Periportal Fibrosis Detection using Machine Learning

Alex Mutebe

MSc Data Science

SUBMITTED AS PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE

Department of Computing

University of Essex

2025

Keywords

- Periportal Fibrosis (PPF)
- Schistosoma mansoni infection
- Chronic liver diseases
- Morbidity and mortality
- Medical imaging
- Diagnostic accuracy
- Ultrasound images (US images)
- Artificial intelligence (AI)
- Machine learning (ML)
- Deep learning
- Feature extraction
- Classification
- Convolutional Neural Networks
- Deep Convolutional Neural Networks
- Support Vector Machine

Abstract

This thesis explores novel deep learning methods to enhance the detection of periportal fibrosis within medical imaging. Schistosoma mansoni infection affects over 54 million people, primarily in sub-Saharan Africa, with 20 million suffering chronic complications. Periportal Fibrosis (PPF) accounts for up to 42% of these cases. PPF is a prevalent outcome of chronic liver diseases, significantly contributing to both morbidity and mortality. Early detection is essential for timely intervention and potential reversibility of the condition. However, traditional imaging techniques, such as ultrasound, remain highly operator-dependent and require significant expertise for accurate diagnosis. Advances in artificial intelligence, particularly machine learning (ML), have shown great potential in medical imaging and diagnosis. Results highlight ML's ability to recognise complex patterns in imaging data, enhancing diagnostic accuracy for liver disease.

This research aims to develop and evaluate a machine learning-based approach for the accurate and efficient detection of periportal fibrosis in individuals with Schistosoma mansoni infection, using non-invasive diagnostic imaging data, specifically ultrasound images. This will be achieved by:

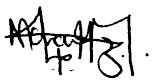
1. Developing and implementing a machine learning method for extracting relevant features indicative of periportal fibrosis from ultrasound images;
2. Training and validating the model using the extracted features to detect the presence of periportal fibrosis; and
3. Evaluating the model's performance using metrics such as accuracy, precision, sensitivity, F1 score, and specificity.

This research has led to the development of a novel deep learning model for periportal fibrosis detection, demonstrating high accuracy and efficiency in preliminary evaluations. Ultimately, this research seeks to contribute to improved patient outcomes by providing a reliable and efficient tool for periportal fibrosis detection.

Statement of Original Authorship

I, Alex Mutebe, confirm that this submission is my own work and, to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at the University of Essex or any other educational institution, except where due acknowledgement is made in the thesis.

I acknowledge any contribution made to the research by colleagues, both within and outside of the University of Essex, during my candidature. I further declare that the intellectual content of this thesis is the product of my own work, except where assistance from others in the project's design and conception, or in style, presentation, and linguistic expression, is explicitly acknowledged.

Signature  Date 20/02/2025

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Dr. Samuel Danso and Dr. Bakhtiyar Ahmed, for their invaluable guidance, support, and encouragement throughout this research project. Their expertise and insights were instrumental in shaping the direction of my research and ensuring its successful completion.

I would also like to extend my sincere appreciation to the Uganda Schistosomiasis Multidisciplinary Research Collaboration (U-SMRC) project team, specifically the study PI Professor Allison, for granting me access to the U-SMRC imaging dataset. Their generosity and support were essential for the successful completion of this research.

I am also grateful to Mr. Simon Mpooya, the study sonographer, for his invaluable support in processing the ultrasound images. His expertise and assistance were crucial in ensuring the quality and consistency of the data used in this research.

I am grateful to the University of Essex for providing the resources and facilities that enabled me to conduct this research. I would also like to acknowledge the Medical Research Council/Uganda Virus Research Institute and London School of Hygiene & Tropical Medicine Uganda Research Unit, for their generous financial support in covering my tuition fees for this degree. Their contribution has been instrumental in enabling me to pursue my academic goals. Finally, I would like to thank my family and friends for their unwavering support and encouragement during my studies.

Nomenclature

Abbreviations

AI	Artificial Intelligence
AUC	Area Under the Curve
CNN	Convolutional Neural Network
CT	Computed Tomography
DL	Deep Learning
FN	False Negative
FP	False Positive
FPR	False Positive Rate
GI	Gastrointestinal
HCC	Hepatocellular Carcinoma
ML	Machine Learning
MRI	Magnetic Resonance Imaging
PPF	Periportal Fibrosis
ROC	Receiver Operating Characteristic
ROI	Region of Interest
TN	True Negative
TP	True Positive
TPR	True Positive Rate
US	Ultrasound

List of Figures

List of Tables

1 Contents

2 Structure of the Document

This thesis is organised into the following chapters:

Chapter 1 - Introduction

Provides an overview of the research problem, objectives, and significance of the study.

Chapter 2 - Literature Review

This review examines existing literature on the causes of liver fibrosis, with a particular focus on periportal fibrosis. It then explores the limitations of current diagnostic techniques before discussing the potential of applying machine learning in medical imaging to address these limitations. Subsequently, it compares different machine learning approaches and ultimately focuses on Convolutional Neural Networks as a proposed model for detecting periportal fibrosis.

Chapter 3 – Methodology

Describes the research design, data collection methods, and the development of the machine learning model for periportal fibrosis detection using US images.

Chapter 4 – CNN Model Implementation

Details the implementation of the Convolutional Neural Network (CNN) model, including the architecture, training process, and evaluation metrics.

Chapter 5 - Discussion and evaluation of the results

Interprets the results, discusses their implications, and compares them with existing literature. It also addresses the limitations of the study and suggests areas for future research.

Chapter 6 - Conclusions and Recommendations

Summarises the key findings, highlights the contributions of the study, and provides recommendations for clinical practice and future research.

3 Chapter 1 Introduction

Schistosomiasis mansoni, caused by the blood fluke Schistosoma mansoni, is a major public health concern, affecting approximately 54 million people annually, primarily in sub-Saharan Africa (Anderson & Enabulele, 2021). The infection leads to intestinal schistosomiasis, with pathological manifestations arising from the formation of granulomas around eggs that become lodged in the liver. This can lead to periportal fibrosis (PPF), a severe complication affecting a significant proportion of infected individuals, particularly in sub-Saharan Africa (Gunda et al., 2020). In Uganda, Schistosoma mansoni infection is highly endemic, with periportal fibrosis prevalent in communities along the shores of Lake Albert and Lake Victoria, affecting nearly 70% of the population (Natukunda et al., 2024).

