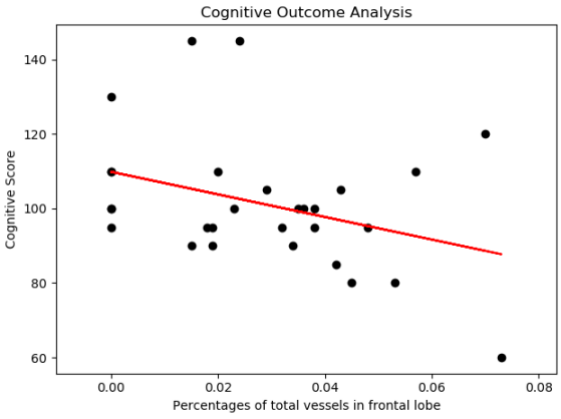
R_{cf2}, R_{ct2}, R_{cp2} or R_{co2}: The ratio of the number of abnormal vessels in the only frontal, temporal, parietal or occipital lobe to the total number of vessels in the only frontal, temporal, parietal or occipital lobe.

R_{cf3}, R_{ct3}, R_{cp3} or R_{co3}: The ratio of the length of abnormal vessels in the only frontal, temporal, parietal or occipital lobe to the length of total vessels in the only frontal, temporal, parietal or occipital lobe.

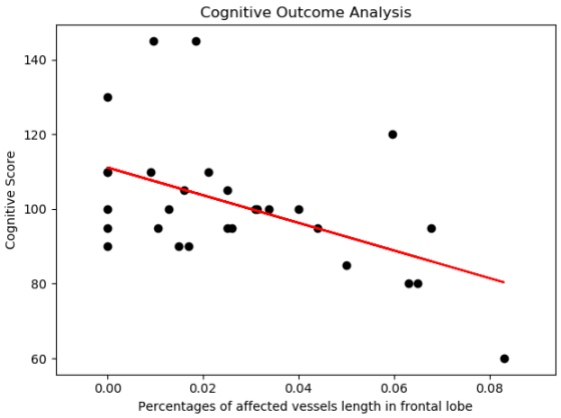
Figure 5.11 shows the ratios of abnormal vessels in only frontal, temporal, parietal and occipital lobes and Table 5.12 shows the correlation between the cognitive score and $R_{cf1-cf3}$, $R_{ct1-ct3}$, $R_{cp1-cp3}$ and $R_{co1-co3}$ with T of 7.09 and q of 3.9%.


Cognitive Outcome Analysis
Cognitive Score
Percentages of total affected vessels in frontal lobe

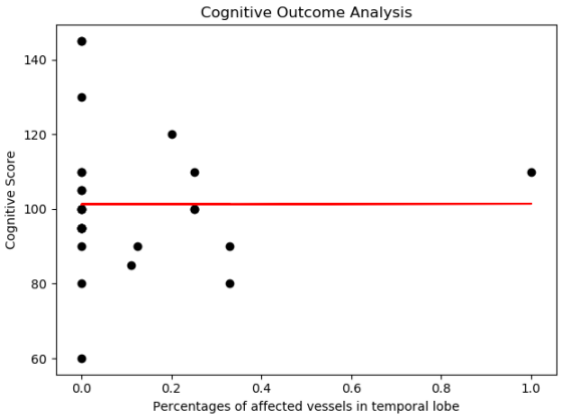
a.


Cognitive Outcome Analysis
Cognitive Score
Percentages of total vessels in frontal lobe

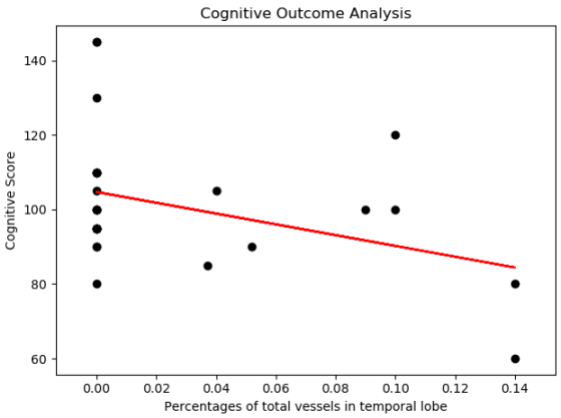
b.


Cognitive Outcome Analysis
Cognitive Score
Percentages of affected vessels length in frontal lobe

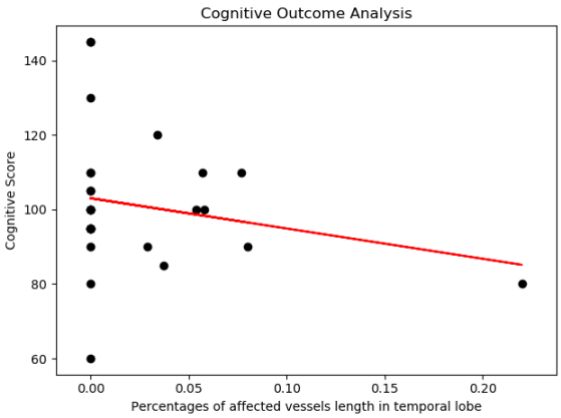
c.


Cognitive Outcome Analysis
Cognitive Score
Percentages of affected vessels in temporal lobe

d.


Cognitive Outcome Analysis
Cognitive Score
Percentages of total vessels in temporal lobe

e.


Cognitive Outcome Analysis
Cognitive Score
Percentages of affected vessels length in temporal lobe

f.

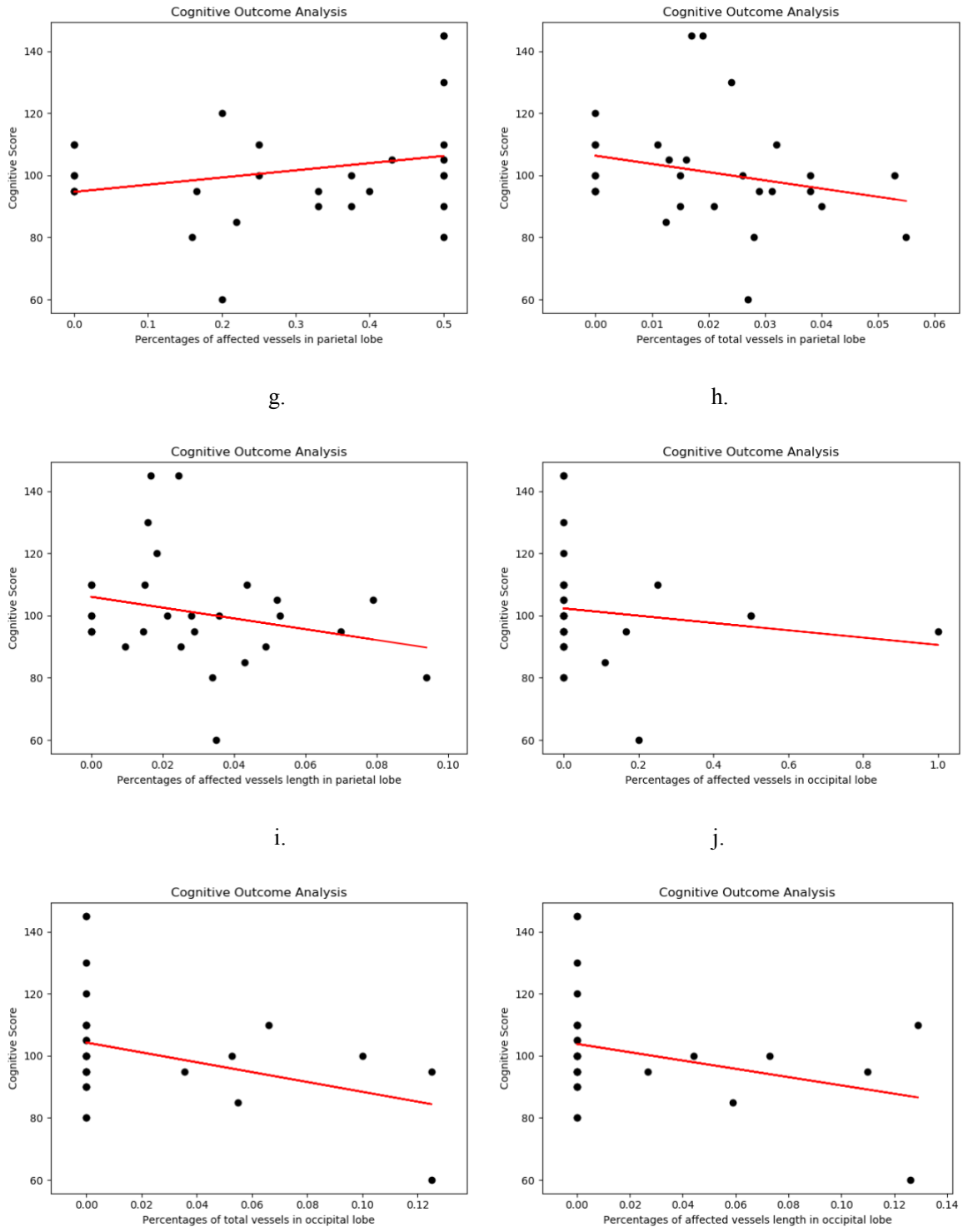


Figure 5. 11 The ratio of the number of abnormal vessels in different lobes to the total number of affected vessels in brain, a) frontal, d) temporal, g) parietal, j) occipital; The ratio of the number of abnormal vessels in the different lobes to the total number of vessels in the different lobes, b) frontal, e) temporal, h) parietal, k) occipital; The ratio of the length of abnormal vessels in different lobes to the total length of vessels in these lobes, c) frontal, f) temporal, i) parietal, l) occipital. All plots in this figure are with T of 7.09 and q of 3.9%.

Table 5. 12 Correlation of ratios with cognitive scores (T of 7.09 and q of 3.9%)

Frontal lobe		
Ratio measurement	Cognitive score	*p*-value
R_{cf1}	-0.12	0.5261
R_{cf2}	-0.36	0.0586
R_{cf3}	**-0.5**	0.006
Temporal lobe		
Ratio measurement	Cognitive score	
R_{ct1}	0.003	0.9893
R_{ct2}	-0.36	0.05
R_{ct3}	-0.21	0.2737
Parietal lobe		
Ratio measurement	Cognitive score	
R_{cp1}	0.26	0.1778
R_{cp2}	-0.24	0.2075
R_{cp3}	-0.24	0.2007
Occipital lobe		
Ratio measurement	Cognitive score	
R_{co1}	-0.15	0.4465
R_{co2}	-0.34	0.0664
R_{co3}	-0.3	0.1121

Various relationships between the ratios of abnormal vessels in the frontal, temporal, parietal and occipital and cognitive development are presented in the scatter plots in Figure 5.11. The cognitive scores for this group are not strongly correlated with our aforementioned measurements, while the correlations between the cognitive scores and the ratios of abnormal vessels length in only the frontal lobe is significant as illustrated in Figures 5.11(c). The correlation coefficient is -0.5 (p-value <0.01) for the ratios of abnormal vessels length in the frontal lobe as tabulated in Table 5.12. In summary, the length of abnormal vessels in the frontal, temporal, and parietal lobes is the best predictor for cognitive scores assessed at the age of 2 years, that the correlation coefficient is -0.52 (p-value <0.005). Our results here is inline with the fact the HIE related brain injury causing cognitive impairments is spread over a few lobes in the brain and is not located in a specific lobe. As shown in Figure 5.12, the abnormal vessels detected by ***Algorithm 1*** in patients with delayed cognitive scores are mainly located in the frontal lobe, temporal lobe and parietal lobe.

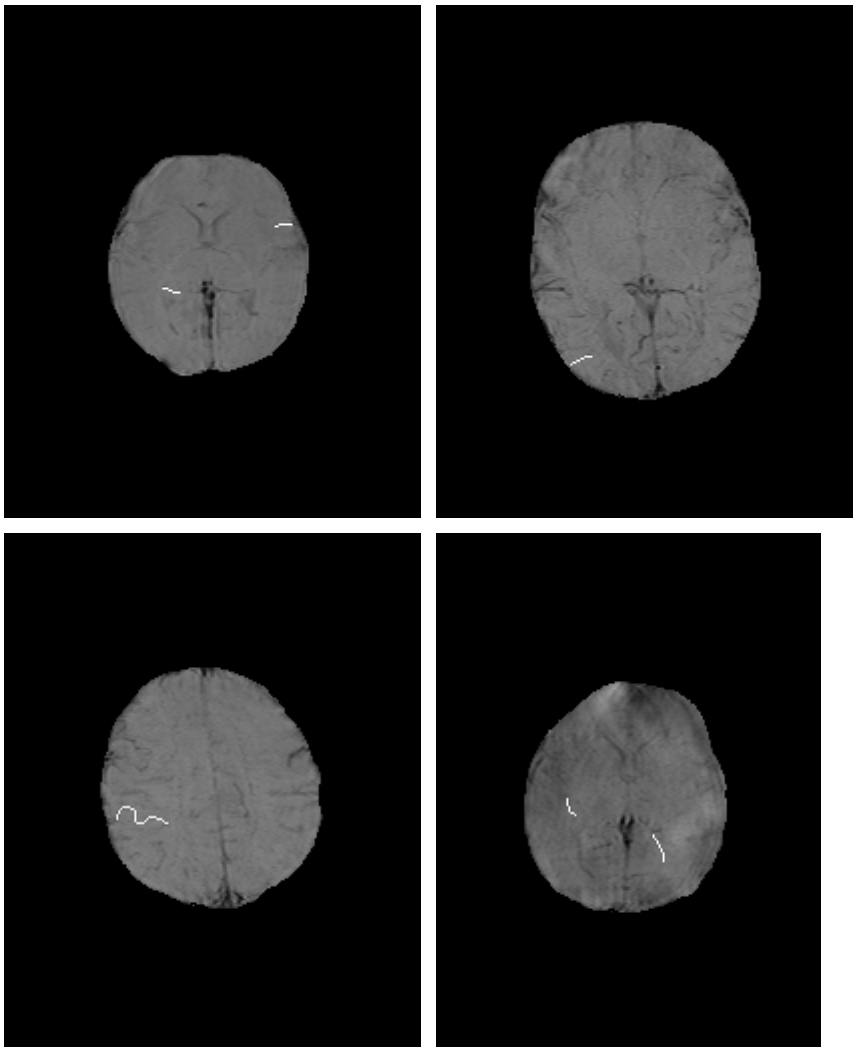
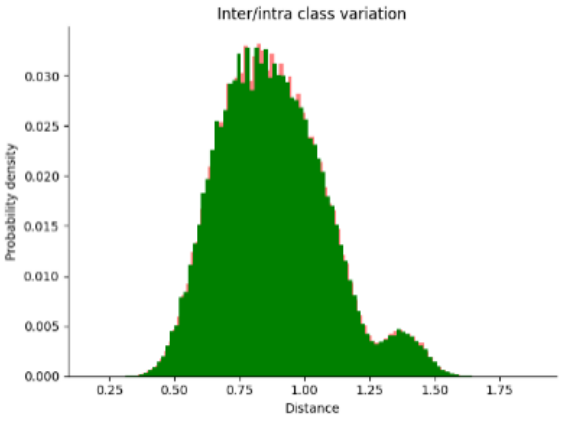


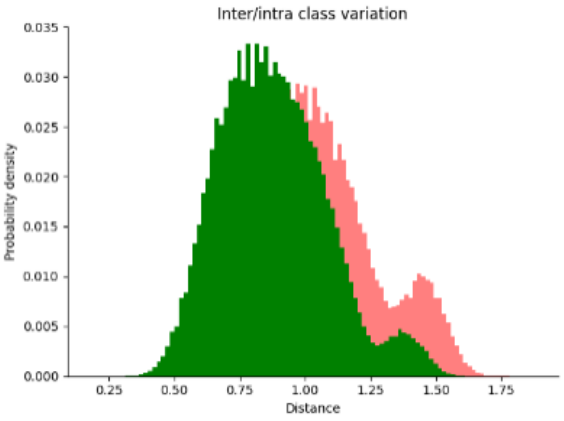
Figure 5. 12 The location of abnormal veins in patients with delayed cognitive scores (<85).

5.3.4. The inter- and intra- class variations

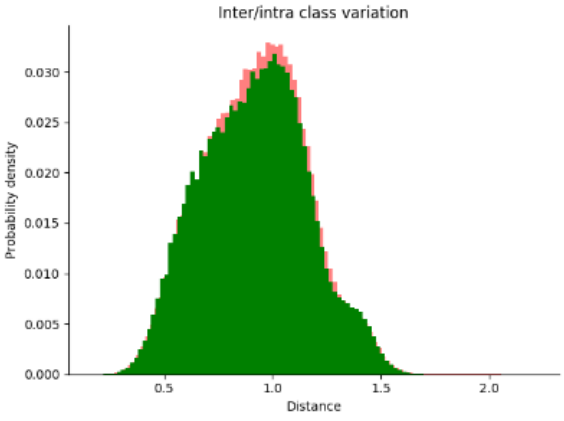
Figure 5.13 shows the inter- and intra- class variations for two normal and abnormal (vessel) groups for neurology outcome (figure 5.13 (a and b)), motor development (figure 5.13 (c, d and e)), and cognitive development (figure 5.13 (g and h)). In Figure 5.13, the green histograms illustrate the intra-class variations for vessels within the same group, while the red histograms represent the inter-class variations for vessels from the two different groups. The histograms of feature points are normalized to the area below each histogram. In figures 5.13 (a, c and f), all vessels from patients with normal neurology outcome, normal motor score and normal cognitive score are considered as normal, and all vessels from patients with abnormal neurology outcome, delayed motor cognitive developments are deemed as abnormal respectively. As observed from figures 5.13(a, c and f), the two histograms are heavily overlapped. The corresponding Receiver Operator Characteristic (ROC) curve for these three cases shown in figure 5.13(h) indicates a low performance (around 50% classification accuracy) for such a vessel classification system. Figure 5.13(b) depicts inter-class and intra-class variation histograms between abnormal and normal vessels groups after the vessels are classified as normal and abnormal for neurology outcomes by our ***Algorithm 1*** with fine-tuning proposed in this paper. Figures 5.13(d and e) present the histograms for inter- and intra- class variations between the abnormal and normal vessel groups after the classification of vessels into normal and abnormal by our ***Algorithm 1*** with thresholds T of 6.17 and T of 6.08 for motor outcomes, while figure 5.13(g) represents the abnormal and normal vessel groups classified by ***Algorithm 1*** with fine-tuning for cognitive outcomes. As observed from figures 5.13(b, d, e, and g), the inter/intra class variation histograms show a better separated features for normal and abnormal vessel groups in the feature space. As observed in figure 5.13(h), the ROC curve corresponding to the intra/inter class variations of figure 5.13(g) demonstrates a higher performance in classification (with an AUC higher than 72%). As shown in Figure 5.13(b, d, e, and g), the overlap between two histograms and the areas under the ROC curves (orange, red, yellow, and brown curves) demonstrate a better performance after ***Algorithm 1*** classifies normal and abnormal vessels in our feature space.


Inter/intra class variation
Probability density
Distance

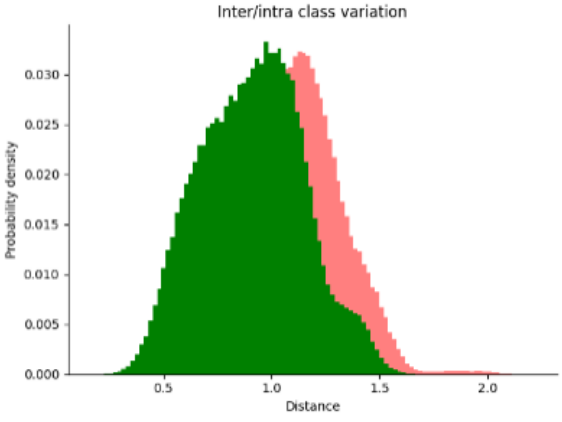
a.


Inter/intra class variation
Probability density
Distance

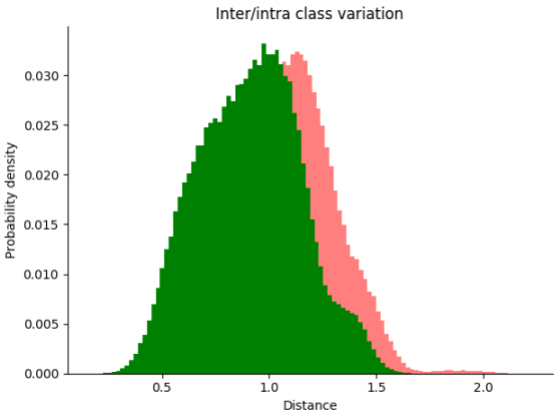
b.


Inter/intra class variation
Probability density
Distance

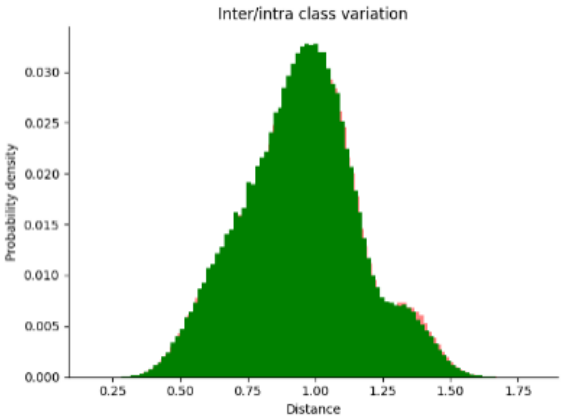
c.


Inter/intra class variation
Probability density
Distance

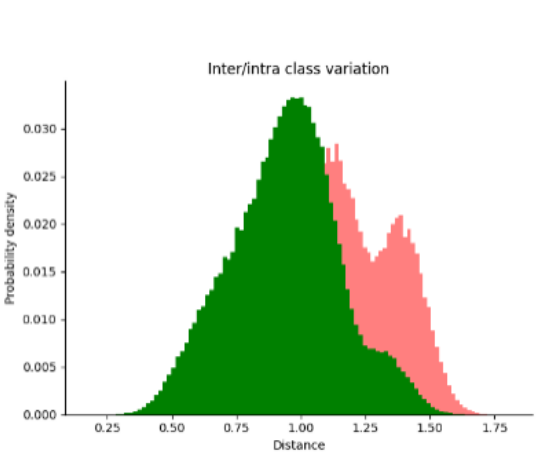
d.

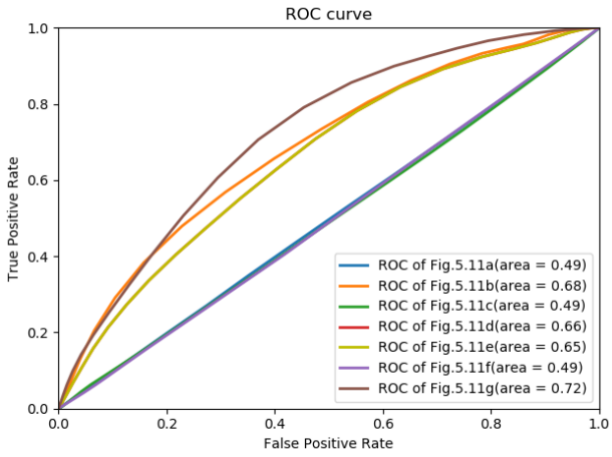

Inter/intra class variation
Probability density
Distance

e.


Inter/intra class variation
Probability density
Distance

f.





g. h.

Figure 5. 13 Intra/inter class variations of neurology outcome for: a) original abnormal group and normal group; b) abnormal group and normal group classified by our Algorithm 1 with fine-tuning; Intra/inter class variations of motor outcome for: c) original abnormal group and normal group; d) abnormal group and normal group classified by our ***Algorithm 1*** with T of 6.17; e) abnormal group and normal group classified by our ***Algorithm 1*** with T of 6.08; Intra/inter class variation of cognitive outcome for: f) original abnormal group and normal group; g) abnormal group and normal group classified by our ***Algorithm 1*** with fine-tuning; h) ROC curve of above inter- and intra- class variation plots.

5.3.5. Regression analysis for motor and cognitive development

For each of 28 patients with motor scores and 29 patients with cognitive scores, we use these different ratios, like R_{m1} to R_{m3}, and R_{c1} to R_{c3}, respectively as new features to predict motor and cognitive development scores. As our earlier work in Chapter 4, we also employ support vector regression (SVR), linear regression (LR) and polynomial regression (PR) to predict the new test. Given a database of *n* patients with ratios, $x_n \in \mathbb{R}^m$, as input and the corresponding Bayley composite scores y_n as output, the motor or cognitive score for a new patient can be predicted by employing different ratios calculated in the previous section through training the LR, PR and SVR model using these ratios (as x_n) and the associated motor or cognitive scores (as y_n). The error of regression method is measured by using a leave-one-out strategy.