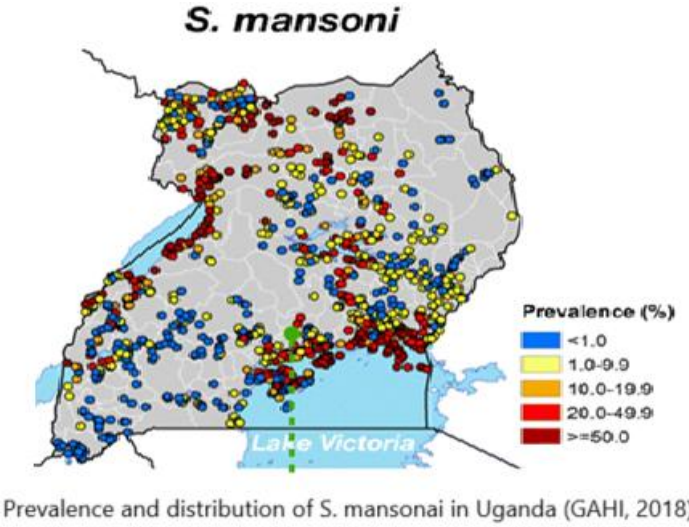


Figure 1. Prevalence and distribution of schistosomiasis in Uganda.

Periportal fibrosis is a common manifestation of chronic liver diseases and significantly impacts morbidity and mortality (Andrianah et al., 2020). Early detection of periportal fibrosis is crucial for timely intervention and the potential for reversibility.

Diagnostic Challenges

Currently, non-invasive diagnostic imaging methods such as ultrasound, CT, MRI, and elastography are used to assess and detect liver damage due to chronic schistosomiasis (Santos et al., 2022). However, these methods have limitations. Conventional scoring methods for liver fibrosis based on these imaging techniques are often time-consuming, subjective, and semi-quantitative, leading to variability in interpretation and potential diagnostic inaccuracies (Masseroli et al., 2002).

The Promise of Machine Learning

To address these challenges, this research explores the application of machine learning (ML) algorithms to enhance non-invasive diagnostic techniques for periportal fibrosis. Studies indicate that ML can support image analysis by autonomously extracting quantitative features associated with fibrosis. This approach aims to minimise operator dependency, improve diagnostic accuracy, and facilitate the earlier detection of periportal fibrosis, ultimately enhancing patient care and outcomes (Lee et al., 2020; Alzubaidi, 2022).

3.1 Research Question

Can machine learning algorithms enhance the accuracy and efficiency of analysing ultrasound images to detect periportal fibrosis in individuals with Schistosoma mansoni infection, compared to traditional image-based scoring methods?

3.2 Aim

To develop and evaluate a machine learning-based approach for the accurate and efficient detection of periportal fibrosis in individuals with Schistosoma mansoni infection, using non-invasive diagnostic imaging data, specifically ultrasound images.

3.3 Objectives

1. To develop and implement a Machine Learning based method for extracting relevant features from the ultrasound images that are indicative of periportal fibrosis periportal fibrosis.
2. To train and validate the model using the extracted features from image data to detect the presence of periportal fibrosis.
3. To evaluate the machine learning model's performance using metrics such as accuracy, precision, sensitivity, specificity, and efficiency.

4 Chapter 2 Literature Review

Periportal fibrosis (PPF) is a significant liver disease with substantial implications for patient health and well-being. Early and accurate detection of PPF is crucial for effective management and improved patient outcomes. This literature review explores the current state of knowledge regarding PPF detection, focusing on the potential of deep learning, specifically Convolutional Neural Networks (CNNs), to enhance diagnostic accuracy using non-invasive imaging modalities.

Literature Review Strategy

The literature was identified through a systematic search of several academic databases, including PubMed, IEEE Xplore, Google Scholar, and Elsevier's ScienceDirect and Scopus. This comprehensive search strategy involved using

specific keywords such as "periportal fibrosis," "diagnostic imaging," "machine learning," and "US images" to ensure a thorough and relevant collection of studies. The search aimed to gather a wide range of articles and papers that discuss the use of machine learning algorithms in the diagnostic imaging of periportal fibrosis, particularly in the context of ultrasound (US) images. By leveraging these databases, the search captured a diverse array of research findings, methodologies, and advancements in the field, providing a solid foundation for the study.

Scope of the Review

This review will cover the following key areas:

1. **Periportal Fibrosis**: Causes and Significance: Investigating the root causes and health impacts of PPF, highlighting the need for early detection.
2. **Diagnostic Methods**: Reviewing current non-invasive and invasive diagnostic methods for PPF, and their limitations.
3. **AI in Medical Imaging**: Summarising the growing role of AI in enhancing diagnostic accuracy and efficiency.
4. **Machine Learning Techniques**: Tracing the development of machine learning, leading to the rise of deep learning for image analysis.
5. **CNNs for PPF Detection**: Evaluating the use of CNNs for detecting PPF in medical images, reviewing studies and their findings.
6. **Future Research**: Identifying areas for future research, including more robust models, new imaging modalities, and integrating AI into clinical workflows.

4.1 Understanding Periportal Fibrosis

Periportal fibrosis (PPF) is a liver condition characterised by the excessive deposition of fibrous tissue around the portal veins. These veins are responsible for carrying blood

from the intestines to the liver. PPF can arise from various factors, including infections, metabolic disorders, and autoimmune diseases. These underlying causes trigger a cascade of events, leading to the accumulation of fibrous tissue around the portal veins, potentially disrupting blood flow and impairing liver function.

Clinical Significance

The clinical significance of periportal fibrosis lies in its impact on liver function and its potential to progress to more severe forms of liver disease, such as cirrhosis (Pinzani et al., 2011). As the fibrosis progresses, it can disrupt blood flow through the liver, leading to significant liver dysfunction and complications like portal hypertension. Portal hypertension is a condition in which the pressure in the portal vein system increases. Patients with periportal fibrosis may experience symptoms such as abdominal pain, fatigue, jaundice, and hepatomegaly (enlarged liver). Early detection and intervention are crucial to prevent the progression of fibrosis and to manage the underlying cause of liver injury.

Schistosomiasis as a Major Cause

Schistosomiasis, caused by the parasite Schistosoma mansoni, is a major cause of periportal fibrosis in endemic regions (Gunda et al., 2020). Despite other contributing factors, schistosomiasis remains a significant global public health concern and leading cause of this liver condition.

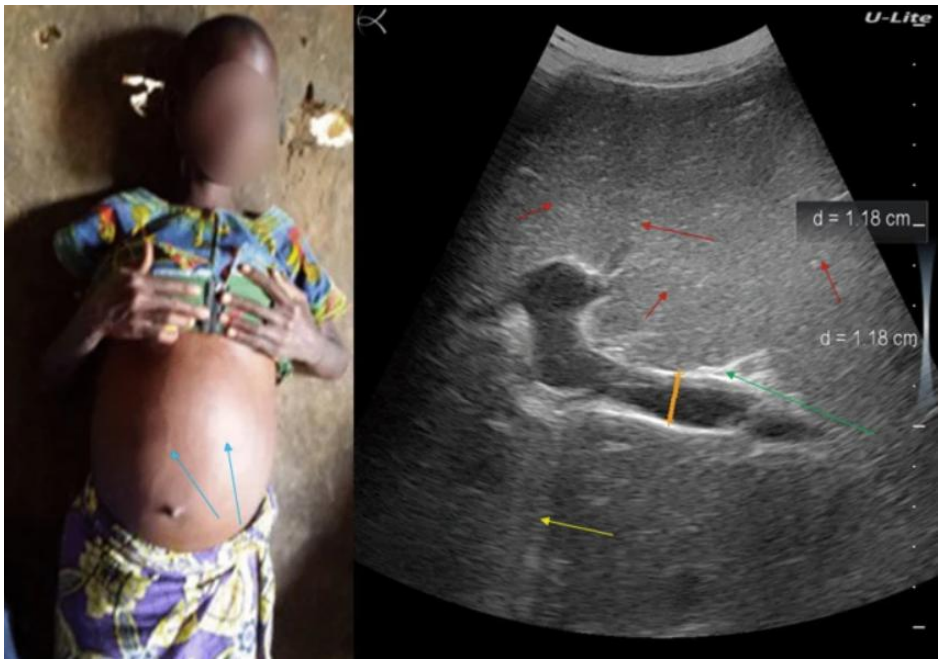


Figure 2. Ultrasound scan with heterogeneous liver architecture, decreased portal vein wall, with a positive schistosomiasis diagnosis (Nigo et al.,2021).

Pathogenesis of Periportal Fibrosis in Schistosomiasis

Infection with Schistosoma mansoni initiates a complex interplay between the parasite and the host's immune system, ultimately leading to the development of periportal fibrosis. Upon infection, adult worms reside in the mesenteric veins and release eggs. Some eggs are excreted, while others become trapped in the liver and intestines. These trapped eggs trigger a vigorous immune response, resulting in the formation of granulomas clusters of immune cells surrounding the eggs.

Chronic inflammation from granulomas leads to fibrous tissue deposition around portal veins, causing periportal fibrosis. This thickens and hardens periportal spaces in the liver, potentially disrupting blood flow and impairing liver function. The severity varies with infection intensity, duration, and host factors. Other causes include chronic hepatitis, alcoholic liver disease, and non-alcoholic fatty liver disease (Bedossa, 2017; Hudson et al., 2023).

4.2 Traditional Approaches to Periportal Fibrosis Detection

Traditionally, the detection of liver fibrosis, including periportal fibrosis, has relied on manual analysis of ultrasound (US) images.

This process involves several key steps:

1. **Image Acquisition**: High-quality US images of the liver are acquired using standardised protocols.
2. **Visual Assessment**: Experienced radiologists or clinicians then visually examine these images, searching for specific features suggestive of fibrosis. These features may include increased liver surface nodularity, periportal thickening, increased liver parenchymal echogenicity, attenuation of ultrasound beams, and the presence of nodules or masses (Friedrich-Rust et al., 2009).

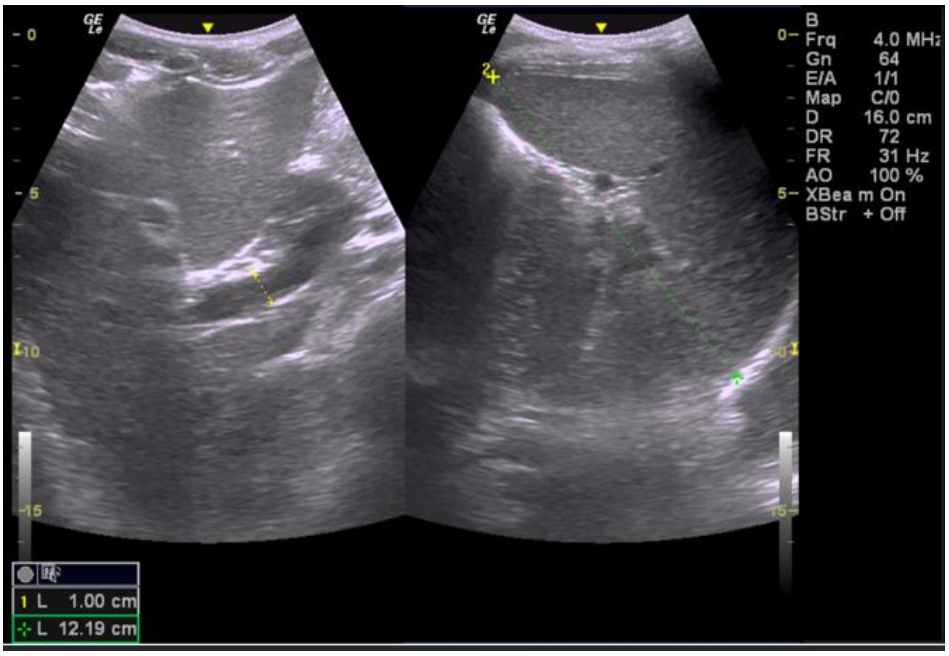


Figure 3: Sample Ultrasound Image of the Liver, Highlighting Key Anatomical Features for Fibrosis Assessment.

Based on this visual assessment, clinicians assign a qualitative grade to the severity of fibrosis, often using categories such as mild, moderate, and severe (Castéra et al., 2015).