To measure the performance of these regression models, the error for each test patient is calculated as the absolute difference between the true score for cognitive or motor scores and the corresponding predicted score provided by using trained different regression models with these ratios, and the mean relative error (MRE) as defined below:

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|p_i - t_i|}{t_i}$$

where n is the number of patients in our database, p_i is the regression model prediction value, t_i is the true value (ground truth). A mean and standard deviation are calculated for both absolute errors and MRE.

All ratios used for regressions are generated by registering SWI images in order to avoid human errors. The mean relative errors measured by LR model, PR model and SVR model are presented in Table 5.13 and Table 5.14 to measure the performance of predicting motor scores, and in Table 5.15 to measure the performance of predicting cognitive scores. Here the MRE of SVR model is calculated using a polynomial kernel with degree = 2 to compare the results trained with the four vascular features by the same SVR model in Chapter 4. PR model only shows the lowest MRE measured by the associated degree. Due to the lack of sufficient data, the Bayley-III composite scores of infants with HIE appears to be discrete rather than continuous. This makes the standard deviation values very large.

Table 5. 1 Regression prediction errors for motor scores (T of 6.17 and q of 4.1%)

Regression		Mean error	MRE
LR model	R_{m1}	10.92±9.54	0.108±0.095
	R_{m2}	**8.94±7.43**	**0.088±0.073**
	R_{m3}	11.30±7.17	0.112±0.071
	R_{mf1}	13.22±9.82	0.13±0.097
	R_{mf2}	9.92±9.08	0.099±0.09
	R_{mf3}	10.81±8.78	0.107±0.087
	R_{mp1}	12.57±10.62	0.125±0.105
	R_{mp2}	12.83±11.13	0.127±0.11
	R_{mp3}	12.69±11.1	0.126±0.11
	R_{mo1}	12.28±11.32	0.12±0.11
	R_{mo2}	12.17±11.42	0.12±0.11
	R_{mo3}	12.46±11.8	0.123±0.12
PR model	R_{m1} (degree = 4)	9.77±6.93	0.098±0.069

	R_{m2} (degree = 4)	9.65±7.055	0.098±0.07
	R_{m3} (degree = 2)	11.39±6.12	0.11±0.061
	R_{mf1} (degree = 2)	12.19±10.61	0.12±0.105
	R_{mf2} (degree = 2)	10.77±7.18	0.106±0.071
	R_{mf3} (degree = 2)	12.37±10.33	0.123±0.102
	R_{mp1} (degree = 3)	11.73±10.67	0.116±0.106
	R_{mp2} (degree = 2)	12.61±11.57	0.125±0.114
	R_{mp3} (degree = 2)	12.79±11.39	0.127±0.113
	R_{mo1} (degree = 2)	12.56±11.67	0.124±0.116
	R_{mo2} (degree = 2)	12.86±11.72	0.127±0.116
	R_{mo3} (degree = 2)	12.88±12.12	0.127±0.12
SVR model	R_{m1}	12.53±8.26	0.124±0.082
	R_{m2}	9.86±8.73	0.097±0.086
	R_{m3}	9.84±8.41	0.097±0.079
	R_{mf1}	14.18±14.29	0.14±0.14
	R_{mf2}	9.73±9.02	0.097±0.089
	R_{mf3}	12.46±12.8	0.12±0.13
	R_{mp1}	12.11±10.73	0.12±0.106
	R_{mp2}	12.07±10.78	0.11±0.107
	R_{mp3}	11.89±10.94	0.118±0.108
	R_{mo1}	11.99±11.0	0.11±0.109
	R_{mo2}	11.10±11.35	0.11±0.112
	R_{mo3}	11.86±11.07	0.117±0.109

Table 5. 14 Regression prediction errors for motor scores (T of 6.08 and q of 4.7%)

Regression		Mean error	MRE
LR model	R_{m1}	11.85±9.31	0.117±0.09
	R_{m2}	9.63±7.67	0.97±0.076
	R_{m3}	11.5±7.36	0.114±0.073
	R_{mf1}	13.23±10.04	0.12±0.099
	R_{mf2}	10.92±9.42	0.108±0.09
	R_{mf3}	10.78±9.26	0.107±0.092
	R_{mp1}	11.86±9.53	0.117±0.094
	R_{mp2}	12.39±10.84	0.12±0.1
	R_{mp3}	12.47±10.98	0.12±0.109
	R_{mo1}	12.94±11.19	0.128±0.11
	R_{mo2}	12.39±10.97	0.123±0.108
	R_{mo3}	12.36±11.58	0.12±0.115
PR model	R_{m1} (degree = 4)	10.96±5.92	0.108±0.058
	R_{m2} (degree = 2)	11.56±8.88	0.115±0.088
	R_{m3} (degree = 2)	10.69±8.36	0.106±0.083
	R_{mf1} (degree = 2)	12.58±10.54	0.124±0.104
	R_{mf2} (degree = 2)	11.81±8.44	0.117±0.083
	R_{mf3} (degree = 2)	12.93±9.4	0.109±0.063
	R_{mp1} (degree = 3)	13.06±10.79	0.129±0.107
	R_{mp2} (degree = 2)	12.58±11.11	0.124±0.11
	R_{mp3} (degree = 2)	12.97±10.82	0.128±0.107

	R_{mo1} (degree = 2)	13.05±11.6	0.129±0.115
	R_{mo2} (degree = 2)	12.73±11.15	0.126±0.11
	R_{mo3} (degree = 2)	13.06±11.68	0.129±0.116
SVR model	R_{m1}	13.53±9.306	0.134±0.092
	R_{m2}	10.98±9.71	0.109±0.096
	R_{m3}	10.86±8.86	0.107±0.088
	R_{mf1}	14.78±16.64	0.146±0.165
	R_{mf2}	10.64±9.4	0.105±0.093
	R_{mf3}	11.56±11.97	0.11±0.118
	R_{mp1}	13.93±14.98	0.138±0.148
	R_{mp2}	14.2±10.50	0.141±0.104
	R_{mp3}	13.66±11.3	0.135±0.11
	R_{mo1}	13.72±12.88	0.136±0.127
	R_{mo2}	14.33±14.63	0.142±0.145
	R_{mo3}	13.0±12.25	0.129±0.121

Table 5. 2 Regression prediction errors for cognitive scores (T of 7.09 and q of 3.9%)

Regression		Mean error	MRE
LR model	R_{c1}	13.18±13.03	0.13±0.131
	R_{c2}	12.35±11.23	0.12±0.111
	R_{c3}	12.58±10.52	0.124±0.103
	R_{cf1}	12.77±13.21	0.126±0.13
	R_{cf2}	13.06±12.53	0.129±0.123
	R_{cf3}	11.67±11.85	0.115±0.117

	R_{ct1}	13.58±13.57	0.134±0.13
	R_{ct2}	13.59±11.61	0.134±0.114
	R_{ct3}	13.14±12.64	0.129±0.124
	R_{cp1}	12.75±12.61	0.126±0.124
	R_{cp2}	13.24±12.11	0.13±0.119
	R_{cp3}	13.33±12.49	0.13±0.12
	R_{co1}	13.84±11.5	0.136±0.114
	R_{co2}	13.88±12.48	0.137±0.123
	R_{co3}	11.32±10.42	0.111±0.103
PR model	R_{c1} (degree= 2)	13.74±13.35	0.135±0.131
	R_{c2} (degree= 2)	11.32±10.42	0.119±0.103
	R_{c3} (degree= 2)	13.158±11.98	0.13±0.118
	R_{cf1} (degree= 2)	13.15±13.49	0.129±0.133
	R_{cf2} (degree= 2)	14.15±14.43	0.139±0.142
	R_{cf3} (degree= 2)	12.07±11.34	0.119±0.112
	R_{ct1} (degree= 2)	19.66±35.67	0.19±0.35
	R_{ct2} (degree= 2)	13.42±12.05	0.132±0.119
	R_{ct3} (degree= 2)	13.73±13.34	0.135±0.131
	R_{cp1} (degree= 2)	12.9±11.97	0.127±0.118
	R_{cp2} (degree= 2)	13.23±12.81	0.13±126
	R_{cp3} (degree= 2)	13.65±12.26	0.134±0.12
	R_{co1} (degree= 2)	18.42±25.83	0.18±0.252
	R_{co2} (degree= 2)	14.48±12.01	0.143±0.118
	R_{co3} (degree= 2)	14.46±13.05	0.145±0.131

SVR model	R_{c1}	**10.25±11.84**	**0.101±0.11**
	R_{c2}	14.74±17.05	0.145±0.16
	R_{c3}	13.35±13.03	0.132±0.13
	R_{cf1}	12.31±12.86	0.121±0.127
	R_{cf2}	14.74±17.05	0.145±0.168
	R_{cf3}	10.26±11.85	0.109±0.117
	R_{ct1}	21.39±51.9	0.211±0.513
	R_{ct2}	11.95±11.94	0.118±0.117
	R_{ct3}	12.39±13.42	0.125±0.132
	R_{cp1}	12.47±12.8	0.123±0.126
	R_{cp2}	12.77±12.68	0.12±0.125
	R_{cp3}	12.41±12.88	0.122±0.127
	R_{co1}	12.08±12.98	0.119±0.128
	R_{co2}	12.95±13.22	0.128±0.13
	R_{co3}	13.08±15.69	0.136±0.155
Predicting motor scores in Chapter 4		10.98±7.67	0.109±0.067
Predicting cognitive scores in Chapter 4		11.40±13.24	0.113±0.13

As observed from Table 5.13, Table 5.14 and Table 5.15, SVR achieves a better performance in predicting motor scores than in predicting cognitive scores in terms of both mean error value of 9.84±8.41 and MRE value of 0.097±0.079 for patients with motor scores. Both of mean errors and MRE computed by using ratios obtained from ***Algorithm 1*** (p-value <0.001) outperforms the regression results presented in Chapter 4 (p-value <0.001). Also the single variable linear regression model for R_{m2} shows lower mean error value of 8.94±7.43 and MRE value of 0.088±0.073 for motor scores, and the polynomial regression (degree=2) with R_{c1}

leads to the mean error value of 11.32±10.42 and MRE value of 0.111±0.103 for cognitive scores. It is noted that the best regression result in Table 5.13 is reported in bold. For a clearer comparison, we use the error bars to present the results of the different tables as shown in figure 5.14.

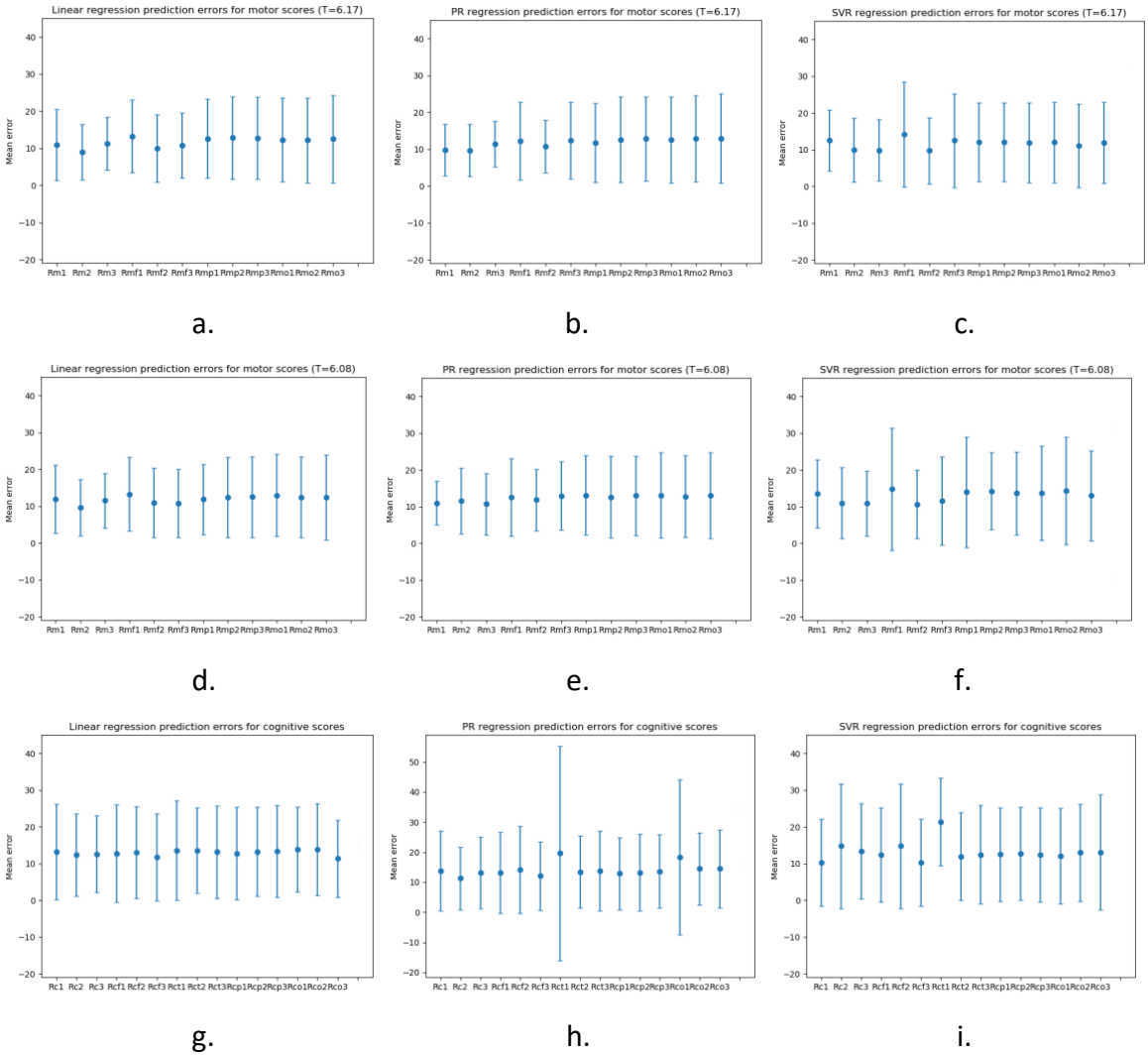


Figure 5. 14 (a -c): Regression prediction errors for motor scores (T of 6.17 and q of 4.1%); (d-f): Regression prediction errors for motor scores (T of 6.08 and q of 4.7%); (g-i): Regression prediction errors for cognitive scores (T of 7.09 and q of 3.9%).

5.4. Summary

Such a vessel classification has helped us analyse and identify brain regions responsible for abnormal outcomes based on neurological, motor and cognitive assessments. It is interesting to note that our vessel classification is based on the outcomes of clinical assessments rather than the purely image based scoring system proposed by Kitamura [45]. We have measured

the performances of our algorithm in various scenarios. By using our algorithm proposed in this paper, in brain regions implemented in the outcomes of interest, we have found some measurements showing strong correlations with outcomes determined by clinical experts who have examined the patients two years after SW images have been taken. The accuracy of early prediction of outcome at the age of two years using SW images in newborns with HIE by ***Algorithm 1*** is approximately 70%. Our experiments therefore demonstrate that ***Algorithm 1*** outperforms traditional Random Forest and *kNN* classifiers in the datasets associated with cognitive and motor outcomes. It is also competitive in the dataset associate with neurological outcome. ***Algorithm 1*** is also used here to find the brain areas affected by the HI injury.

Table 5. 16 Classification performance comparison

	Neurology outcome	Motor outcome	Cognitive outcome
Random forest	**78.67±2.58**	69.91±18.92	61.25±13.75
kNN	75.45±5.81	48.33±9.46	53.75±15.64
Algorithm 1	72.7±5.6	**75.0±13.9**	**65.0 ± 9.8**

In patients with HIE, abnormalities of blood vessels in the cortical motor areas are closely associated with abnormal neurological outcomes and motor development, with the highest correlation coefficients of -0.76. The detection of damaged blood vessels within this region in early SW images can help determine the type and severity of neurological impairment and motor development in infancy. The correlations of our measurements associated with the abnormal vessels (in the frontal, temporal and parietal lobes) with cognitive outcomes are lower. The length of vessels affected by HIE in the frontal, temporal, and parietal lobes shows the highest correlation of -0.52 with cognitive scores. Due to a lack of data, the lowest relative error of our regression model is related to motor scores, with an MRE of 0.088±0.073. Therefore, our analysis of SW images of patients with HIE provides additional information to support prognostication and identification of those who may benefit from early intervention.

Chapter 6

Hypoxic-Ischaemic Encephalopathy Prognosis Using Susceptibility Weighted Image Analysis Based on Histogram Orientation Gradient

In the previous chapters we described the ability of vascular structures on weighted magnetic resonance imaging (SWI) images to predict long-term neurodevelopmental outcome in hypoxic-ischaemic encephalopathy (HIE) survivors. And we mainly analyzed the vessel features of patients with HIE. Furthermore, we want to explore the relationship between vessel features combined with the surrounding tissue of SWI images and the neurological outcome of infants with HIE. The histogram of oriented gradients (HOG) descriptor [100] is able to effectively local features of the image. Here the clear vessel appearance provided by SWI image can be detected by the HOG descriptor. In this chapter, we use HOG operator to analyse SWI images to identify areas of the neonatal brain affected by HIE.

6.1. Introduction

As mentioned previously, magnetic resonance imaging (MRI) has become the standard for the assessment and study of neonatal HI injury and developmental abnormalities [17]. SWI image is increasingly used in clinical practice because of its sensitivity to haemorrhage and calcification [101]. As SWI images can sensitively capture the blood vessels and vascular structures in the brain, we apply the histogram orientation gradient (HOG) feature descriptor for object detection. HOG is a powerful feature extraction technique that calculates the occurrences of gradient orientation in local parts of an image. Dalal and Triggs [100] first presented a histogram of oriented gradients to identify the people in images. Subsequently, HOG feature descriptor has been used in the brain tumor detection [102] and the early diagnosis of Alzheimer's disease [103]. However, HOG has not been applied in the detection and diagnosis of the SWI image with HIE.

A method in [104] assessed the severity of ischaemic injury by quantitative analysis of deep medullary venous structures in SWI images, and the first-order texture parameters derived from SWI were employed to distinguish between infants with ischaemic injury and infants without normal brains. 38 infants with normal brain MRI and 7 neonates suffering from ischaemic were collected. A block shape was manually drawn in the white matter area on the SWI image of each infant to capture the region of interest (ROI) containing the deep medullary venous. The texture parameter values, entropy, skewness and kurtosis were obtained by calculating the grey scale value of the ROIs with deep medullary veins. Boxplots showed that the ischaemic group had significantly higher skewness than the normal group, with no significant difference in entropy or kurtosis between the two groups. The accuracy of classification for normal and ischaemic group based on skewness is 0.832. From the above study, we see that there are two major issues in SWI analysis in the context of neonatal HIE: (a) unbalanced data and (b) segmentation and extraction of different brain regions.

In the present chapter, we propose an automatic framework to detect neonatal hypoxic-ischaemic brain injury by extracting the HOG features of the brain and vessels in SWI images to analyse SWIs of HIE infants. Then, an image registration technique [98] is used to identify the brain regions by matching the SWIs with a brain template. The HOG is utilised to extract features of these brain regions. All extracted feature vectors are used in *kNN* and random forest algorithms for classification of HIE infants with neurodevelopmental outcome at the age of 24 months. We also employ the linear regression, polynomial regression, and support vector

regression (SVR) model to predict motor and cognitive outcomes by using HOG features. Figure 6.1 shows the flow chart of our method.

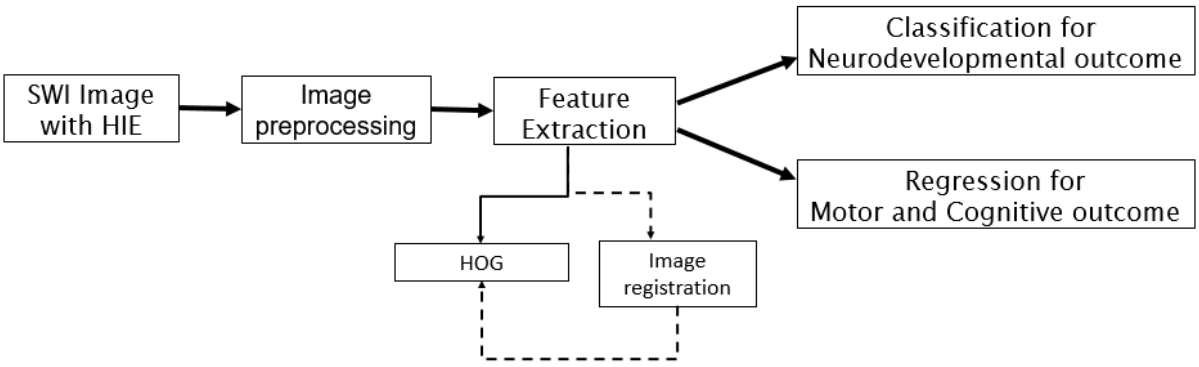


Figure 6. 1 The flow chart of our method

6.2. Methods

6.2.1 Image Pre-processing

The data and ethical certification as used in this chapter are obtained from Chapter 3, which is 42 infants with neonatal HIE. We apply an active contour model [83] for the brain segmentation to remove the skull, eyes and the background from the SWI images to reduce the noise in the images as shown in Figure 6.2(b). Before proceeding with the calculation of HOG feature vectors, we crop the SWI images into an image of 110×130 pixels (110 pixels width and 130 pixels height) to avoid the effect of redundant HOG features from the background in SWI images, as shown in Figure 6.2(c). SWI images are 2D slices obtained by scanning different locations of the brain. Since the shape of the brain is variable in a set of SWI images of one patient, we compared different SWI image sizes to obtain HOG features. The HOG feature vector obtained using the size of 110×130 pixels contains the least redundant information, i.e. the HOG vector value is 0.

6.2.2 Feature Extraction of HOG

After we crop the SWI images, HOG descriptor is used to calculate the feature vectors of SWI images. The first step of HOG is to calculate the gradient of each pixel. We denote $I(x, y)$ to be the SWI image and use a Sobel kernel of size (3×3) to obtain the horizontal and vertical gradients of each pixel. The gradient is composed of magnitude and angle from SWI images by using the following formulae for the absolute gradient:

$$M(x,y) = \sqrt{G_x^2 + G_y^2}$$

$$\theta = Arctan\frac{G_y}{G_x}$$

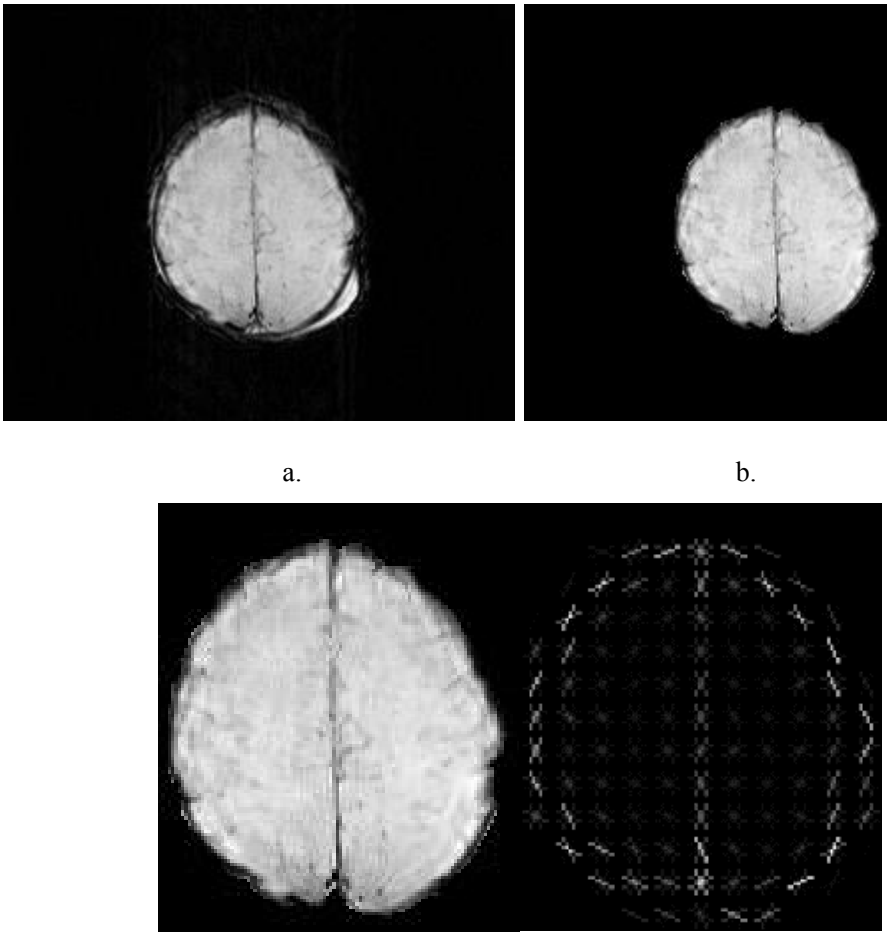
Here, G_x and G_y are the gradients of each pixel in *x* and *y* directions. *M(x, y)* denotes the magnitude and θ denotes gradient direction for the pixel. After obtaining the gradient (including magnitude and direction) of each pixel, the cropped SWI images are divided into 10×10 pixels to form a cell. For each cell, a histogram with four bins and an angle range of 45 degrees is developed. Finally, one cell is formed into a block. The histogram vector and the normalization process can be calculated as follows:

$$v = \{b_1, b_2, \ldots, b_i\}$$

$$v' = v \,/\, \sqrt{\|v\|_2^2 + \varepsilon^2}$$

where *b* is the value of each bin, ε is a small positive value used for regularization to avoid division by zero.