4.3 Limitations of traditional imaging analysis

While conventional image analysis methods have proven valuable in assessing liver health, they have limitations in accurately detecting and staging periportal fibrosis (PPF). These limitations stem from the inherent subjectivity in interpreting imaging features and the reduced sensitivity of these methods in identifying the subtle signs of early-stage fibrosis.

Ultrasound

Ultrasound is a widely available and cost-effective imaging modality often used for initial liver assessments (Barr, 2017). However, its reliance on operator expertise and limited ability to detect early-stage fibrosis can present significant challenges (Friedrich-Rust et al., 2009). Accurately interpreting ultrasound images for fibrosis assessment demands a keen eye and extensive experience, as subtle textural changes and subtle increases in echogenicity can easily be missed by less experienced clinicians.

CT and MRI

A more thorough assessment is made possible by the more detailed visualisation of the liver's architecture provided by computed tomography (CT) and magnetic resonance imaging (MRI). With superior spatial resolution and tissue contrast compared to ultrasound, these modalities provide a clearer picture of liver morphology and potential abnormalities (Castéra et al., 2015). However, their high cost and limited

accessibility in resource-constrained settings hinder their widespread adoption, leaving many patients without access to these advanced imaging techniques.

Elastography

Elastography, a non-invasive method, evaluates liver fibrosis by measuring liver stiffness through shear wave speed. It's valuable but dependent on specialised equipment and can be influenced by operator experience and patient characteristics, leading to result variability (Barr, 2017; Friedrich-Rust et al., 2009).

Overall Limitations

In summary, traditional imaging analysis methods for periportal fibrosis detection face several challenges:

1. **Subjectivity**: Interpretation of imaging features often relies on the subjective assessment of the clinician, leading to potential inter-observer variability.
2. **Limited Sensitivity**: Traditional methods may not be sensitive enough to detect subtle changes associated with early-stage fibrosis, potentially delaying diagnosis and treatment.
3. **Accessibility and Cost**: Advanced imaging techniques like CT, MRI, and elastography can be expensive and have limited accessibility in resource-constrained settings.
4. **Operator Dependency**: The accuracy of traditional methods can be influenced by the operator's experience and expertise.

4.4 The Emergence of Artificial Intelligence in Medical Imaging

Consequently, there has been a growing drive to enhance diagnostic techniques by developing more objective, automated, and quantitative methods for assessing liver

fibrosis. These emergent methods aim to mitigate the subjectivity inherent in traditional approaches, enhance diagnostic efficiency, and provide more accurate and reproducible results, ultimately improving patient care and outcomes (Masseroli et al., 2002; Lee et al., 2020; Alzubaidi, 2022).

Artificial Intelligence and Machine Learning

The application of Artificial Intelligence (AI), particularly machine learning (ML) algorithms, is proving to be a promising approach in this field. ML can:

I. **Enhance Image Quality**: By reducing noise and artefacts, it improves the clarity of images for analysis.
II. **Automate Feature Extraction**: It automatically identifies and quantifies features indicative of liver fibrosis, facilitating more objective assessments.
III. **Improve Diagnostic Accuracy**: It leads to more precise and reliable diagnoses.
IV. **Minimise Operator Dependency**: It reduces reliance on subjective interpretation, enhancing consistency and reproducibility in fibrosis staging.

These ML capabilities aim to overcome the limitations of traditional methods, ultimately improving patient care and outcomes (Lee et al., 2020; Alzubaidi, 2022).

4.5 The Role of Artificial Intelligence in Liver Fibrosis Assessment

AI has significantly impacted modern medical imaging research, enhancing both diagnostics and therapeutics. As Du-Harpur et al. (2020) describe, AI mimics human intelligence using machines, particularly computer systems. It includes technologies such as machine learning (ML), natural language processing (NLP), and computer vision, enabling machines to perform tasks requiring human intelligence, like pattern recognition, prediction, and adaptation.

Within the healthcare domain, AI has found widespread applications, enhancing efficiency and accuracy across numerous fields. In medical imaging, AI plays a crucial role in analysing medical images, aiding in disease detection, and supporting clinical decision-making (Du-Harpur et al., 2020). AI's transformative potential is particularly evident in areas like sonography, where it assists in interpreting complex ultrasound images and guiding diagnostic assessments.

4.6 Machine Learning in Medical Imaging

Machine learning (ML), a branch of AI, enables computers to recognise data patterns for prediction and decision-making (Lee et al., 2017). Imagine a computer analysing numerous medical images, learning to identify subtle liver disease signs without explicit programming. This is ML, allowing computers to learn from data and improve their performance over time.

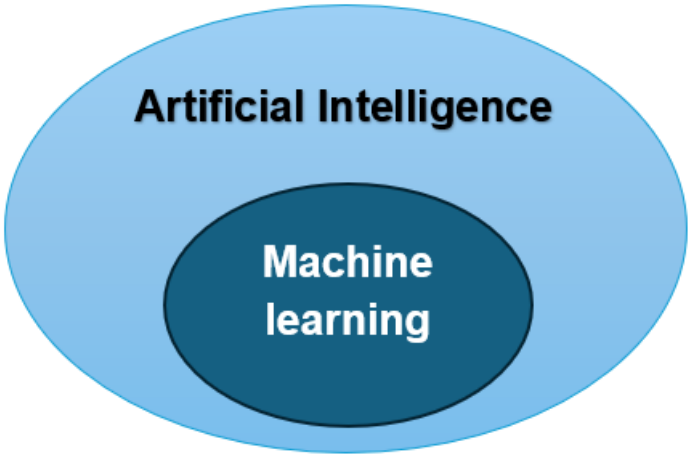


Figure 4: Illustration of Machine learning (ML) as subset of AI

4.7 Traditional Machine Learning techniques

Traditional machine learning techniques have been instrumental in advancing medical imaging, supporting tasks such as classification, segmentation, and registration (Yamashita et al., 2022). Typically, these methods involve manual feature extraction,

where domain experts carefully identify and select relevant features from images, such as texture, shape, or intensity, that signify the condition being analysed.

Manual Feature Extraction and Classification

In the field of medical imaging, expert radiologists carefully identify important visual features, such as the irregular shape of a tumour or the unique textural patterns of a fibrous lesion. These identified features are then fed into machine learning algorithms like Support Vector Machines (SVMs), which help automate the classification or segmentation of these images (Khan et al., 2020; Iwendi et al., 2020).

4.8 Limitations of Traditional Machine Learning Approaches

Traditional machine learning algorithms, while valuable, face inherent limitations, particularly in the context of complex medical images. As Alzubaidi (2022) points out, these algorithms rely heavily on manual feature extraction and selection, a process that is inherently subjective and sensitive to variations in image appearance, such as size, texture, colour, and shape. This dependence on manual feature engineering can be tedious and prone to bias, as different experts may select different features, leading to inconsistencies in results (Yasaka et al., 2020).

Moreover, traditional methods often struggle with the high dimensionality and intricate patterns found in medical images (Zhou et al., 2021). Imagine trying to find a specific grain of sand on a vast beach; traditional algorithms might miss subtle but crucial details that could indicate the presence of a disease. Their limited ability to capture complex patterns and their sensitivity to variations in image appearance hinder their performance in challenging medical image analysis tasks.

In tasks like image recognition, even powerful algorithms like Support Vector Machines (SVMs) can be less efficient in feature extraction compared to newer approaches (Lai, 2019).

The following section explores this evolution from traditional methods to deep learning, highlighting the advantages of DNNs and their transformative impact on medical image analysis.

4.9 The Rise of Deep Learning in Medical Image Analysis

While traditional machine learning techniques have made significant contributions to medical image analysis, their limitations have led to the development of more advanced methods, notably deep learning.

Deep Learning

Deep learning (DL), a subfield of machine learning, has revolutionised medical image analysis. Inspired by the human brain, DL algorithms overcome limitations and reveal new possibilities in medical imaging (Tang, 2019). DL automates feature extraction and handles the complexity of medical image data, unlike traditional methods dependent on manual feature engineering. DL algorithms autonomously learn complex patterns and representations from raw image data, identifying subtle details and relationships that traditional techniques might miss (Alzubaidi, 2022).

4.10 Why Deep Learning?

Deep neural networks (DNNs) excel in automatically learning complex patterns and features from raw image data, outperforming traditional machine learning methods in numerous applications.

DNNs as Digital Assistants

DNNs are increasingly employed as digital assistants in the early diagnosis of liver diseases, demonstrating an impressive ability to detect disease, often rivalling or even surpassing the expertise of specialists (Tang, 2019). This remarkable performance stems from their ability to:

1. **Detect Complex Patterns**: Deep neural networks (DNNs) can autonomously recognise complex patterns and subtle features within medical images that might be challenging for human experts to detect.
2. **Provide Precise Measurements**: DNNs can provide accurate quantitative measurements of anatomical structures and tissue characteristics, enabling more objective and reliable assessments.
3. **Automate Feature Extraction**: Unlike traditional methods, DNNs eliminate the need for manual feature engineering, saving time and reducing subjectivity.

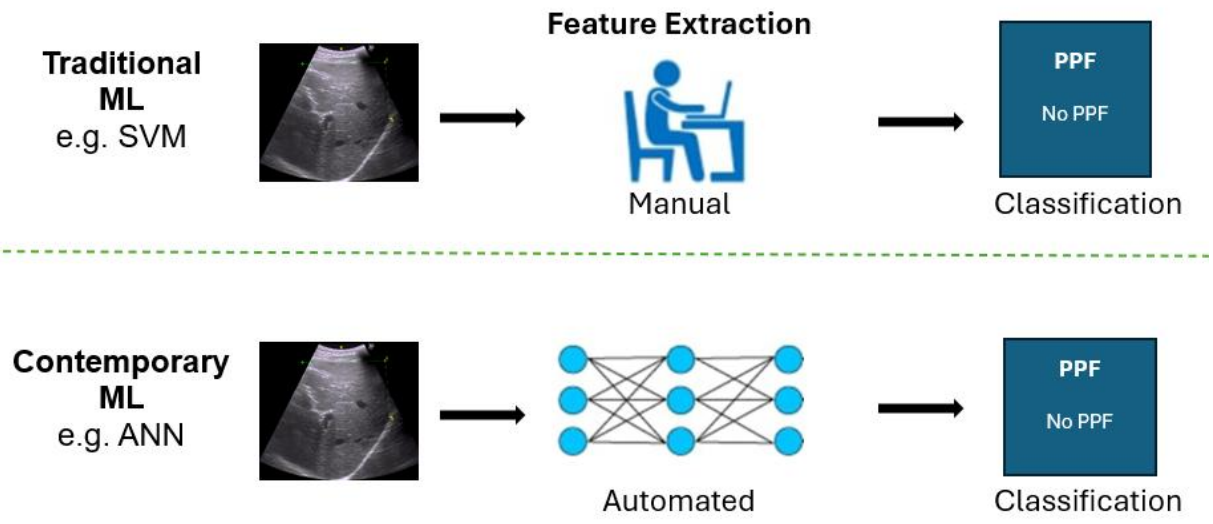


Figure 9: illustrates the contrast in feature selection approaches between traditional and DNNs.

4.11 Deep Neural Networks

Brain-Inspired Computing: Mimicking the Human Brain

The human brain, an organ of unparalleled complexity, possesses the remarkable ability to process vast amounts of information, recognise intricate patterns, and make decisions with astonishing efficiency. The fundamental building block of this intricate network is the neuron, a specialised cell capable of transmitting information through electrical and chemical signals. Neurons are interconnected in a vast and complex web, forming the basis for all cognitive functions, including learning, memory, and perception (Kandel et al., 2013).