For each SWI image of each infant with HIE, the total length of HOG features is 4×11×13 = 572. We sum up the HOG features of each SWI image belonging to the same infant with neonatal HIE to create a feature vector, V_{whole}, describing the image, and then the feature vector for each infant is normalised. Figure 6.2(d) shows HOG of SWI image with HIE.



a. b.

c. d.

Figure 6. 2 a) Original SWI image; b) SWI image after active contour; c) Cropped SWI image; d) HOG image.

6.2.3 Image Registration

To find brain regions affected by hypoxic ischaemia, we use the atlas and image registration tools in Chapter 5 to extract HOG features from different regions in SWI images. We convert all SWI images of each infant in our dataset into 3D brain images to be registered. The strategy on ANTs registration programme for which we opt for, is to map the SWI images to the template brain images using similarity transform and to obtain the registered SWI images. We eventually consider the primary motor area, premotor area, and supplementary motor area of 3D images to explore the relationship between the HOG features of SWI images in these areas and neurology and motor outcomes at 24 months of age. To explore the relationship with cognitive outcomes frontal lobe, parietal lobe and temporal lobe of the brain as a 3D brain are examined.

As shown in Figure 6.3(b), the area in blue represents the motor area. The motor areas of the brain in SWI images are therefore selected by registering SWI images to the template brain and the motor areas are left in the 3D image with the rest of the brain being ignored (i.e. set to zero).

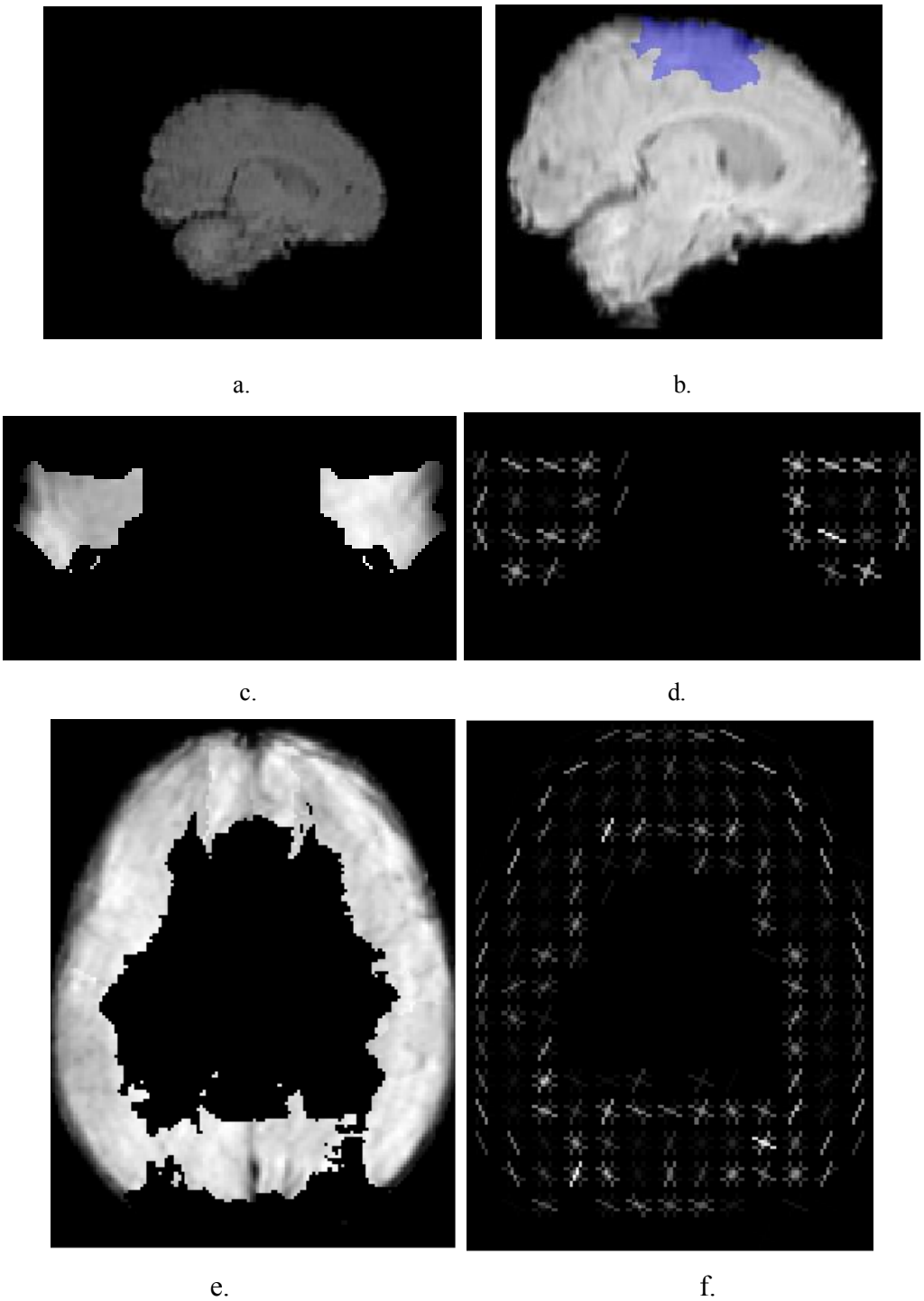


Figure 6. 3 a) Original brain image; b) Brain image after registration and motor area covered by blue; c) 2D SWI image of motor area; d) HOG of motor area; e) 2D SWI image of combined frontal lobe, parietal lobe and temporal lobe; f) HOG of cognitive area.

Finally, we consider the 3D motor region images as slices and compute HOG feature vector V_{motor} associated with the motor region. In a similar fashion, we select frontal, parietal and temporal lobes for cognitive regions of the brain by using the aforementioned registration method. Then by only considering these three lobes on the 3D SWIs and ignoring the rest of the brain, we measure the HOG feature vectors V_{lobes} associated with these three lobes.

6.3. Result of Classification

6.3.1 Classification for Neurology Outcome

For 42 infants with HIE who were assessed with neurological examination, 31 infants have normal neurology outcome, and 11 infants have abnormal neurology outcome. In this experiment, 42 infants are categorised into two groups: normal and abnormal. HOG feature vectors V_{whole} and V_{motor} are assigned to each patient in both groups for classification. Two different types of classifiers, *kNN* and random forest (RF) are used to evaluate the classification performance.

Before using the classification algorithms, we randomly select 11 patients from 31 normal patients to balance our dataset. This selection of 11 out of 31 patients with normal neurology outcome is repeated ten times, and 11 patients from each selection are sorted into classification tests with the 11 patients from the raw abnormal group. The mean and standard deviation of classification accuracies are finally calculated. Normalised HOG feature vectors V_{whole} and V_{motor} are considered as input, respectively, to feed into *kNN* and RF classifiers. Meanwhile, we apply the unnormalised HOG feature vector to classify the neurology outcome. Leave-one-out strategy is employed here. The final accuracy is reported in Table 6.1.

Table 6. 1 Classification accuracy for neurology outcome

	HOG Feature vector	*kNN*	RF
Unnormalised			
Mean(SD)	V_{whole}	57.71±8.24	67.72±9.51
Mean(SD)	V_{motor}	58.62±8.94	62.12±8.06
Normalised			
Mean(SD)	V_{whole}	58.18±7.04	68.58±7.53
Mean(SD)	V_{motor}	58.63±6.23	65.28±6.82

6.3.2 Classification for Motor Outcome

For the 28 infants with HIE who were assessed with Bayley-3 scales, 25 infants have normal motor development (scores>85), two infants have mild motor delay (scores between 77.5–85), and one has severe motor delay (score<77.5). The normal group with normal motor outcomes and the abnormal group with mild motor delays and severe motor delays are used as two classes for classification.

We employ HOG feature vectors V_{whole} and V_{motor} for each infant of two classes as training data. *kNN* and RF classifications are performed based on these feature vectors.

Likewise, balanced data based on three infants with delayed motor scores and three infants with normal motor scores randomly selected from a group of 25 infants with normal motor scores has been used for classification. By repeating the above process ten times with random selections from normal group, the mean and standard deviation of classification accuracies are calculated. Leave-one-out strategy is employed here, and final accuracy is reported in Table 6.2.

Table 6. 2 Classification accuracy for motor outcome

	HOG Feature vector	*kNN*	RF
Unnormalised			
Mean(SD)	V_{whole}	56.67±7.31	63.45±5.36
Mean(SD)	V_{motor}	62.23±9.28	70.41±8.31
Normalised			
Mean(SD)	V_{whole}	55.01±8.05	63.33±7.03
Mean(SD)	V_{motor}	61.67±15.81	71.67±11.24

6.3.3 Classification for Cognitive Outcome

29 infants, of which 25 have normal cognitive outcome (scores>85), three had mild cognitive delay (scores between 77.5–85) and one had severe cognitive delay (scores <77.5), are partitioned by two groups: the normal group with normal cognitive outcomes and the abnormal group with mild and severe cognitive delay outcomes. Here, we use HOG feature vectors V_{whole} and V_{lobes} for each infant with cognitive outcome for *kNN* and RF classifications.

Again, we utilise the balanced data, in which four infants with delayed cognitive scores and four infants with normal cognitive scores were randomly selected from the 25 infants in the normal cognitive score group to measure the performance of *kNN* and RF classification with a leave-one-out strategy. By repeating ten times the aforementioned classification, the mean and standard deviation of final accuracy is computed in Table 6.3.

Table 6. 3 Classification accuracy for cognitve outcome

	HOG Feature vector	*kNN*	RF
Unnormalised			
Mean(SD)	V_{whole}	55.61±14.14	69.17±5.94
Mean(SD)	V_{lobes}	65.49±7.73	75.28±10.04
Normalised			
Mean(SD)	V_{whole}	53.75±13.24	70.00±6.45
Mean(SD)	V_{lobes}	57.5±8.74	76.25±10.9

6.3.2 Experimental Result

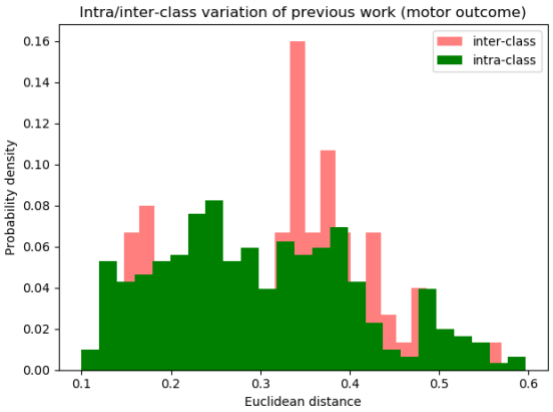
The classification analysis based on the motor outcome and cognitive outcome of infants with HIE all are a two class classification task: normal group and abnormal group. Normalised HOG feature vectors from whole brain $V_{whole}$and from different functional areas, including motor area $V_{motor}$and cognitive regions of the brain V_{lobes}, are fed to *kNN* and RF classification for comparison. All classification performances are compared to our previous work in Chapter 4, which means we extract features from motor area and frontal, temporal and parietal lobes by using the ridge detection method to classify.

Table 6.4 shows the different classification results using features by using different approaches.. In the patient group with neurology outcome, extracted features using the method presented in Chapter 4 show better classification results rather than HOG features in both whole brain area and motor area of the brain. For motor development outcome, features from ridge detection still produce better accuracy than HOG features represented by vector V_{whole} in whole brain area. But in the motor area of the brain, the classification performance of feature vector V_{motor} using HOG feature descriptor exceeds that of feature vectors discussed in chapter4. Compared with methods in Chapter 4, HOG features we propose in this chapter perform better with higher accuracies in both whole brain area and the combined frontal, temporal and parietal lobes of the brain for cognitive outcome. From Table 6.4, HOG features obtained from only the motor areas show better classification accuracies for HIE motor development outcome prognosis in

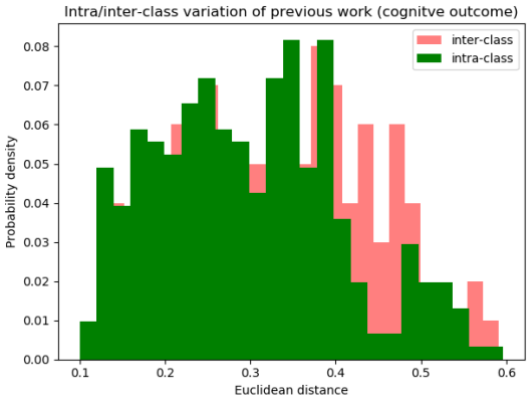
SWI images. Also the feature vectors generated by HOG descriptors in frontal, parietal and temporal lobes for cognitive regions of the brain are effective in classifying infants with cognitive outcome in SWI images.

Table 6. 4 Classification performance comparison

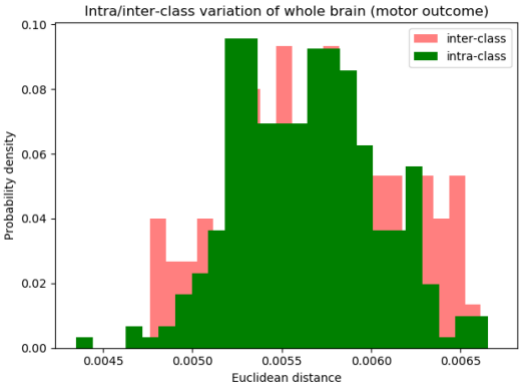
	Features	*kNN*	RF
Neurology outcome			
Whole brain	Ridge detection	75.45±5.81	**78.67±2.58**
	V_{whole}	58.18±7.04	68.58±7.53
Motor Area	Ridge detection	68.83±4.59	71.56±5.71
	V_{motor}	58.63±6.23	65.28±6.82
Motor outcome			
Whole brain	Ridge detection	48.33±9.46	69.91±18.92
	V_{whole}	55.01±8.05	63.33±7.03
Motor Area	Ridge detection	41.52±16.14	68.83±14.59
	V_{motor}	61.67±15.81	**71.67±11.24**
Cognitive outcome			
Whole brain	Ridge detection	53.75±15.64	61.25±13.75
	V_{whole}	53.75±13.24	70.00±6.45
Frontal+ temporal+ parietal lobes	Ridge detection	56.25±13.5	62.5±10.2
	V_{lobes}	57.5±8.74	**76.25±10.9**


Intra/inter-class variation of previous work (motor outcome)
inter-class
intra-class
Probability density
Euclidean distance

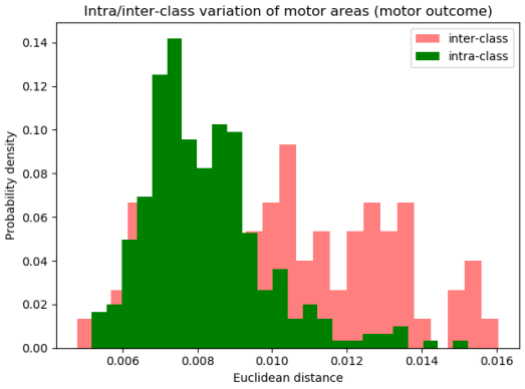
a.


Intra/inter-class variation of previous work (cognitve outcome)
inter-class
intra-class
Probability density
Euclidean distance

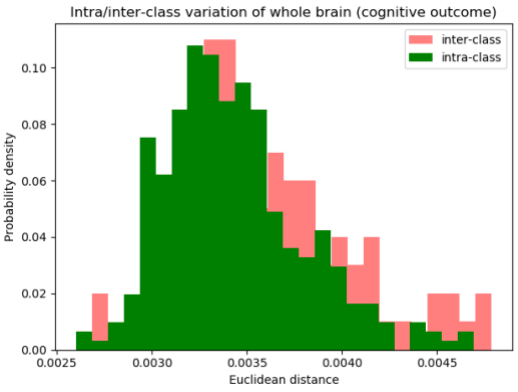
b.


Intra/inter-class variation of whole brain (motor outcome)
inter-class
intra-class
Probability density
Euclidean distance

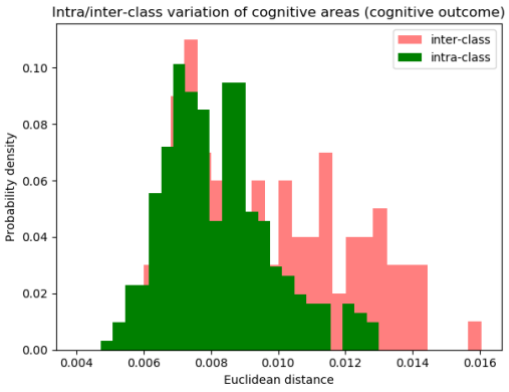
c.


Intra/inter-class variation of motor areas (motor outcome)
inter-class
intra-class
Probability density
Euclidean distance

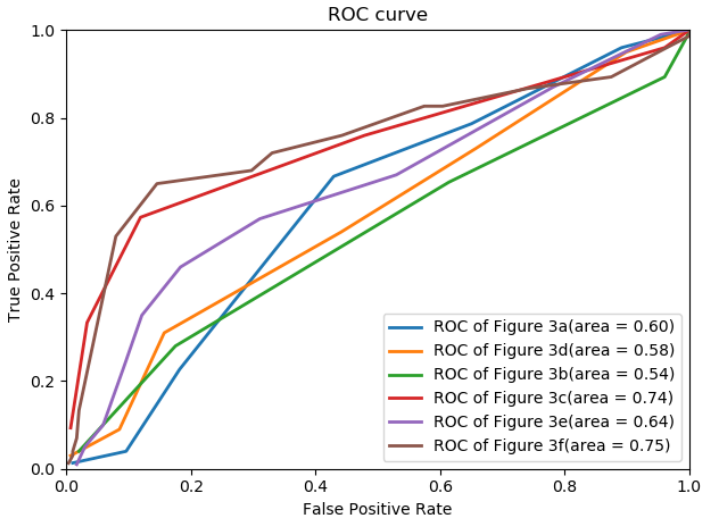
d.


Intra/inter-class variation of whole brain (cognitive outcome)
inter-class
intra-class
Probability density
Euclidean distance

e.


Intra/inter-class variation of cognitive areas (cognitive outcome)
inter-class
intra-class
Probability density
Euclidean distance

f.



g.

Figure 6. 4 Inter/intra-class variations associated with motor outcome for: (a) features obtained from Chapter 4; (c) HOG features of whole brain with motor outcome; (d) HOG features of motor area. Inter/intra-class variations associated with cognitive outcome for: (b) features obtained from Chapter 4; (e) HOG features of whole brain with cognitive outcome; (f) HOG features of cognitive regions. (g) ROC curves of above inter/intra- class variations histograms

Figure 6.4(a, c and d) show the inter-intra class variations for motor outcome and Figure 6.4(b, e and f)) for cognitive outcome. The green histograms represents the intra-class variation for infants in the same group and the red histograms show the inter-class for infants from two different groups. Histograms of features extracted using methods in Chapter 4 (figure 6.4(a and b)) and using methods in this chapter (figure 6.4(c, d, e, and f)) are normalised to the area below the histogram curve. As shown in figure 6.4(a and b), the two histograms are highly overlapped. Figure 6.4(c) presents the inter- and intra- class variation histograms of HOG feature vectors from whole brain area between normal and abnormal groups in infants with motor outcome, and furthermore figure 6.4(d) shows the inter- and intra-class variations of HOG features from motor area of brain. Figure 6.4(e) shows the inter/intra-class variation histograms of HOG features from whole brain area between normal and abnormal groups in infants with cognitive outcome, while figure 6.4(f) shows the variations of HOG features from the combined frontal, temporal and parietal lobes of the brain. The Receiver Operator Characteristic (ROC) curves corresponding to the inter- and intra- class variation histograms are shown in figure 6.4(g). As observed from figure 6.4(d and f), the overlap between two histograms and the areas under the ROC curves (brown and red curves) illustrate a better performance after using image registration to extract the HOG features of motor areas and cognitive areas.

6.4. Result of Regression

For each of 28 patients with motor composite scores and 29 patients with cognitive composite scores, HOG feature vectors are used here to predict motor and cognitive composite scores respectively. A similar work has been presented in Chapters 4 and 5. Here we continue to use support vector regression (SVR), linear regression (LR) and polynomial regression (PR) to predict a score for an unseen test data. A database with *n* patients is presented with the HOG feature vectors as input and the corresponding composite scores as output for regression algorithms. By training the SVR, LR and PR models using these features and the related motor or cognitive scores, the HOG feature vectors can be used to predict a motor or cognitive score for an unseen patient's data. The error of regression method is measured by using a leave-one-out strategy.

The mean relative error (MRE) is used to measure the performance of these regression models.

$$MRE = \frac{1}{n}\sum_{i=1}^{n} \frac{|p_i - t_i|}{t_i}$$

where *n* is the number of patients in our database, p_i is the regression model prediction value, t_i is the true value (ground truth). A mean and standard deviation are calculated for both absolute errors and MRE. For comparison with previous work, we use the same SVR model as in Chapter 4, here using a polynomial kernel of degree = 2. The PR model shows only the lowest MRE measured.

Table 6.5 presents the mean relative errors measured by the LR, PR and SVR models to measure performance in predicting motor scores, and Table 6.6 presents performance in predicting cognitive scores.

Table 6. 5 Regression prediction errors for motor scores

Regression	HOG Features	Mean error	MRE
LR model	V_{whole}	13.63±13.73	0.128±0.10
	V_{motor}	14.51±12.58	0.137±0.06
PR model (degree=2)	V_{whole}	13.60±13.75	0.135±0.136
	V_{motor}	14.46±12.49	0.139±0.073
SVR model	V_{whole}	**11.91±10.97**	**0.118±0.10**
	V_{motor}	11.91± 11.05	0.118±0.11

Table 6. 6 Regression prediction errors for cognitive scores

Regression	HOG Features	Mean error	MRE
LR model	V_{whole}	15.83±12.04	0.156±0.118
	V_{lobes}	16.52±14.47	0.163±0.143
PR model (degree=2)	V_{whole}	15.98±12.07	0.157±0.119
	V_{lobes}	16.61±14.42	0.164±0.142
SVR model	V_{whole}	**11.91±13.05**	**0.117±0.128**
	V_{lobes}	12.22±13.07	0.121±0.129
Predicting motor scores in Chapter 4		10.98±7.67	0.109±0.067
Predicting cognitive scores in Chapter 4		11.40±13.24	0.113±0.13

As presented in Table 6.5 and 6.6, SVR model achieves a better performance than other regression models. SVR perform better in predicting cognitive scores than in predicting motor scores in term of MRE value of 0.117±0.128 for patient with HOG feature, V_{whole}. Overall, the use of HOG features in predicting motor scores is better than that for the prediction of cognitive scores. The mean error and MRE of the regression results in Chapter 4 are superior to regression result by using the HOG feature vectors. It is noted that the best regression results in Table 6.5 and 6.6 is reported in bold.

6.5. Summary

In this chapter, a HOG feature extraction method for infant HIE detection in SWI images has been used. We have here employed a HOG feature descriptor to gain feature vectors to predict motor and cognitive assessments of infants with HIE at the aged of 24 months by classifying HIE SWI images along with kNN and random forest classifiers into normal and abnormal groups based on neurological,. In addition, HOG feature vectors of some brain regions obtained from brain registration, which fit our SWI data to a brain template containing different function regions, are trained for classification, thus helping us to determine which areas of the brain are responsible for abnormal outcome. Compared to our previous work in Chapter 4, we achieve outstanding performance in the classification experimentations on HIE infants with motor outcome by using HOG features of motor areas of the brain in SWI images, 71.67±11.24, and similarly HOG features of frontal, temporal and parietal lobes of the brain show better classification performance for cognitive outcome, with 76.25±10.9 accuracy. In the future, we prefer to explore the relationship between other regions of the brain and assessment outcome at two years of age. Another interesting future work is to combine our previous method with the method presented in this paper to improve the performance of our system.