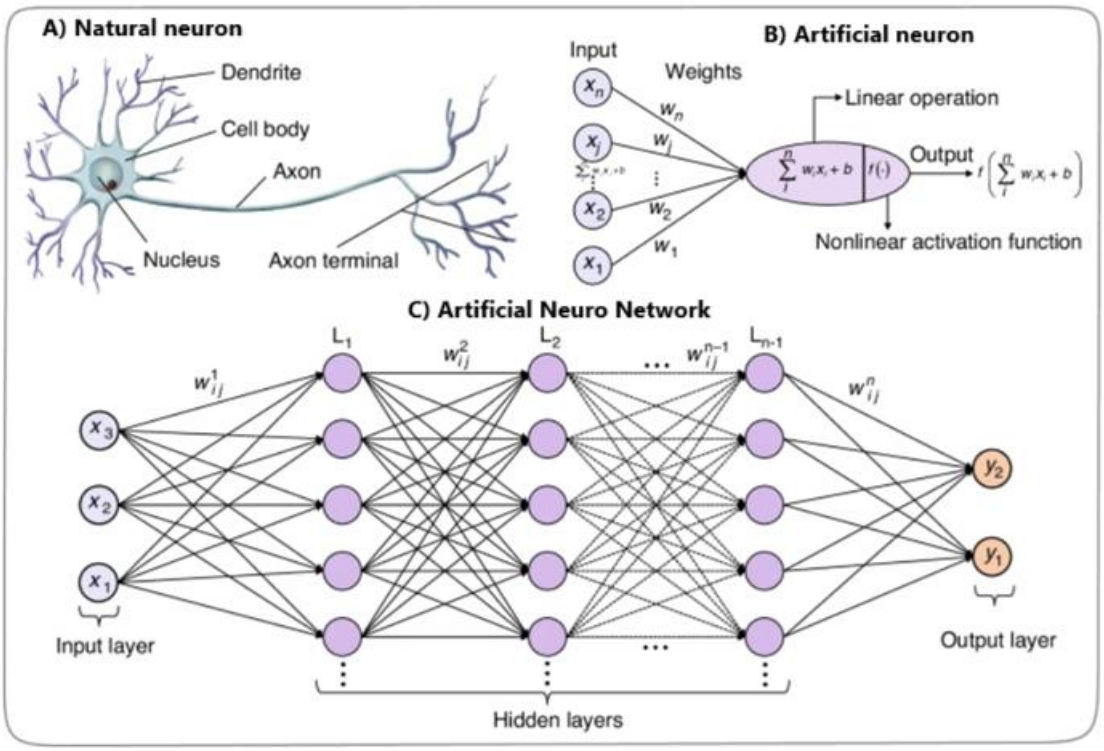


Figure 5: Displays an artificial neuro network, inspired by the functioning of the brain (Fu et al. 2024).

DNNs are inspired by this intricate network of interconnected neurons. They are composed of layers of interconnected nodes, or artificial neurons, organised in a hierarchical structure resembling the organisation of the human brain. These nodes,

much like their biological counterparts, receive input signals, process them, and transmit output signals to other nodes in the network (Kufel et al., 2023).

Hierarchical Structure and Learning

DNNs have an intricate network of connections that allows them to understand complex relationships between inputs and outputs, excelling in tasks like image recognition and natural language processing. Their hierarchical structure helps them progressively learn complex features from data, similar to the human brain.

Comparing Biological and Artificial Neural Networks

While both biological and artificial neural networks share a fundamental structure of interconnected processing units and hierarchical organisation, they differ significantly in their underlying mechanisms and capabilities.

4.11.1 Biological Neural Networks

Biological neural networks, exemplified by the human brain, rely on neurons and synapses for signal transmission and learning. Learning occurs through synaptic plasticity, a process of strengthening or weakening connections between neurons based on experience (Mishra & Srivastava, 2014). These networks exhibit:

- **Adaptability**: They effortlessly learn new tasks and generalise knowledge across domains.
- **Energy Efficiency**: They operate with remarkable energy efficiency.
- **Fault Tolerance**: They are highly resilient to errors and damage.

4.11.2 Artificial Neural Networks

In contrast, artificial neural networks, such as deep learning models, employ nodes and weighted connections, relying on numerical values and backpropagation for learning. These networks are typically:

- **Task-Specific**: Trained for specific tasks and may struggle to adapt to new situations.
- **Less Energy Efficient**: Require significantly more energy to operate compared to biological networks.
- **Less Fault Tolerant**: More susceptible to errors and disruptions.

Inspiration and Innovation

Despite these differences, the inspiration drawn from biological neural networks has been instrumental in the development of powerful deep learning models. While artificial networks still have limitations compared to their biological counterparts, they are rapidly evolving and transforming medical imaging and other fields.

Table 1: A summary table comparing biological neural networks and artificial neural networks

Feature	Biological Neural Network	Artificial Neural Network
Basic Unit	Neuron	Node (Artificial Neuron)
Connection	Synapses	Numerical values
Signal Transmission	Electrical and Chemical	Numerical values
Learning Mechanism	Synaptic plasticity (changes in synapse strength)	Backpropagation (adjusting weights)
Architecture	Complex, layered, interconnected regions	Layered, interconnected nodes

The Power of Interconnected Systems and the Training of DNNs

Despite these inherent differences, both biological and artificial neural networks showcase the remarkable power of interconnected systems for information processing and learning, paving the way for advancements in both biological and artificial

intelligence. This convergence of inspiration and innovation fuels the ongoing quest to understand and replicate the intricacies of the human brain.

Training Deep Neural Networks

The key to the success of DNNs lies in their ability to learn from data through a process called training. During training, the network is presented with a vast dataset of examples, and the connections between nodes, known as weights, are iteratively adjusted to minimise errors in prediction.

Parallels and Limitations

Overall, the parallels between Deep Neural Networks (DNNs) and the human brain are striking, particularly in their hierarchical organisation, distributed processing, and adaptive learning capabilities. However, it's crucial to acknowledge that DNNs remain a simplified model of the brain's vast and intricate complexity. While DNNs have achieved remarkable success in various domains, they are yet to fully replicate the inherent flexibility and adaptability of human cognition the effortless way humans learn and generalise knowledge across diverse contexts.

Transformative Impact

Despite these limitations, DNNs have undeniably revolutionised artificial intelligence, paving the way for significant advancements in fields like medical imaging, robotics, and natural language processing. Their capacity to learn complex patterns from data has opened new frontiers in automation, decision-making, and problem-solving, promising a future where intelligent machines can augment human capabilities and drive innovation across various industries.

4.12 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a type of deep neural network that can autonomously learn features from data, removing the necessity for manual feature engineering. This ability makes CNNs especially apt for medical image analysis.