Chapter 7

Deep Learning for HIE Prognoses

In recent years, with the continuous development and advancement of medical image analyses and computer technology, deep learning, especially in relation to deep convolutional neural networks (CNNs), has rapidly become an indispensable tool and technical approach in medical research and clinical disease diagnoses and treatment by automatically extracting image features implicitly and directly from medical image datasets [108]. When comparing our previous traditional machine learning methods (e.g. K-nearest neighbours, support vector machines and random forest) to ***Algorithm 1*** that we designed in previous chapters, it can be seen that deep learning involves task modules, such as feature learning and feature abstraction, in addition to model learning with the use of multi-layer task modules to complete the final learning task [105]. Despite its excellent performance in medical image analysis, the overfitting problem of deep learning on small and medium sized datasets and the lack of interpretability of neural network algorithms still exist [106].

This chapter will introduce several classical network models of convolutional neural networks and how they can be applied to our susceptibility weighted images (SWI) of hypoxic ischaemic encephalopathy (HIE). Our aim is to demonstrate that deep neural networks can be used for neonatal HIE neurodevelopmental outcome prognoses in SWI images and to explore the areas of the neonatal brain that are affected by hypoxic ischemic injuries. Although there are

currently studies using deep learning for vein segmentation from SWI images [109] and using deep neural networks for automated diagnoses of neonatal encephalopathy [82], there is no current study to use deep learning for HIE prognoses based on SWI images. Given the proven performance of different deep neural networks, we employ the CNN models to our datasets and calculate the gradient-weighted class activation mapping (Grad-CAM) [106] to highlight the important areas of the neonatal brain on the SWI images for predictions.

7.1. Convolutional Neural Networks

In early 1980, Kunihiko built on previous work to simulate the biological visual system, and he proposed a hierarchical, multi-layered, artificial neural network known as 'neurocognition' to handle handwritten character recognition [110]. Subsequently, LeCun et al. [111] proposed a gradient learning-based convolutional neural network algorithm in 1998 (see Figure 7.1) and used it successfully for handwritten numeric character recognition. Hinton et al. used Alex-Net [112] to win the ImageNet image classification competition [113]. Since then, convolutional neural networks have shown excellent performance in a wide range of computer vision tasks, such as image classification, image semantic segmentation, image retrieval and object detection.

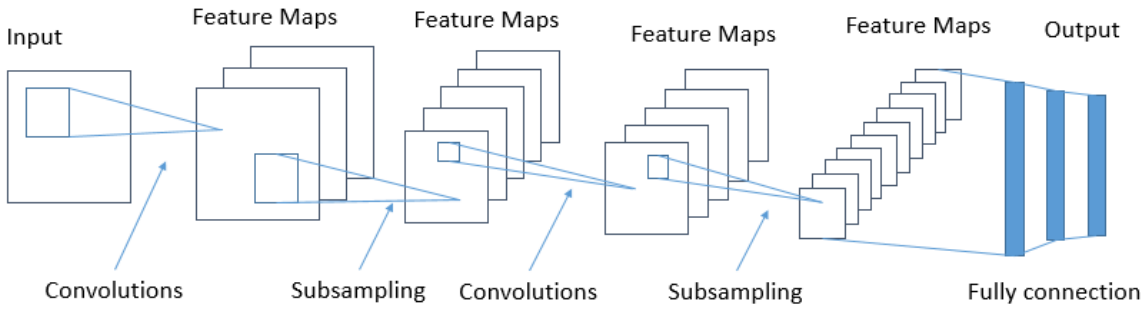
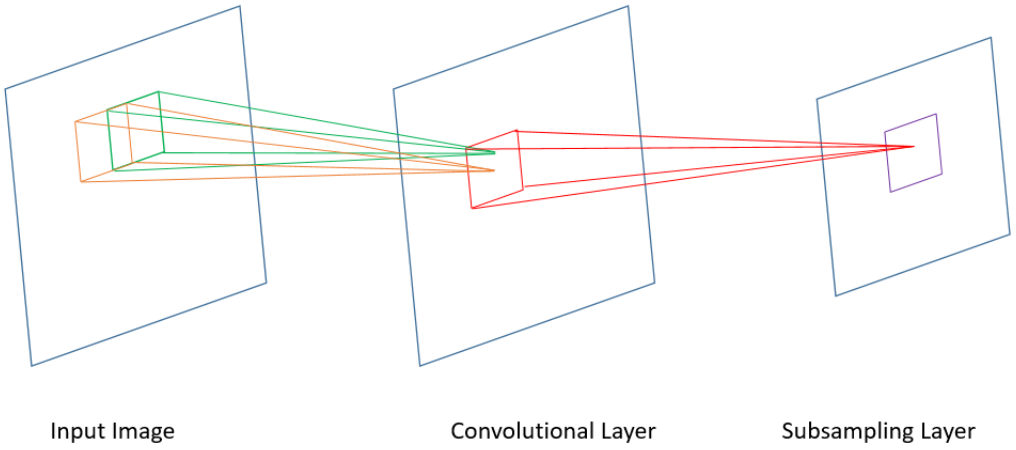


Figure 7. 1 Architecture of LeNet

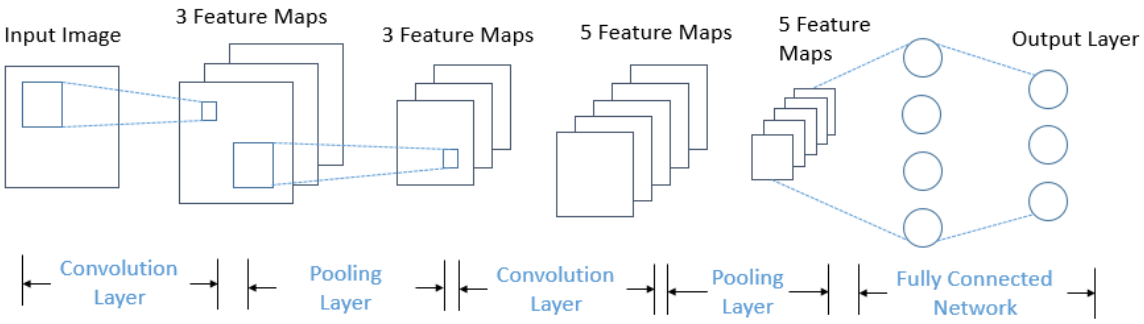
7.1.1. The Basic Structure and Components

CNNs are a type of artificial neural network [108], and they are usually composed of several alternating convolutional and pooling layers with a fully connected layer at the end (see Figure 7.2). Each of the convolutional layers has a fixed input size, and the CNNs achieve shift, scale and distortion invariance through local receptive fields, shared weights and sub-sampling, greatly reducing the degrees of freedom of the model [114].

In general, convolutional neural networks extract high-level semantic information from the original data input layer by stacking layers of operations, such as convolution, pooling and non-linear activation function mapping, and this is known as 'feed-forward'. The different types of operations are normally referred to as 'layers' in a convolutional neural network, and convolution operations correspond to 'convolutional layers', pooling operations correspond to 'pooling layers' and so on. Finally, the final layer of a convolutional neural network formalises its target task (classification, regression, etc.) as an objective function. By calculating the error or loss between the predicted value and the true value, the error or loss is passed backwards from the last layer to update the parameters of each layer by using the backpropagation algorithm. This process is repeated after the parameters have been updated until the network model converges, thus achieving the purpose of model training.



a.



b.

Figure 7. 2 a) The convolutional layer; b) The structure of convolution neural networks

The architecture of a typical CNNs is shown in Figure 7.2. In recent years, with the development of computer technology, more and more variant network models are created for

different image processing tasks [114], such as VGG-Nets [115], network-in-network [116], GoogLeNet (Inception) [117], residual networks [118] and squeeze-and-excitation networks [119].

7.2. Residual Networks

For the design of deeper convolutional neural networks, there is always a problem in that as the depth of the networks increase, the training error of the training data increases rather than decreases [120].

To address this degradation problem, He et al. proposed residual learning that creates a direct link from input to output, the residual module [118]. For a stacked layer structure, its learned features are denoted as $H(x)$ when the input is x, and the optimisation objective of the network changes from the original fitted output $H(x)$ to the difference between the output and the input ($F(x) = H(x) - x$). The formulation of $F(x) + x$ can be realised by feed-forward neural networks with skip connections (see Figure 7.3). The residual module of ResNet [118] is defined by the following formula:

$$y = F(x, \{W_i\}) + x$$

A deformation of the above formula is as follows:

$$F(x, \{W_i\}) = y - x$$

With x and y as the input and output, the function $F(x)$ that the network needs to learn is the residual term $y - x$ and is known as the 'residual function'. There are two branches in residual learning: the residual function and the identity mapping to the input. These two branches are then passed through the rectified linear unit (ReLu) activation function to form the residual learning module. A network structure formed by stacking multiple residual learning modules is called the residual network.

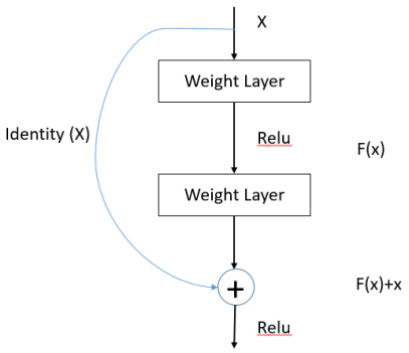


Figure 7. 3 The Inception module

The ResNet network is a reference to the VGG19 network, modified from it and adding residual units through the shortcut connection (see Figure 7.4). Additionally, the plain network shows degradation, but the ResNet network effectively solves the degradation problem.

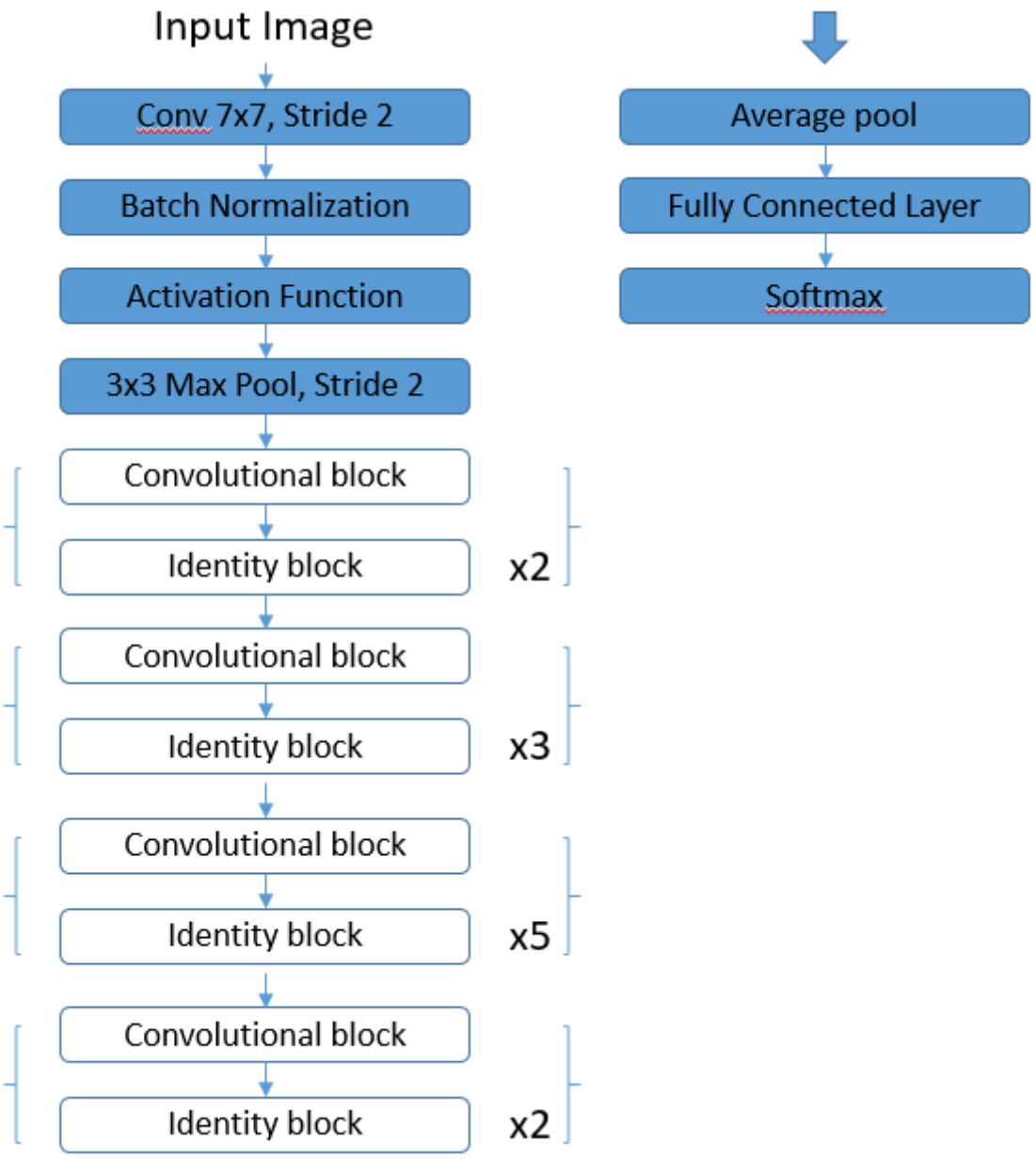


Figure 7. 4 The architecture of ResNet 50

As the depth of the network further increases, ResNet uses two types of residual modules (see Figure 7.5). The graph on the left corresponds to a simple network, and the graph on the right corresponds to a deep network known as a 'bottleneck residual block'.

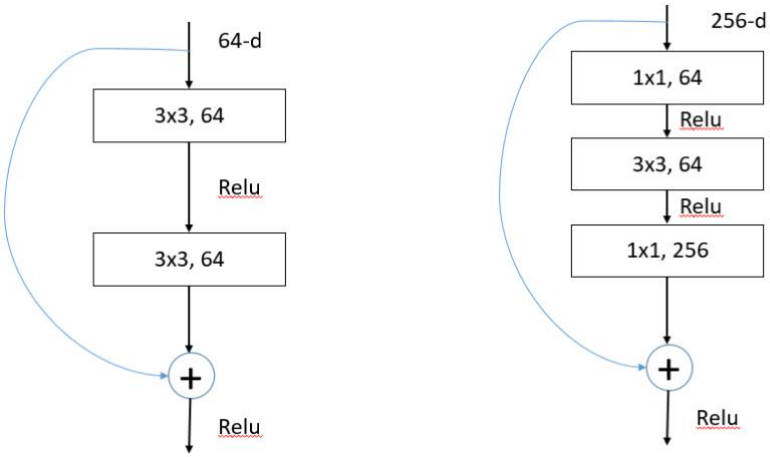


Figure 7. 5 Different residual units

7.3. Transfer Learning

As previously described, CNNs require a large amount of labelled training data, and in medical domain, the datasets usually contain a small number of samples which is not enough for the training of a CNN [121]. This issue can be effectively solved by using the transfer learning technique [120] with a CNN model that has been pre-trained on other large datasets. Transfer learning can start with a pre-trained network in the source domain, including the network parameters and hyperparameters, and train it with the small dataset. This also implies that transfer learning reduces the time spent on training data and avoids overfitting.

As we have a limited number of samples in the small dataset for training, we use residual networks with 50 layers (ResNet-50) that are pre-trained with the natural images dataset ImageNet [113], and we transfer the learning of a pre-trained model to our SWI images with HIE. ImageNet consists of more than 1.5 million images and spans over one thousand object classes. We replace the last fully connected layer and the classification layer in the pre-trained model to a set of layers that can classify our SWI images with HIE (see Figure 7.6).

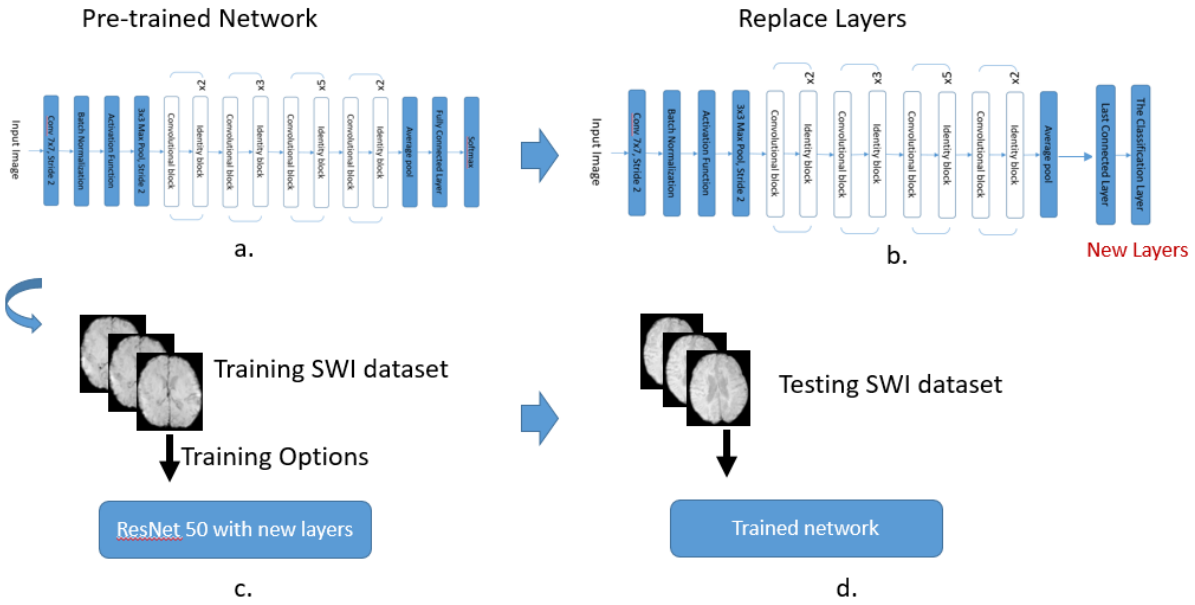


Figure 7. 6 The flow chart of transfer learning. a) A ResNet 50 network pre-trained on the ImageNet dataset is retrained. b) The last fully connected layer and the classification layer in the pre-trained model are replaced. c) This network is retrained by SWI images. d) The transfer learning network is used to predict the testing SWI images.

7.4.Visual illustrations

To visualise the regions in input images which are important in the classification accuracy, we use the gradient-weighted class activation mapping (Grad-CAM) technique [106] to enhance the interpretability of the networks, meaning that the gradient of any target is used to produce a rough localisation map in the last convolutional layer, highlighting the important regions of the images used to predict the prognosis.

To obtain the class discriminant localisation map Grad-CAM $L^c_{Grad-CAM} \in \mathbb{R}^{u \times v}$ of width u and height v for any class c, the score gradient for the class is calculated first, which is the corresponding y^c to the feature mapping A^k of the convolutional layer $\frac{\partial y^c}{\partial A^k}$. The neuron importance weights of a_k^c are as follows:

$$a_k^c = \frac{1}{Z}\sum_i \sum_i \frac{\partial y^c}{\partial A_{ij}^k}$$

These are obtained by a global average pooling of the gradients derived from backpropagation. The notation y^c means that the model determined the score of the input image as class c before entering SoftMax. The weights of a_k^c represents the partial linearisation of the neural network following A and captures the 'importance' of the feature map k for a target class c. Then, the

weighted combination of the forward activation map is implemented and obtained by the following ReLU function:

$$L^{c}_{Grad-CAM} = ReLU(\sum_{k} a^{c}_{k} A^{k})$$

The purpose of ReLU is to focus on areas that have a positive impact on the classification of the model. As shown in Figure 7.7, given an image and a target class as input, the image is propagated through the CNNs and then computed by a specific task to obtain the raw score for that class. For all classes, the gradient is set to zero, except for the gradient of the needed class that is set to one. This signal is then back-propagated to the trained convolutional feature map of interest to obtain a Grad-CAM localisation (blue heat map).

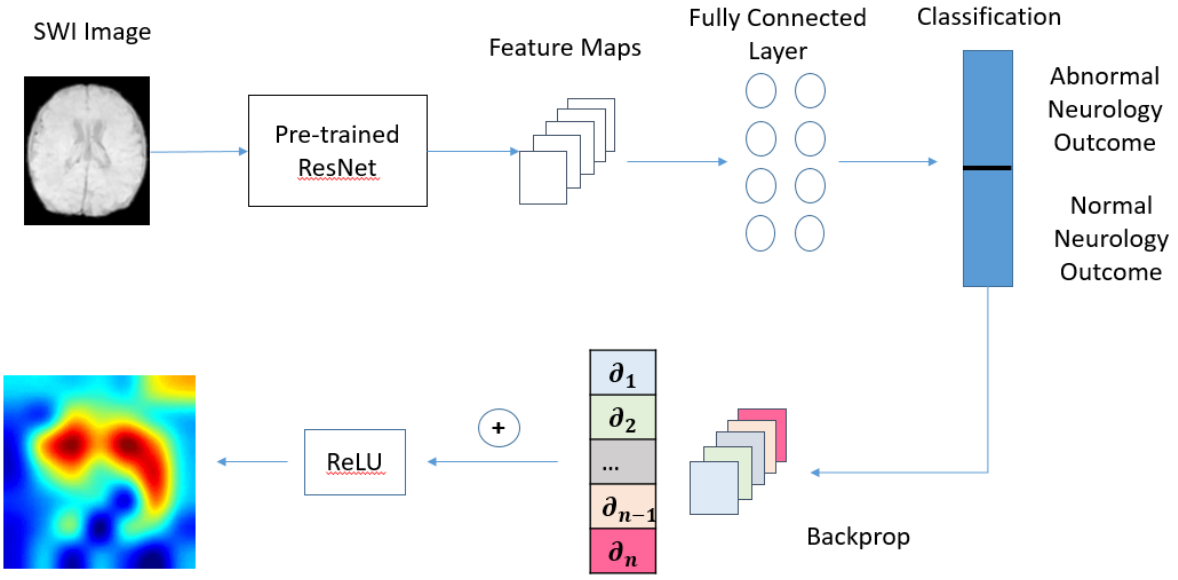


Figure 7. 7 Grad-CAM overview in our network

7.5. Deep Learning for SWI Images with HIE

7.5.1. The Dataset

The data for this study includes 42 infants with neonatal HIE born at a gestational age > 36 + 6 weeks. The ethical approval and MRI protocol has been described previously in Chapter 3. Each of the infants has a standardised neurological examination at 24 months of age by means of follow-up.

The neurological examinations consisted of the assessment of cranial nerve function, movements, posture, reflexes and muscle tone, and neurological status was then considered as either normal or abnormal (cerebral palsy). Out of the 42 infants that are neurologically assessed at two years of age, 31 of the infants (73.8%) have a normal neurological outcome, and 11 of the infants (26.2%) has an abnormal neurological outcome.

For each infant with neonatal HIE, there are around 40~60 SWI slices after undergoing an MRI scan. To ensure robust training data, we select seven SWI images in both directions from the centre slice of each set of each patient, thus avoiding the interference of some images containing noisy information (i.e. nose and eyes) in the model, meaning that the data used for the deep leaning was 15 SWI images for each patient. Figure 7.8 shows the SWI images that were chosen and the deprecated SWI images.