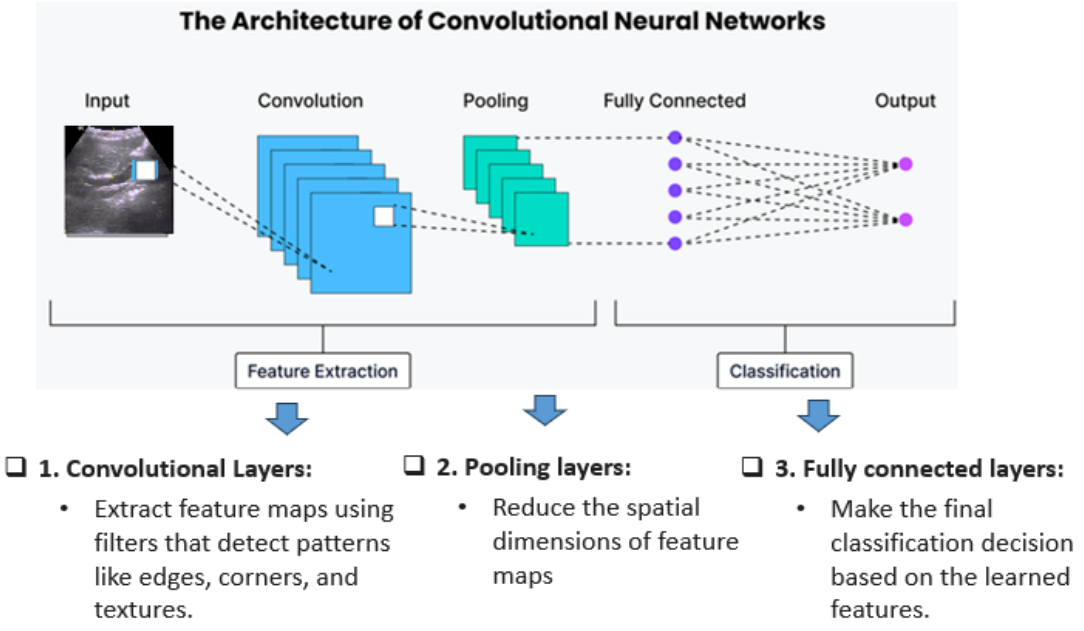


Figure 6: Illustration of a CNN architectural with hierarchical layers

Prominent CNN Architectures and Advantages

Several widely recognised CNN architectures have emerged as powerful tools for medical image analysis, including:

- **AlexNet**: One of the early breakthroughs in deep learning for image recognition, known for its relatively simple architecture and significant performance improvement over traditional methods.
- **VGGNet**: Known for its deeper architecture and use of smaller convolutional filters, leading to improved feature extraction and accuracy.
- **GoogLeNet**: Was a significant breakthrough in deep learning, particularly in image classification and detection tasks. The network is deeper than its

predecessors, with 22 layers in total, and it effectively addresses the vanishing gradient problem, which is common in deep networks.

- **ResNet**: It was designed to address the vanishing gradient problem, which often occurs when training very deep neural networks. ResNet utilises residual learning, where the network learns residual functions with reference to the layer inputs, rather than learning unreferenced functions.
- **DenseNet**: Further enhanced the connectivity between layers by connecting each layer to all subsequent layers, promoting feature reuse and improving efficiency.

Backpropagation and Weight Adjustment

Backpropagation is a fundamental algorithm for training neural networks, including CNNs. It stands for "backward propagation of errors" and involves two phases: forward pass and backward pass. This iterative process fine-tunes the weights of the

connections between nodes in the network based on prediction errors (Mishra & Srivastava, 2014).

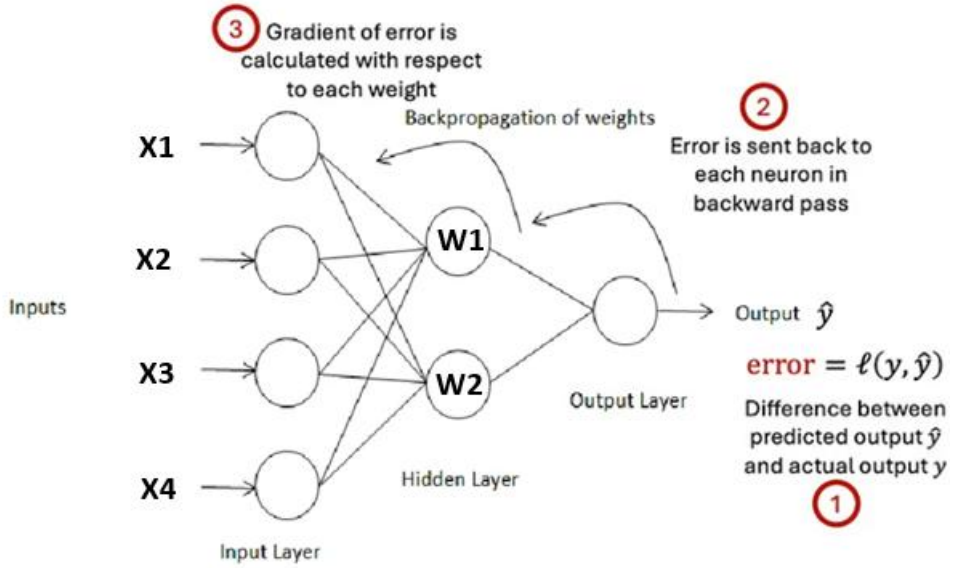


Figure 7: Illustration of backpropagation concept in a CNN

The Feedback Loop of Learning

The feedback loop of learning involves presenting the network with a large dataset of images and labels. The network makes predictions and uses backpropagation to adjust the weights if predictions are incorrect. This process, akin to feature selection, prioritises relevant features for classification (LeCun et al., 2015). The network enhances its predictions by assigning higher weights to better features and lower weights to less relevant ones. Through repeated iterations, the network optimises its performance for accurate image classification.

Essential for Accuracy

Backpropagation is crucial for CNNs to achieve high accuracy in image analysis tasks. It enables the network to learn and prioritise the most relevant features from the data, resulting in more robust and accurate predictions.

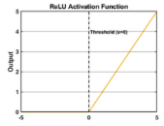
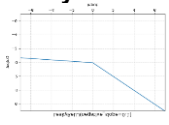
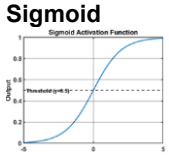
The Role of Activation Functions in CNNs

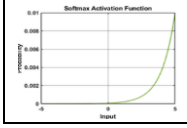
Activation functions in neural networks introduce non-linearity, allowing the model to learn and represent complex patterns. Without activation functions, CNNs would only perform linear transformations, limiting their ability to model intricate relationships in medical images. Non-linearity is essential for capturing complex features indicative of periportal fibrosis. Activation functions provide this non-linearity, allowing the network to detect subtle patterns crucial for accurate disease detection and classification. They act as switches, determining neuron activation based on the input (Dubey, 2023).