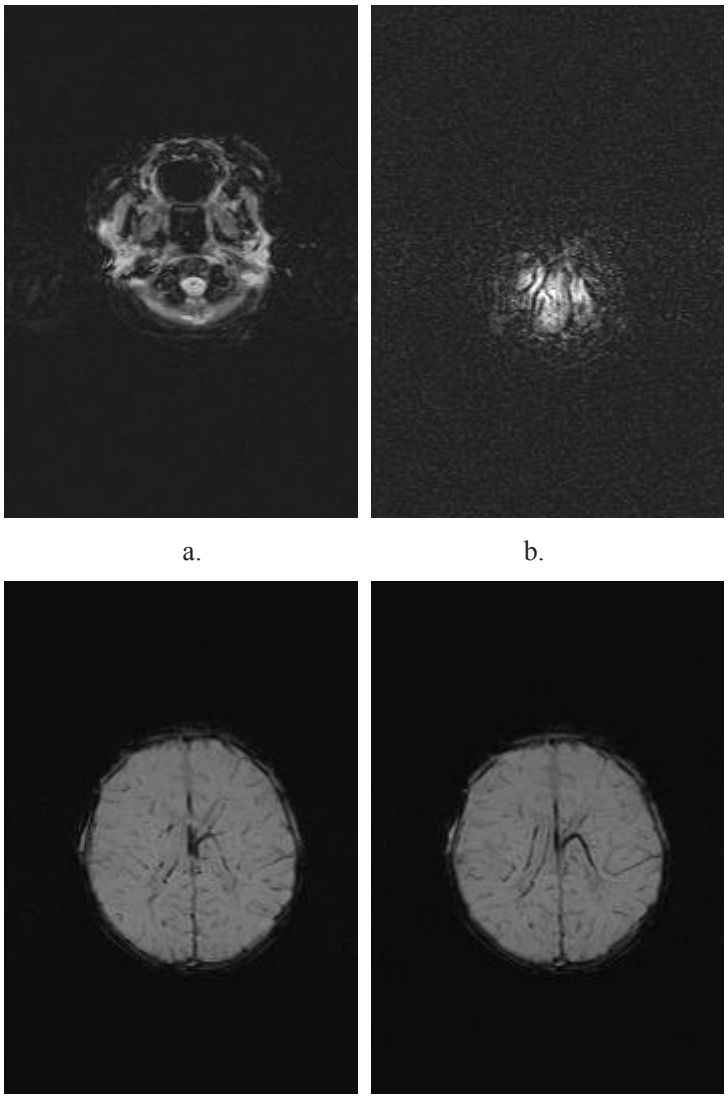
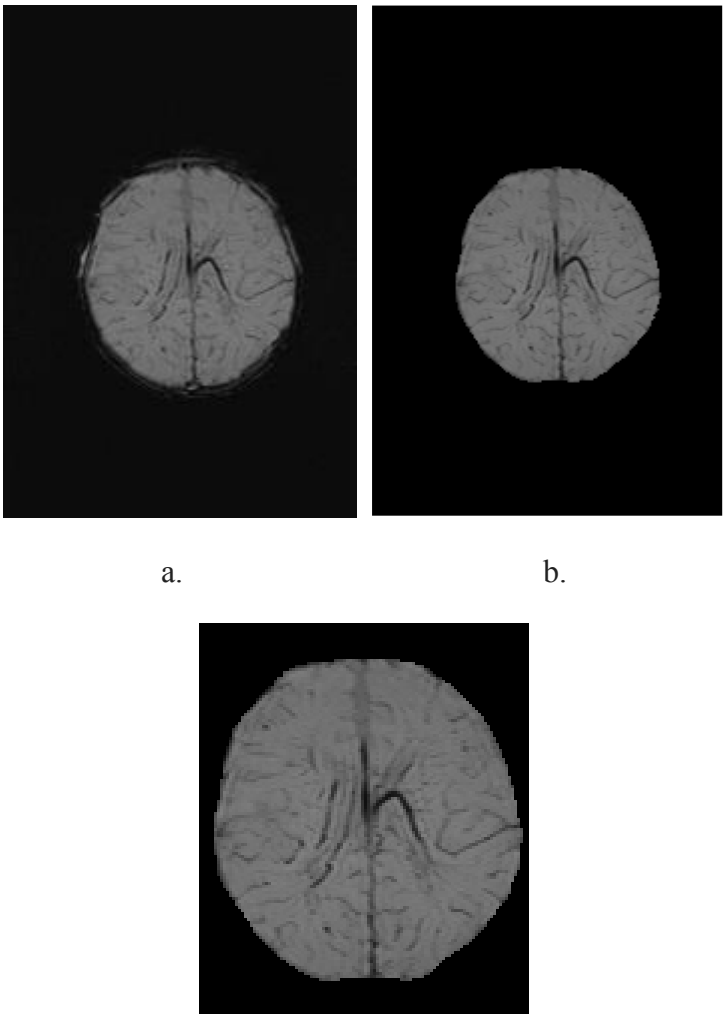


Figure 7. 8 Both a and b show the deprecated SWI images, and c and d show the selected SWI images.

7.5.2. Image Pre-Processing

Before training the data, we used an active contour model for each SWI slice to obtain a binary brain mask that could eliminate background noise from calvarium. We then use a mask with 110×130 (110 pixels width and 130 pixels height) to match the centre of input image automatically and cropped the SWI images into an image of 110×130 pixels to avoid the effect from the background in SWI images (see Figure 7.9).



a. b.

c.

Figure 7. 9 a) an original SWI image, b) the result after segmentation with active contour model, and c) the cropped image after active contour.

7.5.3. Fine Tuning

The 42 infants with information on neurology outcomes were classified into two groups: the normal group (465 SWI images and 31 patients with normal neurological outcomes) and the abnormal group (165 SWI images and 11 patients with abnormal neurological outcomes;

cerebral palsy). To balance the data distribution, 11 out of the 31 patients with normal neurology outcomes were randomly selected for the training dataset. This selection was performed 10 times to train the CNNs model. After each selection, we randomly selected six patients with 90 SWI images (three patients from the normal group and three patients from the abnormal group) for testing and split the remaining patients into training and validation datasets (70% data for training and 30% data for validation). Finally, we transferred the learning of pre-trained ResNet50 to our SWI images with HIE for classification by neurology outcome, as shown in Figure 7.6. As there were not enough datasets of patients with motor and cognitive outcome on which to train the transfer learning network, we only used data from patients with neurology outcome for training.

7.6. Result and Analysis of Classification of Deep Learning

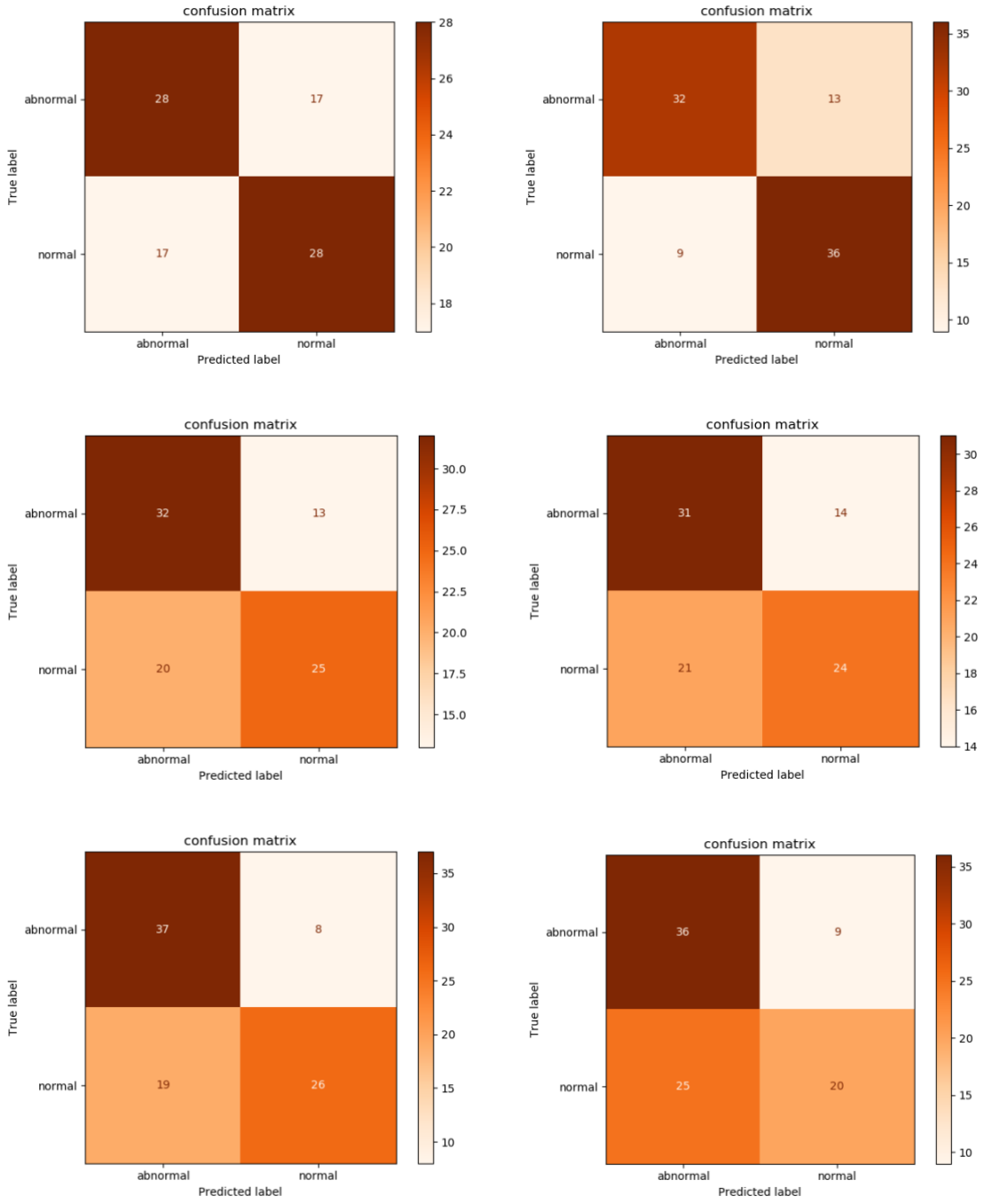
In this subsection, transfer learning in conjunction with a pre-trained ResNet 50 is utilised to retrain the network by the SWI images with normal and abnormal neurology outcomes at 24 months of age in neonatal HIE. Furthermore, the contributions of different brain regions in SWI images will be considered and Grad-CAM will be used to estimate the significance of the regions of the brain.

The pre-trained model was fine-tuned by inputting SWI images, and a classification label will be assigned to each SWI image in the test dataset. This means that each patient will own 15 normal or abnormal classification labels. For all of the slices of the same patient, the slices that obtain the same classification label as the neurology outcome of the patient at two years of age are named here as correctly predicted slices, otherwise the slices are referred to as incorrectly predicted slices.

7.6.1. Result and Analysis of classification based on Networks

After we apply transfer learning to the fine-tuned networks to conduct the classification of neurology outcome by using SWI images, each patient in the test group will receive 15 classified outputs. The clinical neurology outcome of each patient is used as the ground truth for the slices belonging to the patient. According to the true and predicted labels of SWI images in the test group, we can calculate directly the accuracy, sensitivity and specificity. After the ten experiments are performed to measure the classification accuracies by selecting 10 randomly selected patients for normal groups to balance our dataset, the mean and standard

deviation of classification accuracies for all experiments are computed here. The confusion matrices for the classification of every experiment are shown in Figure 7.10.



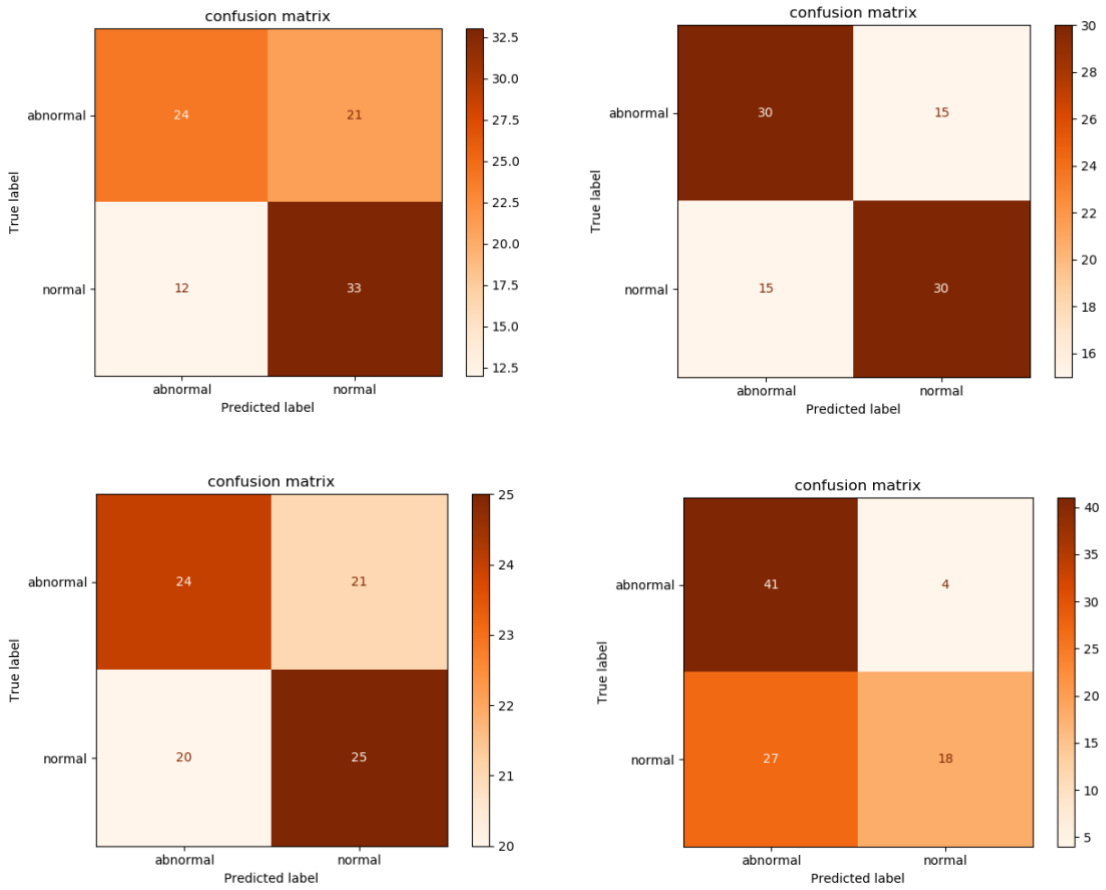


Figure 7. 10 The confusion matrices of ten experiments based on networks

As observed from Figure 7.10, the accuracy, sensitivity and specificity are 0.644±0.056, 0.703±0.121 and 0.598±0.109 respectively through using the classification result of each SWI image produced directly by the network. In terms of the accuracy, the fine-tuned networks have not performed very well. That is because for SWI images of some patients with HIE, there may be no HI injury or some subtle injury that would not lead to abnormal neurology at age 24 months. This will affect the training of the network model. The sensitivity shows good performance, which indicates that the network models have a low miss detection rate, and it can classify most of the SWI images of patients with abnormal neurology outcomes into abnormal labels. Comparatively lower results for specificity mean that more patients with normal labels are falsely reported.

7.6.2. Rule-based classification system

In the previous subsection, we analyse the classification results for the SWI images in the whole test group. In fact we will prefer to consider the classification results for each patient rather than the slices. Given that the SWI images in our dataset are 2D MRI axial images of the brain

injured by hypoxic ischaemic of infants at the birth, we face two issues in this dataset. The first issue is that these SWI images may not have a good resolution and quality, also may be contaminated with noise, which means the slices from the patients with normal neurology outcome will appear abnormal signals causing slices to be predicted into the abnormal group. The second issue is that the amount of brain damage of patients in normal group suffered from hypoxic ischaemic at birth may be too small to affect the neurological outcome at two years of age. It is also noted that the brain damage of an abnormal patient with abnormal neurology outcome may not appear on each single SWI image from that patient. This is to say that an abnormal patient may also contain some SWI images showing normal neurology outcome. Therefore, we need to eliminate the effect of SWI images containing brain damage on the classification of normal patients and the effect of SWI images not containing any brain damage on the classification of abnormal patients.

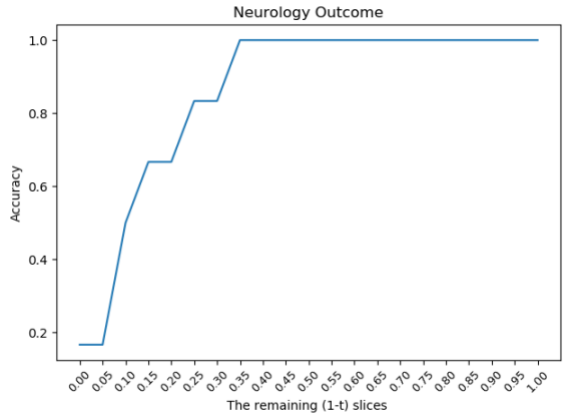
In order to consider the above two issues in the analysis of the classification of patients in the test group, we set the following rules with thresholds t and q:

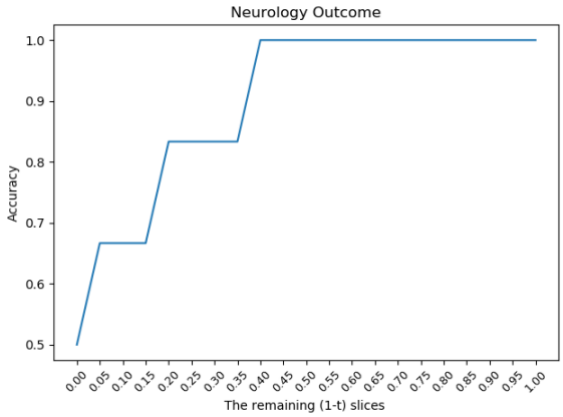
Rule1: When the ratio of (incorrectly predicted slices)/(total slices) of a patient is less than $\%t$, the patient is classified based on the class of $\%(1-t)$ remaining slices.

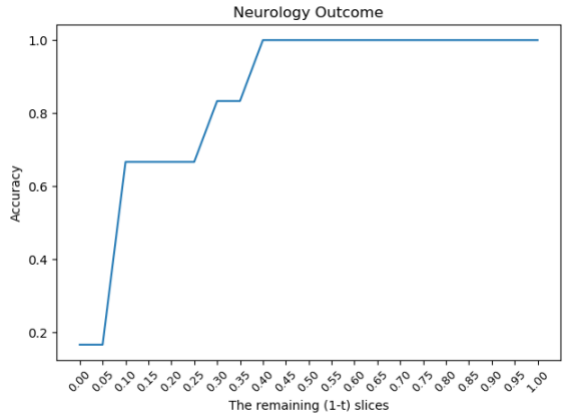
Rule2: If $q\%$ or more slices of a patient are predicted to be abnormal, the patient is considered as abnormal, otherwise the patient is considered to be normal.

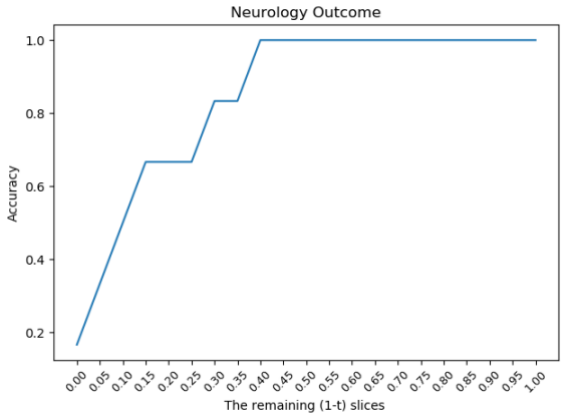
We employ **Rule1** and **Rule2** to consider the first and second issues in our dataset. Based on these two rules, we can calculate the accuracy of the classification for neurology outcome through deep neural networks. The accuracy will be measured by comparing the test outcomes that are determined by our ruled –based system with the outcomes of clinical assessments (ground truth).

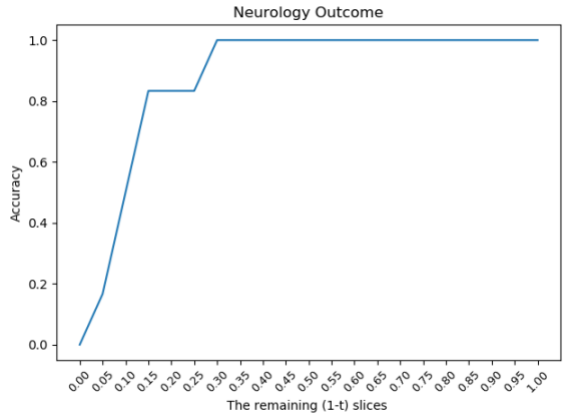
At the first step, we select a fixed t and then remove the incorrectly classified slices to measure the accuracy by following **Rule1**, where t is chosen in steps of 0.05 from $t = 0$. Figure 7.11 shows the different accuracies as t is changed at this step. In Table 7.1, the value of t is fixed and the mean and standard deviation of the classification accuracies are computed by ten experiments. As explained earlier, we repeat the test experiment by randomly choosing the normal patients to be in the normal class to balance our dataset.

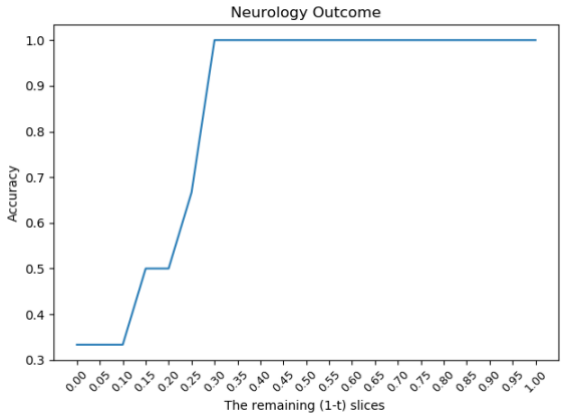

Neurology Outcome
Accuracy
The remaining (1-t) slices


Neurology Outcome
Accuracy
The remaining (1-t) slices


Neurology Outcome
Accuracy
The remaining (1-t) slices


Neurology Outcome
Accuracy
The remaining (1-t) slices


Neurology Outcome
Accuracy
The remaining (1-t) slices


Neurology Outcome
Accuracy
The remaining (1-t) slices

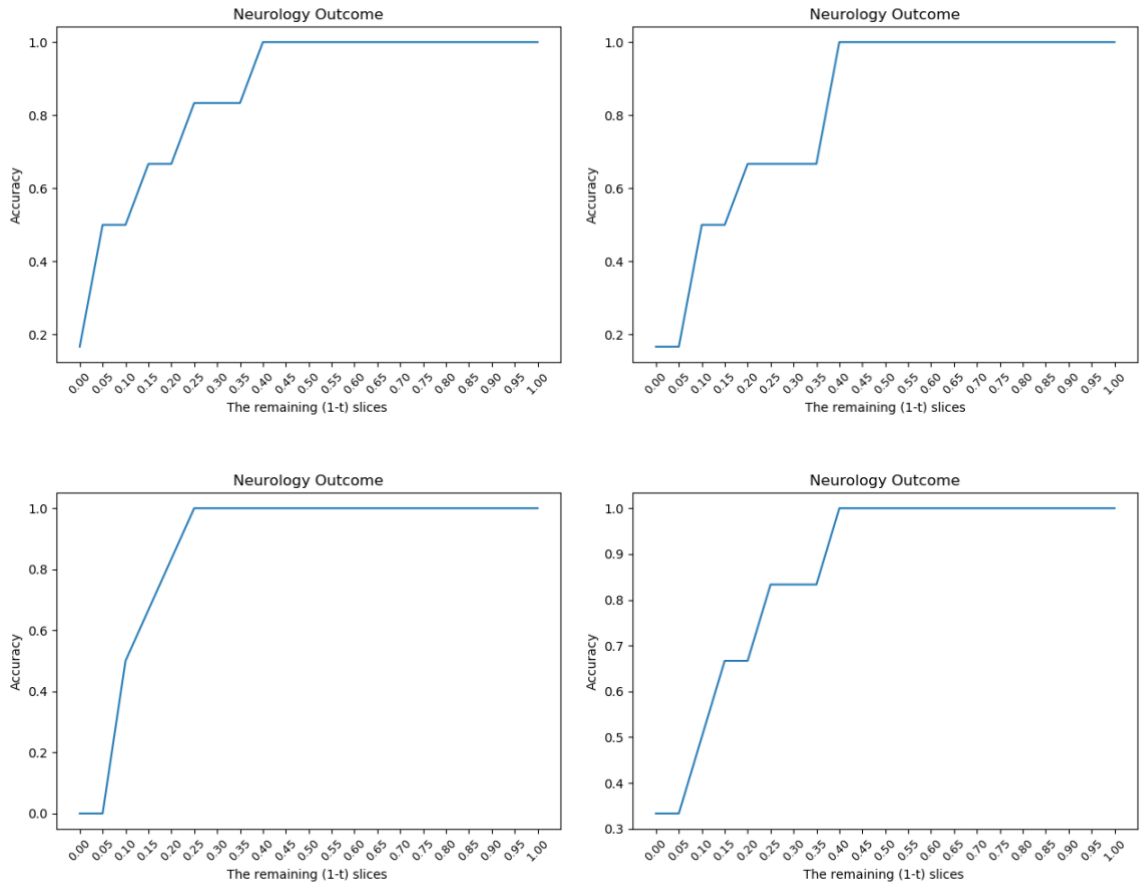


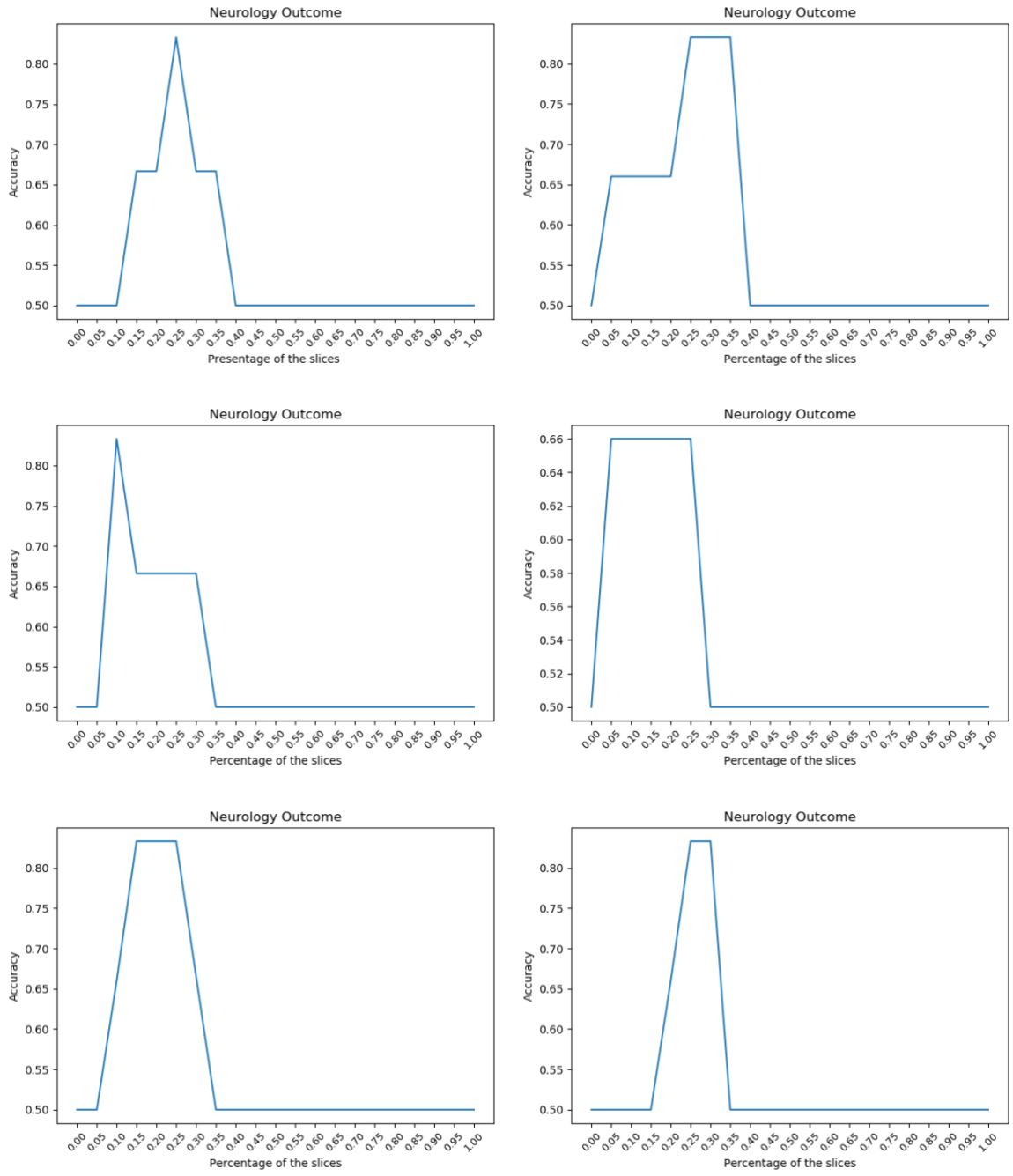
Figure 7. 11 The accuracy results from the ten experiments under **Relu1**

Table 7. 1 The accuracy of classification under **Rule1** by fixed t

Fixed *t*	*t=0*	*t=0.05*	*t=0.1*	*t=0.15*	*t=0.2*	*t=0.25*	*t=0.3*	*t=0.35*	***t=0.4***
Accuracies	0.2 ±0.153	0.28 ±0.193	0.51 ±0.094	0.65 ±0.095	0.70 ±0.105	0.783 ±0.112	0.867 ±0.105	0.883 ±0.112	**1.0 ±0.0**
Fixed *t*	*t=0.45*	*t=0.5*	*t=0.55*	*t=0.6*	*t=0.65*	*t=0.7*	...	*t=0.95*	*t=1.0*
Accuracies	1.0 ±0.0	1.0 ±0.0	1.0 ±0.0	1.0 ±0.0	1.0 ±0.0	1.0 ±0.0	…	1.0 ±0.0	1.0 ±0.0

As observed from Table 7.1, with *t =0.45*, the accuracy under **Rule 1** reaches the maximum, 1.0±0.0, and maintains this accuracy until *t = 1.0*.

We then explore the classification result under **Rule 2** without *t*. Figure 7.12 shows the accuracies by considering **Rule 2** in one experiment as *q* is changed. Similarly, *q* is incremented in steps of 0.05 from *q =0*. In Table 7.2, we fix the value of *q* and the mean and standard deviation of the classification accuracies are derived by ten experiments for ten random selections of normal paitents.



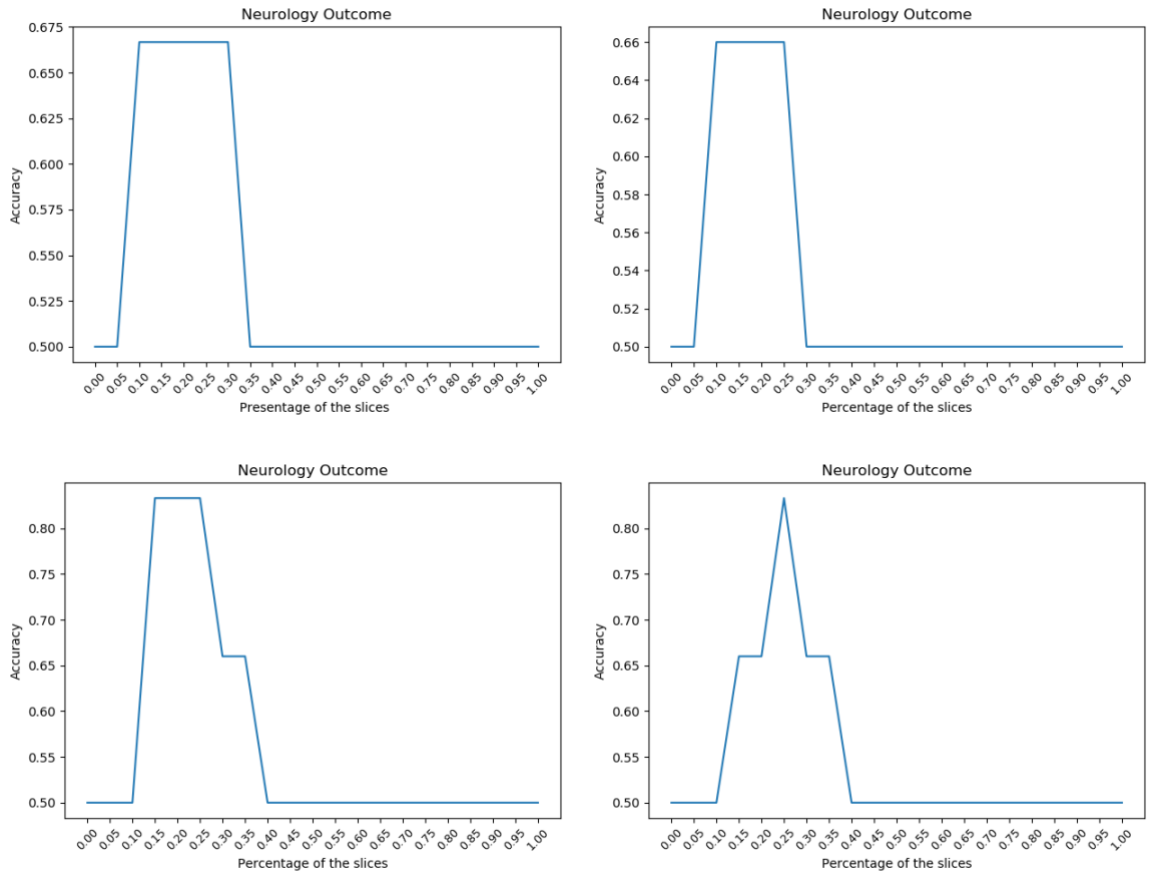


Figure 7. 12 The accuracy results from the ten experiments based our **Rule 2**

Table 7. 2 The accuracy of classification under **Rule 2** for various values of q

Fixed *q*	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	***q=0.25***	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.516 ±0.095	0.65 ±0.123	0.683 ±0.095	0.667 ±0.078	**0.766 ±0.086**	0.633 ±0.131	0.567 ±0.116	0.5 ±0.0
Fixed *q*	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	...	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	…	0.5 ±0.0	0.5 ±0.0

As shown in Table 7.2, the best accuracy of classification under **Rule 2** is 0.766 ± 0.086 when *q = 0.25*.

Limited by the amount of valid data, it is found that the accuracy of the test group stays the same after increasing the value of *t* or *q* in a certain range, while using the single **Rule 1** or **Rule 2**. To better resolve the two issues in our dataset which we discussed earlier, we combined **Rule 1** and **Rule 2** to improve the classification performance of our ruled based deep learning system. For each value of *t,* the slices which have been incorrectly classified are removed from the dataset. Then **Rule 2** is applied on the remaining slices. This is to say, we select a value for *t* in steps of 0.05 in sequence from 0 to 0.4, then remove the corresponding incorrectly predicted slices and introduce *q* to the (1-*t*) remaining slices in each experiment, which means that in the remaining slices if *q%* or more slices of a patient are predicted to be abnormal, the patient is considered as abnormal, otherwise the patient is considered to be normal. Table (7.3-7.10) show the accuracies of classification by various values of *t* after we combine **Rule 1** and **Rule 2.**

Table 7. 3 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.05

t = 0.05	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.683 ±0.123	0.717 ±0.112	0.70 ±0.087	0.750 ±0.087	**0.783 ±0.137**	0.683 ±0.122	0.65 ±0.123	0.5 ±0.0
t = 0.05	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	...	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

Table 7. 4 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.1

t = 0.1	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.749 ±0.117	0.733 ±0.086	0.766 ±0.086	**0.816 ±0.123**	0.783 ±0.137	0.716 ±0.112	0.70 ±0.105	0.5 ±0.0

t = 0.1	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	...	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

Table 7. 5 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.15

t = 0.15	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.766 ±0.086	0.799 ±0.07	0.833 ±0.136	0.816 ±0.123	**0.849 ±0.094**	0.75 ±0.087	0.70 ±0.105	0.5 ±0.0
t = 0.15	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	...	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

Table 7. 6 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.2

t = 0.2	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.833 ±0.111	0.866 ±0.105	0.833 ±0.136	**0.883 ±0.08**	0.849 ±0.094	0.75 ±0.087	0.70 ±0.105	0.5 ±0.0
t = 0.2	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	...	*q=0.95*	*q=1.0*
Accuracies	0.5	0.5	0.5	0.5	0.5	0.5		0.5	0.5

	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0		±0.0	±0.0

Table 7. 7 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.25

t = 0.25	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.866 ±0.105	0.866 ±0.105	**0.883 ±0.08**	**0.883 ±0.08**	0.833 ±0.111	0.716 ±0.08	0.666 ±0.136	0.5 ±0.0
t = 0.25	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	*...*	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

Table 7. 8 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.3

t = 0.3	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.866 ±0.105	**0.899 ±0.116**	0.883 ±0.08	0.866 ±0.07	0.816 ±0.122	0.70 ±0.105	0.666 ±0.136	0.5 ±0.0
t = 0.3	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	*...*	*q=0.95*	*q=1.0*

Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

Table 7. 9 The accuracy of classification after combine **Rule1** and **Rule2** by fixed t =0.35

t = 0.35	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.899 ±0.116	**0.933 ±0.086**	0.883 ±0.08	0.866 ±0.07	0.816 ±0.122	0.70 ±0.105	0.666 ±0.136	0.5 ±0.0
t = 0.35	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	*…*	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

Table 7. 10 The accuracy of classification after combine **Rule 1** and **Rule 2** by fixed t =0.4

t = 0.4	*q=0*	*q=0.05*	*q=0.1*	*q=0.15*	*q=0.2*	*q=0.25*	*q=0.3*	*q=0.35*	*q=0.4*
Accuracies	0.5 ±0.0	0.899 ±0.116	**0.933 ±0.086**	0.883 ±0.08	0.866 ±0.07	0.816 ±0.122	0.70 ±0.105	0.666 ±0.136	0.5 ±0.0

t = 0.4	*q=0.45*	*q=0.5*	*q=0.55*	*q=0.6*	*q=0.65*	*q=0.7*	...	*q=0.95*	*q=1.0*
Accuracies	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0	0.5 ±0.0		0.5 ±0.0	0.5 ±0.0

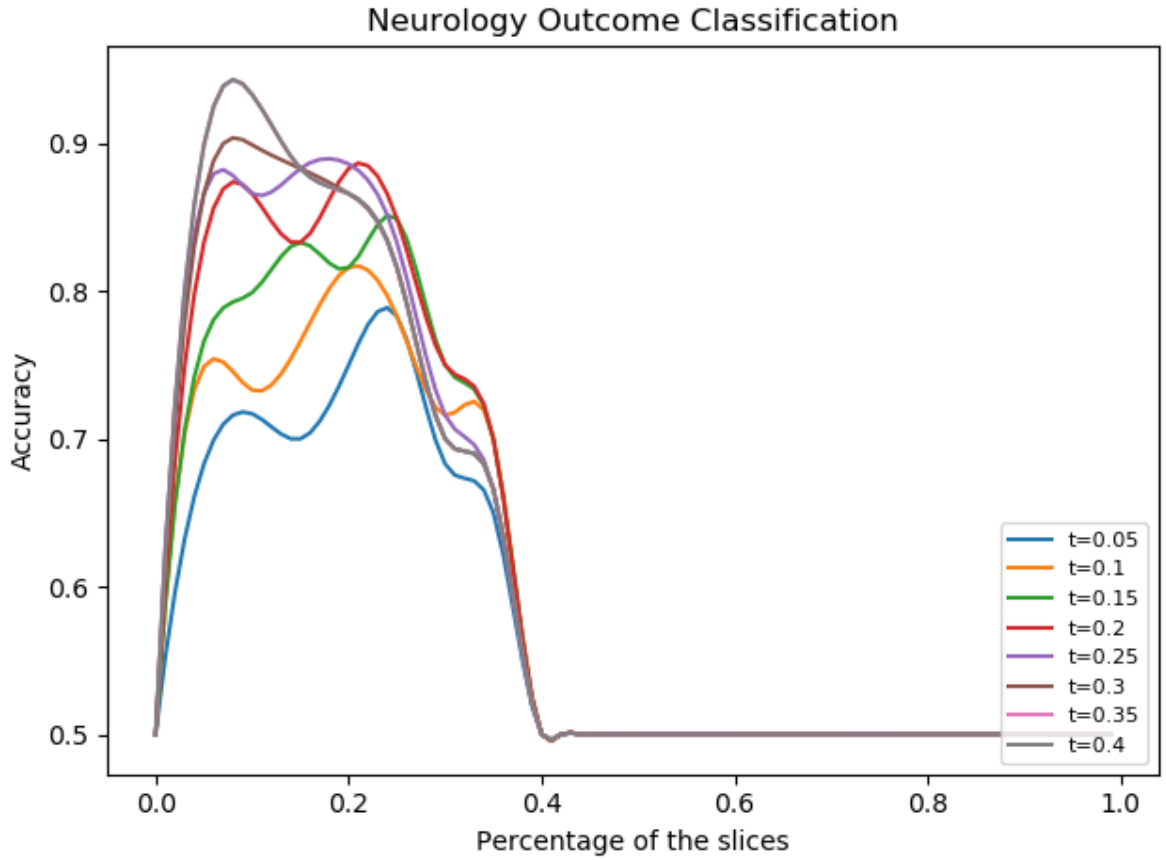


Figure 7. 13 Choosing t and *q* where the accuracy is maximum

As demonstrated in Table 7.10 and 7.11, the best accuracy of classification under combining **Rule 1** and **Rule 2** is 0.933 ±0.086 when $t = 0.35$ and $q = 0.1$ *or* $t = 0.4$ and $q = 0.1$. The combination of the rules and transfer learning with ResNet 50 produces a better performance compared to the accuracy of ***Algorithm 1***. Figure 7.13 shows how accuracies change with respect to q by changing *t* under combining **Rule 1** and **Rule 2** based on deep learning. The Receiver Operator Characteristic (ROC) curves of rule based on deep learning and the ***Algorithm 1*** for neurology outcome are shown in Figure 7.14. The ROC curve corresponding to the combined Rule 1 and Rule 2 demonstrates a higher performance in classification (with an area under curve higher than 92%).

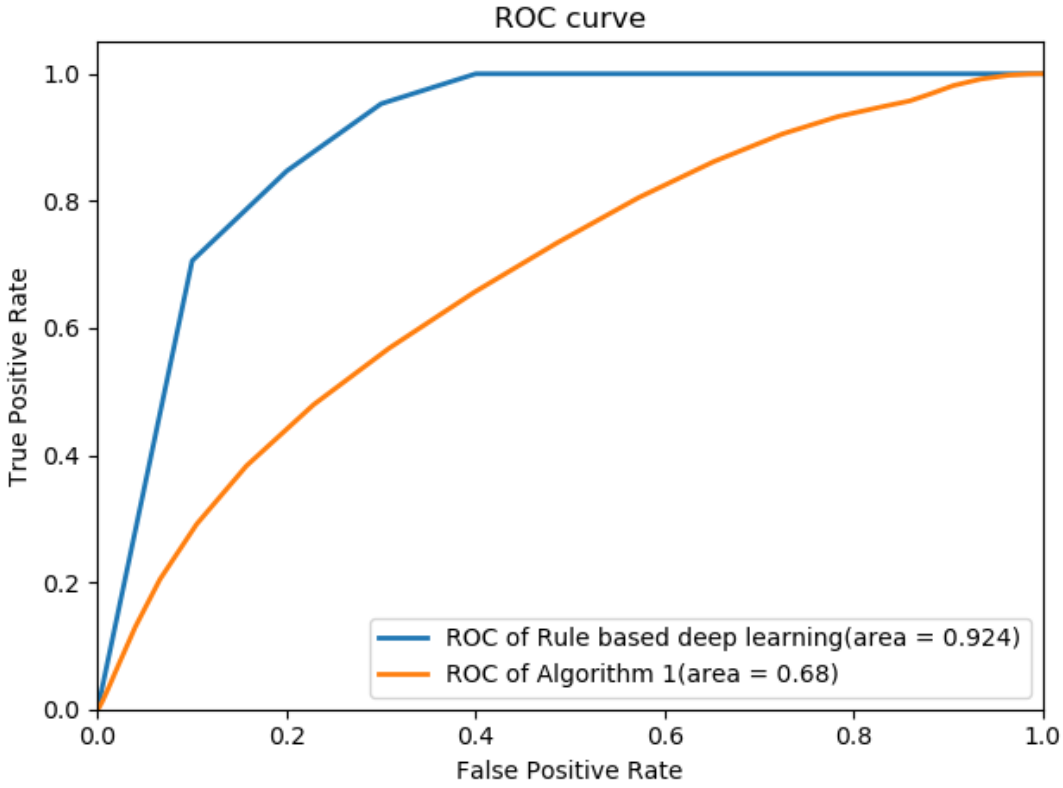


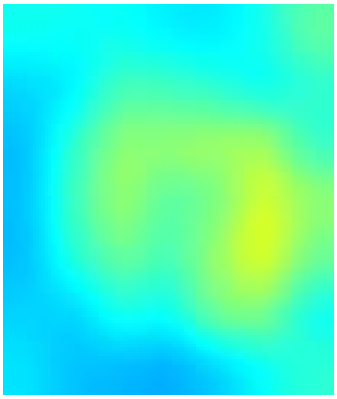
Figure 7. 14 The ROC curve of our rule based deep learning and the ***Algorithm 1***

Table 7. 11 Classification performance comparison

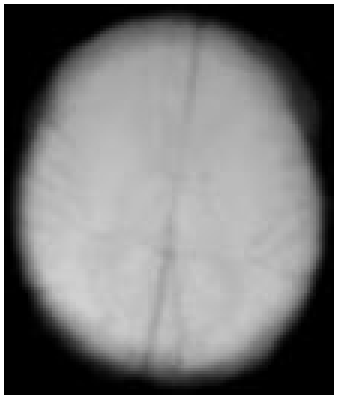
Methods	***Algorithm 1***	Rule based Deep Learning
Accuracy	0.727 ± 0.056	0.933 ±0.086

7.6.3. Result and Analysis of Grad-CAM

We employ Grad-CAM to explore which parts are important for identification of a SWI brain image with neonatal HIE. These blue heatmaps use different colours in order to show which regions of the brain are significant in assigning the images to a certain class. Figure 7.15 and 7.16 show the average heatmap images for both the abnormal test group and the normal test group. The red areas in the heatmaps are the most important, and the blue parts are the least important regions of the images for assigning them to a certain group.

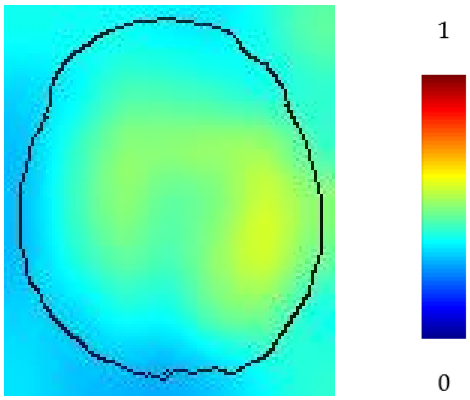


a. Average heatmap



b. Average total test images

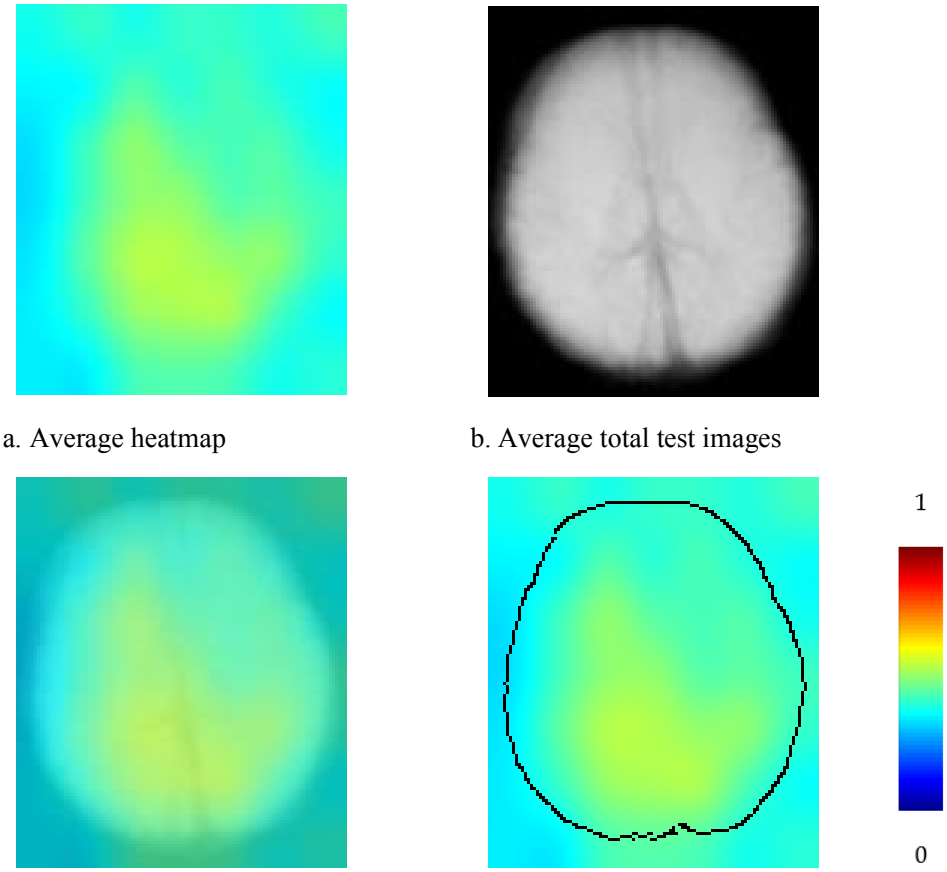
c. Cover heatmap and test image



d. Heatmap with brain outline

Figure 7. 15 Average heatmaps of abnormal group for SWI image classification

As Figure7.15 demonstrates, it looks like that the motor area of the brain is more important and the other regions are less important for classification of SWI images based on ResNet-50. However, in the average heatmaps of normal group based on the same deep network, the parts of parietal lobe of the brain make more contribution than other parts (see Figure 7.16). In order to clearly analyse which part of the SWI images from the infants with neonatal HIE in abnormal group is more valued in the classification task of the ResNet-50, we use Grad-CAM to calculate the average heatmap for each patient in the abnormal group and the heatmap of the central slice of a set of SWI images for each patient.



a. Average heatmap b. Average total test images

c. Cover heatmap and test image d. Heatmap with brain outline

Figure 7. 16 Average heatmaps of normal group for SWI image classification

In these heatmaps, there is activation outside of the brain. This is because when using transfer learning, the distribution of data in the source and target domains is different. Our pretrained network is based on the ImageNet dataset. This dataset may already contain information from MRI images. The performance of transfer learning depends on the similarity between datasets. The difference between the properties of the ImageNet dataset and the medical image dataset means that the pretrained model has weaker generalization on the small SWI dataset. Due to the limited data available, the network parameters of the ImageNet dataset affect the contribution of SWI images to the classification of patients.

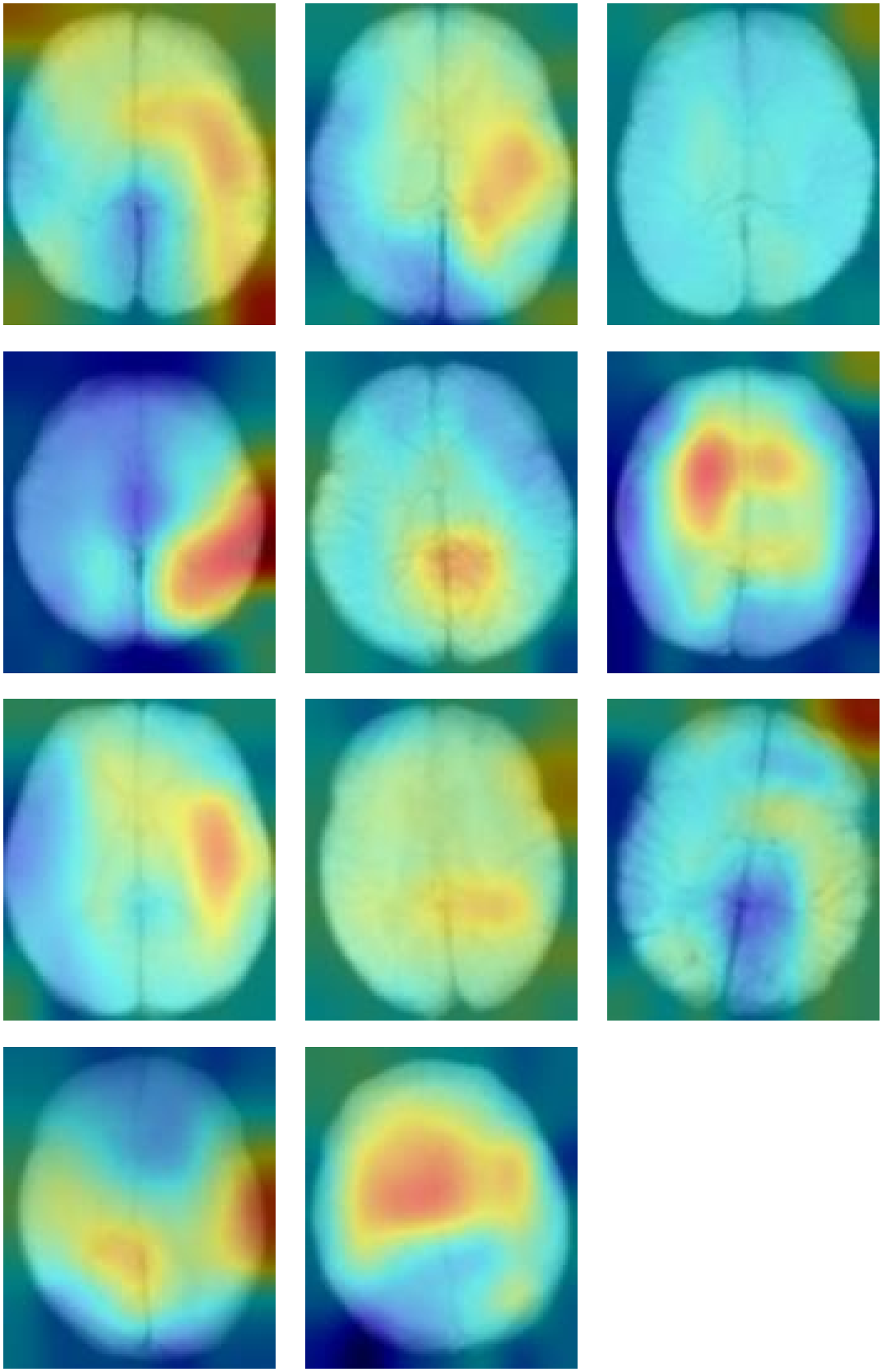
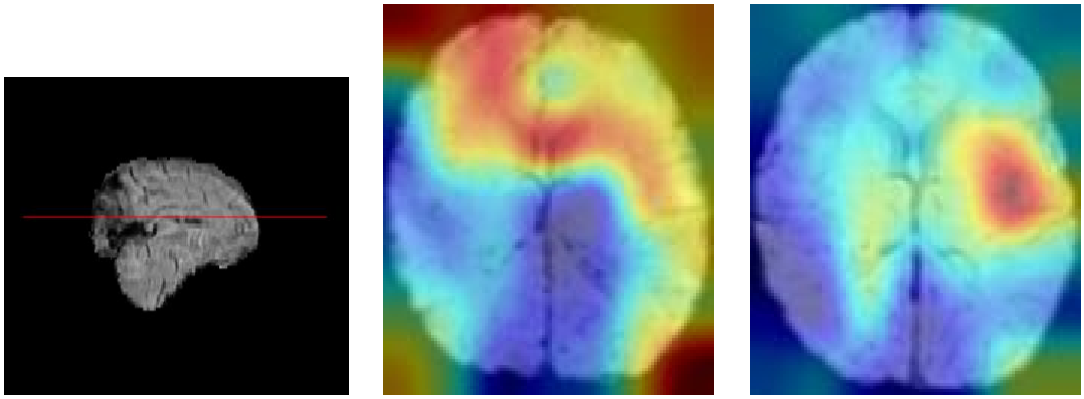


Figure 7. 17 Average heatmaps of each patient in abnormal group with brain outlines



a.

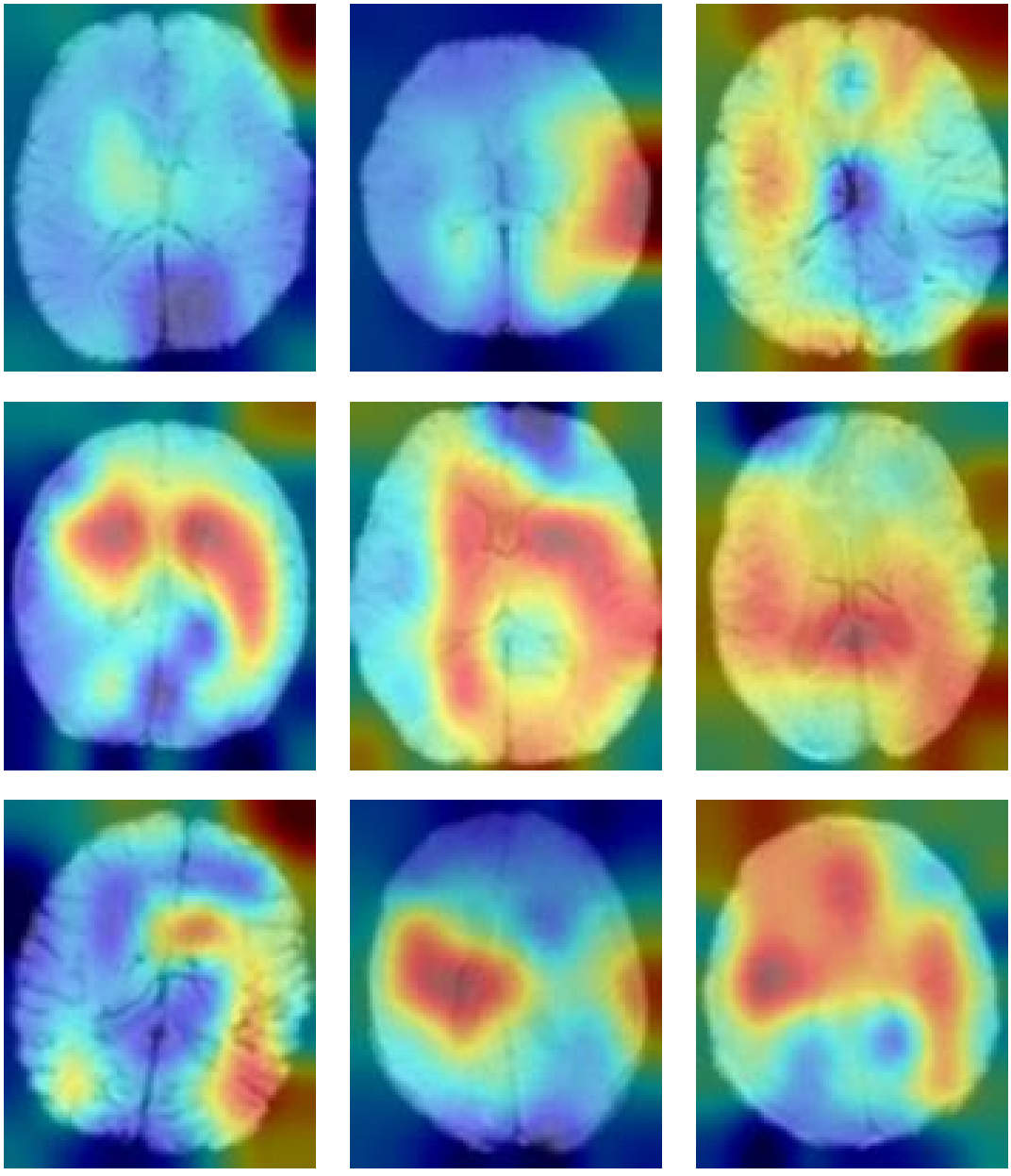


Figure 7. 18 The heatmaps of the middle slice of SWI images of each patient in abnormal group with brain outlines. a) Location of the centre SWI slice in the brain

Figure 7.17 shows the average heatmaps of the 11 abnormal patients. Figure 7.18 illustrates the heatmaps of the middle slice (red line in Figure 7.18(a)) of SWI images of each patient in abnormal group. Although the different heatmaps show the contribution of different regions based on ResNet-50, the contribution of the motor area differs significantly from the other regions. In almost every heatmap of the patients in abnormal group, the motor area is the field of interest for classification decision based on ResNet-50 (red part). We also notice that these heatmaps are also under the influence of the large data set (ImageNet) they have been trained for.

7.7. Summary

In this chapter, a convolutional neural networks, known as residual networks (ResNet) with fine-tuned networks are employed to classify HIE patients into normal or abnormal groups by using their SWI images, and the main emphasis is on describing the residual networks that were applied in this study and the visualisation of the classification results. Here by fine-tuning a pre-trained ResNet 50 model, we use transfer learning method to train our SWI dataset. We acknowledge that our SWI dataset is too small even to use, the transfer learning. In order to improve the accuracy of system, we introduce a rule-based system to use two rules known as **Rule 1** and **Rule 2** in this thesis to improve the classification performance of our deep learning system for patients with neurology outcomes and we also use the Grad-CAM to present the brain regions important in each group. The deep learning approaches have better performances than the ***Algorithm 1*** based on combining **Rule1** and **Rule2.** The best accuracy for classification of patients by using our rule based deep learning method in the test data is 0.933 ±0.086.

Additionally, we considered which parts of the brain in SWI HIE images contribute the most for allocating the patients to a particular class by presenting the average heat maps for each group. This is also the first study to use Grad-CAM to analysis which region of the SWI images with HIE contributes more to the classification results. We have plotted the average heatmaps for patients in the abnormal and normal groups, the average heatmaps for each patient in the abnormal group, and individual heatmap for the same location slice of each patient in the abnormal group. We compared the average heatmaps of patients in the abnormal and normal groups, which showed a greater contribution of the motor area to the classification of abnormal patients, while the parietal lobe play an important role in the classification of normal patients. The average heatmaps of each patient in abnormal group also support the greater contribution

of the motor area in the classification of abnormal patients to the abnormal group. This also is inline with medical finding [30,36].

In this study, as our small dataset posed a challenge to the deep learning task, we will consider collecting more data in future work to improve the accuracy of using SWI images taken at birth from infants with HIE to predict neurology outcomes assessed two years later.

Chapter 8

Conclusions and Future Works

8.1. Conclusions

MRI imaging has become the standard for detecting neonatal hypoxic-ischemic encephalopathy (HIE) and plays a crucial role in the prognosis of HIE. There are major advances in combining different sequences of MRI images with computer vision algorithms to predict HIE. This thesis uses computer vision and deep learning techniques to extract some features from susceptibility-weighted imaging (SWI) images of infants with HIE for the prognosis of HIE.

Since the SWI images could distinctively visualise the cerebral veins and the haemorrhagic parts of infants with HIE, we first investigated the structure and gray scale intensity of the veins. For this purpose, we have developed a new automated system for detecting ridges representing veins on the SWI images. We calculated the width, intensity value, length, and the eigenvalues of the Hessian matrix of ridges associated with the veins as feature vectors by analysing the ridges in SWIs. Individual or concatenated features from a balanced dataset were used as input for k-nearest neighbours (*kNN*) and random forest classifiers with a leave-one-out validation strategy to classify SWI images for evaluating the neurology outcome of neonates with HIE at 24 months of age. Meanwhile, we applied these feature vectors to train support vector

regression (SVR) and random forest regression models for predicting the composite motor and cognitive scores of infants with HIE at 24 months of age. The best classification accuracy associated with the random forest classifier for concatenated features consisting of width, intensity, length, and Hessian eigenvalues is 78.67 ± 2.58%. Our result is better than the classification accuracy reported in [76], which is 72.27 ± 4.85%. It is noted that we have used the data from the same group of patients as the dataset used in [76], and the dataset is not balanced in [76]. Mean relative error (MRE) was used to measure the performance of the regression models, with the lowest mean relative errors of 0.113 ± 0.13 and 0.109 ± 0.067 for the cognitive and motor composite scores, respectively. This is the first study to train a regression model by extracting features from SWI images of infants with HIE at birth to predict motor and cognitive scores of HIE survivors at 24 months of age.

We further developed ***Algorithm 1*** based on the detected ridges to find where the brain has been damaged by hypoxic-ischaemic (HI) injury to predict the neurodevelopmental outcomes of newborns with HIE as they grow up. ***Algorithm 1*** helped us to classify veins on SWI images into normal and abnormal groups supervised indirectly by clinical assessment outcomes. Experiments to measure the performance of ***Algorithm 1*** in various scenarios showed that ***Algorithm 1*** performed better than traditional *kNN* and random forest classifiers in datasets related to motor and cognitive outcomes. The accuracy for motor outcome using ***Algorithm 1*** was 75.0 ± 13.9% and the accuracy for cognitive outcome was 65.0 ± 9.8%. By analysing and identifying the veins classified in the abnormal group, we found that the abnormal veins on the SWI images of infants with HIE at birth were associated with follow-up outcomes of survivors at two years of age. The analysis of abnormal veins was accomplished by registering the SWI images with a brain template containing all brain lobes. We calculated the ratios of abnormal veins detected by ***Algorithm 1*** in various brain areas of infants with HIE. Of these, there was a significant correlation between motor outcome and abnormal veins in cortical motor areas, with the strongest correlation coefficient of -0.76. Vein length in the frontal, temporal, and parietal lobes affected by HIE had the highest correlation with cognitive scores with a correlation value of -0.52. We also trained the regression models by using the ratios of abnormal veins and motor and cognitive composite scores. The lowest Mean Relative Error (MRE) value for the linear regression model correlated with motor scores, is computed here as 0.088 ± 0.073. The polynomial regression model predicting cognitive scores obtained a minimum MRE value of 0.111 ± 0.103.

In addition, we exploited the histogram of oriented gradients (HOG) technique to obtain feature vectors to improve the classification performance of neurodevelopmental outcomes. Motor and cognitive regions of the brain on SWI images were obtained using the image registration method in Chapter 5. The HOG feature vectors of these regions were used as input for the *kNN* and random forest classifiers for classification. Compared to the previous results, the classification results for motor outcomes showed better performance than Chapter 4, with an accuracy of 71.67 ± 11.24. The combined HOG features of the frontal, temporal, and parietal lobes of the brain illustrated the better performance of classification for cognitive outcomes than Chapter 4 and Chapter 5, with an accuracy of 76.25 ± 10.9.

Furthermore, deep learning approaches were used to classify the SWI images of infants with HIE based on the neurology outcomes. We used the transfer learning method to train the SWI images by fine-tuning a pre-trained ResNet 50 model and introduced a rule-based system, containing two rules: **Rule 1** and **Rule 2**. The performance for the classification was 93.3% ± 8.6. We also considered which parts of the brain contribute the most to the classification of SWI images with neurological outcomes. The heatmaps presented in Chapter 7 demonstrated that the motor areas were important for the classification of neonates with HIE. This is consistent with the measurement results from ***Algorithm 1***. Deep learning methods have shown the best classification performance than previous methods in the literature as well as in the rest of this thesis.

In this thesis, we demonstrate that SWI image analysis of infants with HIE can provide additional information to predict neurodevelopmental outcomes at the age of two years. By developing ***Algorithm 1***, for SWI images that were not annotated by clinical experts, we still identified the veins that were affected by hypoxic-ischaemic injury and demonstrated that the brain regions where the abnormal veins were located, were correlated with neurodevelopmental outcomes at two years of age. Although there are some studies of the degree of hypoxic-ischaemic injury based on SWI images, there are no studies that classify damaged veins based on clinical assessment. We have also used HOG features as well as deep learning methods on SWI images to improve the performance classification and regression. The results supported the relationship between damaged regions of the brain and neurodevelopmental outcomes. This is also the first study to use SWI images with HIE to train a deep learning network.

8.2. Future works

With the development of neuroimaging and a deeper understanding of the pathological mechanisms of HIE, more and more studies are tending to explore the associations between veins and cerebral cortex damaged by hypoxic-ischaemic and long-term neurodevelopmental outcomes. A few studies are now also beginning to use neuroimaging of HIE patients to predict Bayley-III scores. However, the main issue in front of many HIE researchers is the lack of data.

One of our most significant tasks in the future will be to collect more data on clinical assessment outcomes at various time periods. The expansion of the dataset will improve the results of the methods in this thesis. In addition, we would like to use our method to analyse the neurodevelopmental outcomes at 48 months of age or during adolescence for HIE survivors by using our methods. Also, we can use neural network models to predict motor and cognitive scores on a large enough dataset.

In this work, we have focused on exploring the relationship between the functional cortex of the brain with HIE and developmental outcomes. One area for future work is to explore the relationship between damaged veins within other structures of the brain and long-term outcomes, such as veins near the corpus callosum, thalamus, and basal ganglia. We expect to combine these structures with veins in the cerebral cortex to improve the prognostic qualities of SWI images in HIE.

Another interesting future work is to combine HOG features with features obtained from ridge detection to improve the performance of the current automatic diagnosis system for SWI images. The last but not the least direction for future work would be to compare our vessel classification scheme, based on clinical outcomes, with the vessel scoring system proposed by Kitamura [45].

Reference

[1] Nadeem, M., Murray, D. M., Boylan, G. B., Dempsey, E. M., & Ryan, C. A. (2011). Early blood glucose profile and neurodevelopmental outcome at two years in neonatal hypoxic-ischaemic encephalopathy. *BMC pediatrics*, *11*(1), 1-6.

[2] Massaro, A. N., Evangelou, I., Fatemi, A., Vezina, G., Mccarter, R., Glass, P., & Limperopoulos, C. (2015). White matter tract integrity and developmental outcome in newborn infants with hypoxic-ischemic encephalopathy treated with hypothermia. *Developmental Medicine & Child Neurology*, *57*(5), 441-448.

[3] Martinez-Biarge, M., Diez-Sebastian, J., Rutherford, M. A., & Cowan, F. M. (2010). Outcomes after central grey matter injury in term perinatal hypoxic-ischaemic encephalopathy. *Early human development*, *86*(11), 675-682.

[4] Vannucci, R. C. (2000). Hypoxic-ischemic encephalopathy. *American journal of perinatology*, *17*(03), 113-120.

[5] Massaro, A. N., Evangelou, I., Brown, J., Fatemi, A., Vezina, G., McCarter, R., ... & Limperopoulos, C. (2015). Neonatal neurobehavior after therapeutic hypothermia for hypoxic ischemic encephalopathy. *Early human development*, *91*(10), 593-599.

[6] Lennartsson, F., Darekar, A., Maharatna, K., Konn, D., Allen, D., Tournier, J., ... & Vollmer, B. (2018, July). Developing a framework for studying brain networks in neonatal hypoxic-ischemic encephalopathy. In *Annual Conference on Medical Image Understanding and Analysis* (pp. 203-216). Springer, Cham.

[7] Douglas-Escobar, M., & Weiss, M. D. (2015). Hypoxic-ischemic encephalopathy: a review for the clinician. *JAMA pediatrics*, *169*(4), 397-403.

[8] Vannucci, R. C., & Perlman, J. M. (1997). Interventions for perinatal hypoxic-ischemic encephalopathy. *Pediatrics*, *100*(6), 1004-1014.

[9] Busl, K. M., & Greer, D. M. (2010). Hypoxic-ischemic brain injury: pathophysiology, neuropathology and mechanisms. *NeuroRehabilitation*, *26*(1), 5-13.

[10] Dixon, B. J., Reis, C., Ho, W. M., Tang, J., & Zhang, J. H. (2015). Neuroprotective strategies after neonatal hypoxic ischemic encephalopathy. *International journal of molecular sciences*, *16*(9), 22368-22401.

[11] Okereafor, A., Allsop, J., Counsell, S. J., Fitzpatrick, J., Azzopardi, D., Rutherford, M. A., & Cowan, F. M. (2008). Patterns of brain injury in neonates exposed to perinatal sentinel events. *Pediatrics*, *121*(5), 906-914.

[12] Mota-Rojas, D., Villanueva-García, D., Solimano, A., Muns, R., Ibarra-Ríos, D., & Mota-Reyes, A. (2022). Pathophysiology of Perinatal Asphyxia in Humans and Animal Models. *Biomedicines*, *10*(2), 347.

[13] https://my.clevelandclinic.org/health/diseases/6025-cerebral-hypoxia

[14] James, A., & Patel, V. (2014). Hypoxic ischaemic encephalopathy. *Paediatrics and Child Health*, *24*(9), 385-389.

[15] Allen, K. A., & Brandon, D. H. (2011). Hypoxic ischemic encephalopathy: pathophysiology and experimental treatments. *Newborn and Infant Nursing Reviews*, *11*(3), 125-133.

[16] Volpe, J. J. (2001). Perinatal brain injury: from pathogenesis to neuroprotection. *Mental retardation and developmental disabilities research reviews*, *7*(1), 56-64.

[17] Lakatos, A. (2021). *Prognostic Role of Early MRI in Neonatal Hypoxic Ischemic Encephalopathy* (Doctoral dissertation).

[18] Wisnowski, J. L., Wintermark, P., Bonifacio, S. L., Smyser, C. D., Barkovich, A. J., Edwards, A. D., ... & Publications Committee. (2021, October). Neuroimaging in the term newborn with neonatal encephalopathy. In *Seminars in Fetal and Neonatal Medicine* (Vol. 26, No. 5, p. 101304). WB Saunders.

[19] Krishnan, P., & Shroff, M. (2016). Neuroimaging in neonatal hypoxic ischemic encephalopathy. *The Indian Journal of Pediatrics*, *83*(9), 995-1002.

[20] Cowan, F. M., Pennock, J. M., Hanrahan, J. D., Manji, K. P., & Edwards, A. D. (1994). Early detection of cerebral infarction and hypoxic ischemic encephalopathy in neonates using diffusion-weighted magnetic resonance imaging. *Neuropediatrics*, *25*(04), 172-175.

[21] Bekiesińska-Figatowska, M., Szkudlińska-Pawlak, S., Kwaśniewicz, P., Duczkowska, A., Ring, M., Iwanowska, B., & Sawicki, M. (2021). Arterial spin labeling in neonatal magnetic resonance imaging–first experience and new observations. *Polish Journal of Radiology*, *86*(1), 415-424.

[22] Dibble, M., O'Dea, M. I., Hurley, T., Byrne, A., Colleran, G., Molloy, E. J., & Bokde, A. L. W. (2020). Diffusion tensor imaging in neonatal encephalopathy: a systematic review. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, *105*(5), 480-488.

[23] Lally, P. J., Pauliah, S., Montaldo, P., Chaban, B., Oliveira, V., Bainbridge, A., ... & Thayyil, S. (2015). Magnetic Resonance Biomarkers in Neonatal Encephalopathy (MARBLE): a prospective multicountry study. *BMJ open*, *5*(9), e008912.

[24] Haacke, E. M., Mittal, S., Wu, Z., Neelavalli, J., & Cheng, Y. C. (2009). Susceptibility-weighted imaging: technical aspects and clinical applications, part 1. *American Journal of Neuroradiology*, *30*(1), 19-30.

[25] Sehgal, V., Delproposto, Z., Haacke, E. M., Tong, K. A., Wycliffe, N., Kido, D. K., ... & Reichenbach, J. R. (2005). Clinical applications of neuroimaging with susceptibility-weighted imaging. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *22*(4), 439-450.

[26] Tang, Z., Mahmoodi, S., Dasmahapatra, S., Darekar, A., & Vollmer, B. (2020, July). Ridge detection and analysis of susceptibility-weighted magnetic resonance imaging in neonatal

hypoxic-ischaemic encephalopathy. In *Annual Conference on Medical Image Understanding and Analysis* (pp. 307-318). Springer, Cham.

[27] Tang, Z.; Mahmoodi, S.; Darekar, A. and Vollmer, B. (2022). Hypoxic-Ischaemic Encephalopathy Prognosis using Susceptibility Weighted Image Analysis based on Histogram Orientation Gradient. In *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - BIOSIGNALS,* ISBN 978-989-758-552-4; ISSN 2184-4305, pages 57-62.

[28] https://www.abclawcenters.com/practice-areas/prenatal-birth-injuries/fetus-or-newborn-medical-problems/hypoxic-ischemic-encephalopathy/

[29] Wang, G., Shen, X., Tian, Q., Lakshmipriya, T., & Gopinath, S. C. (2020). The Correlation Between Ischemic Stroke and Thrombosis by Nanoscale Biomarker Analysis. *Current Nanoscience*, *16*(5), 676-684.

[30] Kinney, H. C., & Volpe, J. J. (2018). Hypoxic-ischemic injury in the term infant: neuropathology. In *Volpe's neurology of the newborn* (pp. 484-499). Elsevier.

[31] Parmentier, C. E., de Vries, L. S., & Groenendaal, F. (2022). Magnetic Resonance Imaging in (Near-) Term Infants with Hypoxic-Ischemic Encephalopathy. *Diagnostics*, *12*(3), 645.

[32] Sarnat, H. B., & Sarnat, M. S. (1976). Neonatal encephalopathy following fetal distress: a clinical and electroencephalographic study. *Archives of neurology*, *33*(10), 696-705.

[33] Thompson, C. M., Puterman, A. S., Linley, L. L., Hann, F. M., Van der Elst, C. W., Molteno, C. D., & Malan, A. F. (1997). The value of a scoring system for hypoxic ischaemic encephalopathy in predicting neurodevelopmental outcome. *Acta paediatrica*, *86*(7), 757-761.

[34] Abbasi, H., & Unsworth, C. P. (2020). Applications of advanced signal processing and machine learning in the neonatal hypoxic-ischemic electroencephalography. *Neural regeneration research*, *15*(2), 222.

[35] Bourel-ponchel, E., Wallois, F., Flamein, F., & Lamblin, M. D. (2019). P02-F Early EEG characteristics in hypoxic-ischemic encephalopathy and outcome at 2 years. *Clinical Neurophysiology*, *130*(7), e67.

[36] Inder, T. E., & Volpe, J. J. (2018). Hypoxic-ischemic injury in the term infant: clinical-neurological features, diagnosis, imaging, prognosis, therapy. In *Volpe's neurology of the newborn* (pp. 510-563). Elsevier.

[37] Eken, P., Toet, M. C., Groenendaal, F., & De Vries, L. S. (1995). Predictive value of early neuroimaging, pulsed Doppler and neurophysiology in full term infants with hypoxic-ischaemic encephalopathy. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, *73*(2), F75-F80.

[38] McArdle, C. B., Richardson, C. J., Hayden, C. K., Nicholas, D. A., & Amparo, E. G. (1987). Abnormalities of the neonatal brain: MR imaging. Part II. Hypoxic-ischemic brain injury. *Radiology*, *163*(2), 395-403.

[39] Badve, C. A., Khanna, P. C., & Ishak, G. E. (2012). Neonatal ischemic brain injury: what every radiologist needs to know. *Pediatric radiology*, *42*(5), 606-619.

[40] Hwang, M., Haddad, S., Tierradentro-Garcia, L. O., Alves, C. A., Taylor, G. A., & Darge, K. (2022). Current understanding and future potential applications of cerebral microvascular imaging in infants. *The British Journal of Radiology*, *95*(1133), 20211051.

[41] Barkovich, A. J., Hajnal, B. L., Vigneron, D., Sola, A., Partridge, J. C., Allen, F., & Ferriero, D. M. (1998). Prediction of neuromotor outcome in perinatal asphyxia: evaluation of MR scoring systems. *American Journal of Neuroradiology*, *19*(1), 143-149.

[42] Rutherford, M., Pennock, J., Schwieso, J., Cowan, F., & Dubowitz, L. (1996). Hypoxic-ischaemic encephalopathy: early and late magnetic resonance imaging findings in relation to outcome. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, *75*(3), F145-F151.

[43] Weeke, L. C., Groenendaal, F., Mudigonda, K., Blennow, M., Lequin, M. H., Meiners, L. C., ... & de Vries, L. S. (2018). A novel magnetic resonance imaging score predicts neurodevelopmental outcome after perinatal asphyxia and therapeutic hypothermia. *The Journal of pediatrics*, *192*, 33-40.

[44] Shankaran, S., Barnes, P. D., Hintz, S. R., Laptook, A. R., Zaterka-Baxter, K. M., McDonald, S. A., ... & Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. (2012). Brain injury following trial of hypothermia for neonatal hypoxic–ischaemic encephalopathy. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, *97*(6), F398-F404.

[45] Kitamura, G., Kido, D., Wycliffe, N., Jacobson, J. P., Oyoyo, U., & Ashwal, S. (2011). Hypoxic-ischemic injury: utility of susceptibility-weighted imaging. *Pediatric neurology*, *45*(4), 220-224.

[46] Shankaran, S., Laptook, A. R., Ehrenkranz, R. A., Tyson, J. E., McDonald, S. A., Donovan, E. F., ... & Jobe, A. H. (2005). Whole-body hypothermia for neonates with hypoxic–ischemic encephalopathy. *New England Journal of Medicine*, *353*(15), 1574-1584.

[47] Lemyre, B., & Chau, V. (2018). Hypothermia for newborns with hypoxic-ischemic encephalopathy. *Paediatrics & Child Health*, *23*(4), 285-291.

[48] Vollmer, B. (2020). Severe neonatal hypoxic-ischaemic brain injury: still an important cause of infantile spasms. *Developmental Medicine & Child Neurology*, *62*(1), 9-9.

[49] Cawley, P., & Chakkarapani, E. (2020). Fifteen-minute consultation: Therapeutic hypothermia for infants with hypoxic ischaemic encephalopathy—translating jargon, prognosis and uncertainty for parents. *Archives of Disease in Childhood-Education and Practice*, *105*(2), 75-83.

[50] Robertson, C. M., & Perlman, M. (2006). Follow-up of the term infant after hypoxic-ischemic encephalopathy. *Paediatrics & child health*, *11*(5), 278-282.

[51] Vollmer, B. (2019). Neonatal neurology: bridging the gap. *Developmental Medicine & Child Neurology*, *61*(1), 5-5.

[52] Bayley, N. (2009). *Bayley-III: Bayley Scales of infant and toddler development*. Florence, Italy: Giunti OS.

[53] Procianoy, R. S., Corso, A. L., Longo, M. G., Vedolin, L., & Silveira, R. C. (2019). Therapeutic hypothermia for neonatal hypoxic-ischemic encephalopathy: magnetic resonance imaging findings and neurological outcomes in a Brazilian cohort. *The Journal of Maternal-Fetal & Neonatal Medicine*, *32*(16), 2727-2734.

[54] Edmonds, C. J., Helps, S. K., Hart, D., Zatorska, A., Gupta, N., Cianfaglione, R., & Vollmer, B. (2020). Minor neurological signs and behavioural function at age 2 years in neonatal hypoxic ischaemic encephalopathy (HIE). *European Journal of Paediatric Neurology*, *27*, 78-85.

[55] van Schie, P. E., Schijns, J., Becher, J. G., Barkhof, F., van Weissenbruch, M. M., & Vermeulen, R. J. (2015). Long-term motor and behavioral outcome after perinatal hypoxic-ischemic encephalopathy. *european journal of paediatric neurology*, *19*(3), 354-359.

[56] Twomey, E., Twomey, A., Ryan, S., Murphy, J., & Donoghue, V. B. (2010). MR imaging of term infants with hypoxic-ischaemic encephalopathy as a predictor of neurodevelopmental outcome and late MRI appearances. *Pediatric radiology*, *40*(9), 1526-1535.

[57] Haataja, L., Mercuri, E., Guzzetta, A., Rutherford, M., Counsell, S., Frisone, M. F., ... & Dubowitz, L. (2001). Neurologic examination in infants with hypoxic-ischemic encephalopathy at age 9 to 14 months: use of optimality scores and correlation with magnetic resonance imaging findings. *The Journal of pediatrics*, *138*(3), 332-337.

[58] Dag, Y., Firat, A. K., Karakas, H. M., Alkan, A., Yakinci, C., & Erdem, G. (2006). Clinical outcomes of neonatal hypoxic ischemic encephalopathy evaluated with diffusion-weighted magnetic resonance imaging. *Diagnostic and Interventional Radiology*, *12*(3), 109.

[59] Goergen, S. K., Ang, H., Wong, F., Carse, E. A., Charlton, M., Evans, R., ... & Cheong, J. L. Y. (2014). Early MRI in term infants with perinatal hypoxic–ischaemic brain injury: Interobserver agreement and MRI predictors of outcome at 2 years. *Clinical radiology*, *69*(1), 72-81.

[60] Del Balzo, F., Maiolo, S., Papoff, P., Giannini, L., Moretti, C., Properzi, E., & Spalice, A. (2014). Electroencephalogram and magnetic resonance imaging comparison as a predicting factor for neurodevelopmental outcome in hypoxic ischemic encephalopathy infant treated with hypothermia. *Pediatric Reports*, *6*(03), 5532.

[61] Nanavati, T., Seemaladinne, N., Regier, M., Yossuck, P., & Pergami, P. (2015). Can we predict functional outcome in neonates with hypoxic ischemic encephalopathy by the combination of neuroimaging and electroencephalography?. *Pediatrics & Neonatology*, *56*(5), 307-316.

[62] Trivedi, S. B., Vesoulis, Z. A., Rao, R., Liao, S. M., Shimony, J. S., McKinstry, R. C., & Mathur, A. M. (2017). A validated clinical MRI injury scoring system in neonatal hypoxic-ischemic encephalopathy. *Pediatric radiology*, *47*(11), 1491-1499.

[63] Lakatos, A., Kolossváry, M., Szabó, M., Jermendy, Á., Barta, H., Gyebnár, G., ... & Kozák, L. R. (2019). Neurodevelopmental effect of intracranial hemorrhage observed in hypoxic ischemic

brain injury in hypothermia-treated asphyxiated neonates-an MRI study. *BMC pediatrics*, *19*(1), 1-11.

[64] Tharmapoopathy, P., Chisholm, P., Barlas, A., Varsami, M., Gupta, N., Ekitzidou, G., ... & Shah, D. K. (2020). In clinical practice, cerebral MRI in newborns is highly predictive of neurodevelopmental outcome after therapeutic hypothermia. *European Journal of Paediatric Neurology*, *25*, 127-133.

[65] Mittal, S., Wu, Z., Neelavalli, J., & Haacke, E. M. (2009). Susceptibility-weighted imaging: technical aspects and clinical applications, part 2. *American Journal of neuroradiology*, *30*(2), 232-252.

[66] Meoded, A., Poretti, A., Northington, F. J., Tekes, A., Intrapiromkul, J., & Huisman, T. A. G. M. (2012). Susceptibility weighted imaging of the neonatal brain. *Clinical radiology*, *67*(8), 793-801.

[67] Messina, S. A., Poretti, A., Tekes, A., Robertson, C., Johnston, M. V., & Huisman, T. A. (2014). Early Predictive Value of Susceptibility Weighted Imaging (SWI) in Pediatric Hypoxic-Ischemic Injury. *Journal of Neuroimaging*, *24*(5), 528-530.

[68] Tsui, Y. K., Tsai, F. Y., Hasso, A. N., Greensite, F., & Nguyen, B. V. (2009). Susceptibility-weighted imaging for differential diagnosis of cerebral vascular pathology: a pictorial review. *Journal of the neurological sciences*, *287*(1-2), 7-16.

[69] Li, X., Zhang, W., Liu, D., & Zeng, Y. W. (2019). Effect of 3.0 T Magnetic resonance SWI and MRS on early diagnosis of Neonatal HIE and Regression analysis of related predictive factors. *Journal of Hainan Medical University*, *25*(1), 75-78.

[70] Ghosh, N., Recker, R., Shah, A., Bhanu, B., Ashwal, S., & Obenaus, A. (2011). Automated ischemic lesion detection in a neonatal model of hypoxic ischemic injury. *Journal of Magnetic Resonance Imaging*, *33*(4), 772-781.

[71] Ghosh, N., Yuan, X., Turenius, C. I., Tone, B., Ambadipudi, K., Snyder, E. Y., ... & Ashwal, S. (2012). Automated core–penumbra quantification in neonatal ischemic brain injury. *Journal of Cerebral Blood Flow & Metabolism*, *32*(12), 2161-2170.

[72] Sarioglu, F. C., Sarioglu, O., Guleryuz, H., Deliloglu, B., Tuzun, F., Duman, N., & Ozkan, H. (2022). The role of MRI-based texture analysis to predict the severity of brain injury in neonates with perinatal asphyxia. *The British Journal of Radiology*, *95*(1132), 20210128.

[73] Murphy, K., van der Aa, N. E., Negro, S., Groenendaal, F., de Vries, L. S., Viergever, M. A., ... & Išgum, I. (2017). Automatic quantification of ischemic injury on diffusion-weighted MRI of neonatal hypoxic ischemic encephalopathy. *NeuroImage: Clinical*, *14*, 222-232.

[74] Weiss, R. J., Bates, S. V., Song, Y. N., Zhang, Y., Herzberg, E. M., Chen, Y. C., ... & Ou, Y. (2019). Mining multi-site clinical data to develop machine learning MRI biomarkers: application to neonatal hypoxic ischemic encephalopathy. *Journal of Translational Medicine*, *17*(1), 1-16.

[75] Ning, N., Li, X., Gao, J., Zhang, Y., Han, J., Luo, X., ... & Yang, J. (2013). Quantitative Measurement of Deep Medullary Venous in Susceptibility Weighted Imaging: Comparison of

Hypoxic-ischemic and Normal Neonates. *Society of Magnetic Resonance in Medicine Proceedings*.

[76] Wu, S., Mahmoodi, S., Darekar, A., Vollmer, B., Lewis, E., & Liljeroth, M. (2017, July). Feature extraction and classification to diagnose hypoxic-ischemic encephalopathy patients by using susceptibility-weighted MRI images. In *Annual Conference on Medical Image Understanding and Analysis* (pp. 527-536). Springer, Cham.

[77] Bériault, S., Xiao, Y., Collins, D. L., & Pike, G. B. (2015). Automatic SWI venography segmentation using conditional random fields. *IEEE Transactions on Medical Imaging*, *34*(12), 2478-2491.

[78] Ward, P. G., Ferris, N. J., Raniga, P., Dowe, D. L., Ng, A. C., Barnes, D. G., & Egan, G. F. (2018). Combining images and anatomical knowledge to improve automated vein segmentation in MRI. *Neuroimage*, *165*, 294-305.

[79] Citraro, L., Mahmoodi, S., Darekar, A., & Vollmer, B. (2017). Extended three-dimensional rotation invariant local binary patterns. *Image and vision Computing*, *62*, 8-18.

[80] Abbasi, H., Bennet, L., Gunn, A. J., & Unsworth, C. P. (2019, July). 2D wavelet scalogram training of deep convolutional neural network for automatic identification of micro-scale sharp wave biomarkers in the hypoxic-ischemic EEG of preterm sheep. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1825-1828). IEEE.

[81] Wang, J., Ju, R., Chen, Y., Liu, G., & Yi, Z. (2020). Automated diagnosis of neonatal encephalopathy on aEEG using deep neural networks. *Neurocomputing*, *398*, 95-107.

[82] Raurale, S. A., Boylan, G. B., Lightbody, G., & O'Toole, J. M. (2020, July). Grading the severity of hypoxic-ischemic encephalopathy in newborn EEG using a convolutional neural network. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 6103-6106). IEEE.

[83] Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International journal of computer vision*, *1*(4), 321-331.

[84] Illingworth, J., & Kittler, J. (1988). A survey of the Hough transform. *Computer vision, graphics, and image processing*, *44*(1), 87-116.

[85] Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*(2), 197-227.

[86] Mendonca, A. M., & Campilho, A. (2006). Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE transactions on medical imaging*, *25*(9), 1200-1213.

[87] Kirbas, C., & Quek, F. (2004). A review of vessel extraction techniques and algorithms. *ACM Computing Surveys (CSUR)*, *36*(2), 81-121.

[88] Staal, J., Abràmoff, M. D., Niemeijer, M., Viergever, M. A., & Van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging*, *23*(4), 501-509.

[89] Halpin, S., McCusker, C., Fogarty, L., White, J., Cavalière, E., Boylan, G., & Murray, D. (2022). Long-term neuropsychological and behavioral outcome of mild and moderate hypoxic ischemic encephalopathy. *Early Human Development*, *165*, 105541.

[90] Tusor, N., Wusthoff, C., Smee, N., Merchant, N., Arichi, T., Allsop, J. M., ... & Counsell, S. J. (2012). Prediction of neurodevelopmental outcome after hypoxic–ischemic encephalopathy treated with hypothermia by diffusion tensor imaging analyzed using tract-based spatial statistics. *Pediatric research*, *72*(1), 63-69.

[91] Domnick, N. K., Gretenkord, S., De Feo, V., Sedlacik, J., Brockmann, M. D., & Hanganu-Opatz, I. L. (2015). Neonatal hypoxia–ischemia impairs juvenile recognition memory by disrupting the maturation of prefrontal–hippocampal networks. *Experimental Neurology*, *273*, 202-214.

[92] Thorngren-Jerneck, K., Ohlsson, T., Sandell, A., Erlandsson, K., Strand, S. E., Ryding, E., & Svenningsen, N. W. (2001). Cerebral glucose metabolism measured by positron emission tomography in term newborn infants with hypoxic ischemic encephalopathy. *Pediatric research*, *49*(4), 495-501.

[93] Jiang, L., El-Metwally, D., Sours Rhodes, C., Zhuo, J., Almardawi, R., Medina, A. E., ... & Raghavan, P. (2022). Alterations in motor functional connectivity in Neonatal Hypoxic Ischemic Encephalopathy. *Brain injury*, 1-8.

[94] Bhroin, M. N., Kelly, L., Sweetman, D., Aslam, S., O'Dea, M. I., Hurley, T., ... & Bokde, A. L. (2022). Relationship between MRI scoring systems and neurodevelopmental outcome at two years in infants with neonatal encephalopathy. *Pediatric Neurology*, *126*, 35-42.

[95] De Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L. (2000). The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, *50*(1), 1-18.

[96] Finder, M., Boylan, G. B., Twomey, D., Ahearne, C., Murray, D. M., & Hallberg, B. (2020). Two-year neurodevelopmental outcomes after mild hypoxic ischemic encephalopathy in the era of therapeutic hypothermia. *JAMA pediatrics*, *174*(1), 48-55.

[97] Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTS). *Insight j*, *2*(365), 1-35.

[98] Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*, *54*(3), 2033-2044.

[99] Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., ... & Toga, A. W. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage*, *39*(3), 1064-1080.

[100] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886-893). Ieee.

[101] Wagner, F., Haenggi, M. M., Wagner, B., Weck, A., Weisstanner, C., Grunt, S., ... & Verma, R. K. (2015). The value of susceptibility-weighted imaging (SWI) in patients with non-neonatal hypoxic-ischemic encephalopathy. *Resuscitation*, *88*, 75-80.

[102] Alavi, S. E., & javad Rashti, M. (2018). Brain tumors detection on mri images through extracting hog features. *JOURNAL OF ADVANCED APPLIED SCIENTIFIC RESEARCH*, *2*(1), 9-25.

[103] Sarwinda, D., & Bustamam, A. (2018, June). 3D-HOG Features–Based Classification using MRI Images to Early Diagnosis of Alzheimer's Disease. In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)* (pp. 457-462). IEEE.

[104] Kim, H. G., Choi, J. W., Han, M., Lee, J. H., & Lee, H. S. (2020). Texture analysis of deep medullary veins on susceptibility-weighted imaging in infants: evaluating developmental and ischemic changes. *European Radiology*, *30*(5), 2594-2603.

[105] Raurale, S. A., Boylan, G. B., Mathieson, S. R., Marnane, W. P., Lightbody, G., & O'Toole, J. M. (2021). Grading hypoxic-ischemic encephalopathy in neonatal EEG with convolutional neural networks and quadratic time–frequency distributions. *Journal of Neural Engineering*, *18*(4), 046007.

[106] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision* (pp. 618-626).

[107] Wu, K., Otoo, E., & Shoshani, A. (2005, April). Optimizing connected component labeling algorithms. In *Medical Imaging 2005: Image Processing* (Vol. 5747, pp. 1965-1976). SPIE.

[108] Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*, *29*(2), 102-127.

[109] Zhang, X., Zhang, Y., & Hu, Q. (2019). Deep learning based vein segmentation from susceptibility-weighted images. *Computing*, *101*(6), 637-652.

[110] Fukushima, K., Miyake, S., & Ito, T. (1983). Neocognitron: A neural network model for a mechanism of visual pattern recognition. *IEEE transactions on systems, man, and cybernetics*, (5), 826-834.

[111] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

[112] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

[113] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

[114] Yu, H., Yang, L. T., Zhang, Q., Armstrong, D., & Deen, M. J. (2021). Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. *Neurocomputing,* 444, 92-110.

[115] Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv*:1405.3531.

[116] Lin, M., Chen, Q., & Yan, S. (2013). Network in network. *arXiv preprint arXiv:1312.4400*.

[117] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

[118] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[119] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

[120] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

[121] Peng, J., Kang, S., Ning, Z., Deng, H., Shen, J., Xu, Y., ... & Liu, L. (2020). Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging. *European radiology*, 30(1), 413-424.

[122] Vesoulis, Z. A., Trivedi, S. B., Morris, H. F., McKinstry, R. C., Li, Y., Mathur, A. M., & Wu, Y. W. Deep Learning to Optimize MRI Prediction of Motor Outcomes after Hie. *Available at SSRN 4265237*.

[123]Zwiggelaar, R., Astley, S. M., Boggis, C. R., & Taylor, C. J. (2004). Linear structures in mammographic images: detection and classification. *IEEE transactions on medical imaging*, 23(9), 1077-1086.

[124]Melinscak, M., Prentasic, P., & Loncaric, S. (2015, March). Retinal Vessel Segmentation using Deep Neural Networks. In *VISAPP* (1) (pp. 577-582).

[125]Mendonca, A. M., & Campilho, A. (2006). Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE transactions on medical imaging, 25(9)*, 1200-1213.

[126]Moccia, S., De Momi, E., El Hadji, S., & Mattos, L. S. (2018). Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics. *Computer methods and programs in biomedicine*, 158, 71-91.

[127]Straub, S., Stiegeler, J., El-Sanosy, E., Bendszus, M., Ladd, M. E., & Schneider, T. M. (2022). A novel gradient echo data based vein segmentation algorithm and its application for the detection of regional cerebral differences in venous susceptibility. *NeuroImage, 250*, 118931..

[128]Hoover, A. D., Kouznetsova, V., & Goldbaum, M. (2000). Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical imaging, 19(3)*, 203-210.

[129]Tolias, Y. A., & Panas, S. M. (1998). A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Transactions on Medical Imaging, 17(2)*, 263-273.

[130]De, J., Cheng, L., Zhang, X., Lin, F., Li, H., Ong, K. H., ... & Ahmed, S. (2015). A graph-theoretical approach for tracing filamentary structures in neuronal and retinal images. *IEEE transactions on medical imaging, 35(1),* 257-272.

[131]Shokouh, G. S., Magnier, B., Xu, B., & Montesinos, P. (2021). Ridge detection by image filtering techniques: a review and an objective analysis. *Pattern Recognition and Image Analysis, 31(3)*, 551-570.

[132]Hannink, J., Duits, R., & Bekkers, E. (2014, September). Crossing-preserving multi-scale vesselness. In *International conference on medical image computing and computer-assisted intervention (pp. 603-610)*. Springer, Cham.

[133]Walter, T., & Klein, J. C. (2001, October). Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques. In *International symposium on medical data analysis (pp. 282-287)*. Springer, Berlin, Heidelberg.

[134]Mulay, S., Ram, K., Sivaprakasam, M., & Vinekar, A. (2019, March). Early detection of retinopathy of prematurity stage using deep learning approach. In *Medical Imaging 2019: Computer-Aided Diagnosis* (Vol. 10950, pp. 758-764). SPIE.

[135]Zheng, Q., Martin-Saavedra, J. S., Saade-Lemus, S., Vossough, A., Zuccoli, G., Gonçalves, F. G., ... & Hwang, M. (2020). Cerebral pulsed arterial spin labeling perfusion weighted imaging predicts language and motor outcomes in neonatal hypoxic-ischemic encephalopathy. *Frontiers in Pediatrics, 8*, 576489.