Activation Functions in Image Classification

For a comprehensive understanding, Table 2 presents commonly used activation functions in image classification tasks, alongside their specific applications and supporting literature:

Table 2. Application of Activation functions in CNNs

Function	Description	Application
ReLU 	Is a popular activation function in neural networks. It outputs the input if positive, otherwise zero, introducing non-linearity and improving computational efficiency.	Convolution layers
Leaky ReLU 	It overcomes a limitation of the ReLU by permitting small, non-zero gradient for negative inputs, thereby preventing inactive neurons and potentially enhancing performance.	Convolution layers
Sigmoid 	The sigmoid function maps input to a value between 0 and 1, useful for binary classification.	Output layer (Binary classification

Softmax 	Used in the output layer of neural networks, converts raw output scores into probabilities by exponentiating each score and normalising them to sum to 100%. This is useful for multi-class classification problems.	Output layer (multi-class classification)

4.12.1 Convolutional Layers: Extracting Features from Images

Convolutional layers in CNNs use filters to scan input images, creating feature maps that capture elements like edges and textures. They detect local patterns and build them into complex structures, with stride and padding controlling output size.

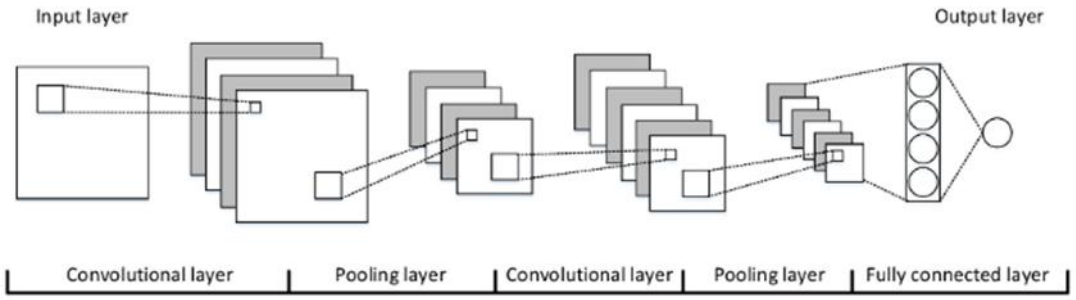


Figure 8: Illustration of a CNN Architecture, depicting the interconnected layers and flow of Information.

Functionality

Convolutional layers function by applying filters to the input image. These filters, which are small matrices of numerical values, are trained to detect specific patterns like edges, corners, and textures. When a filter is applied to a section of the image, it

performs convolution, a mathematical operation that produces a new value indicating the presence or absence of the pattern in that section.

Feature Maps

The outputs of a convolutional layer are feature maps, each indicating the presence of specific patterns throughout the image. These feature maps encapsulate crucial information, allowing the CNN to comprehend the image's content.

Benefits of Convolution

Convolutional operations offer several key advantages in image analysis:

1. **Edge Detection**: Convolutional filters can effectively identify the boundaries of objects within an image.
2. **Noise Reduction**: Convolution can help reduce unwanted noise or variations in an image, improving its clarity for further processing.
3. **Pattern Recognition**: Convolutional layers are adept at recognising recurring patterns, which is crucial for tasks like identifying objects or features in medical images.

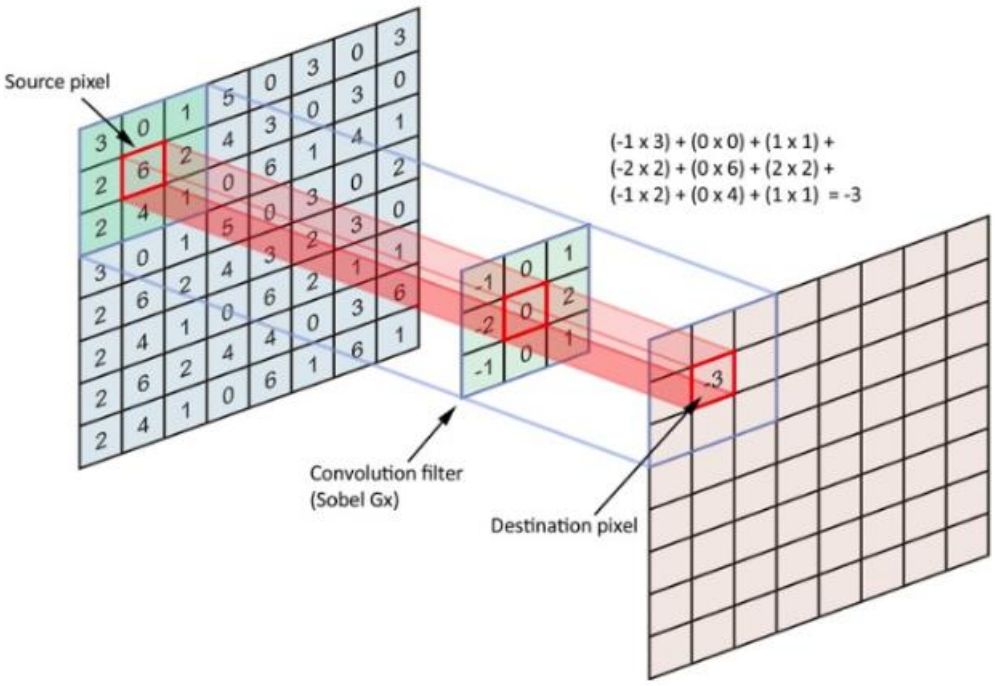


Figure 9: An illustration of a CNN convolution.

Hierarchical Feature Learning in Convolutional Layers

Convolutional layers in a CNN are organised hierarchically, allowing the network to progressively learn more complex features as information passes through each layer.

Early Layers

In the initial stages, convolutional layers extract fundamental features like edges, corners, and textures from the input image. These basic features act as building blocks for more complex representations.

Deeper Layers

As information progresses through deeper layers, the network combines the basic features extracted by earlier layers to identify more complex patterns and shapes. This hierarchical learning process allows the CNN to develop a deeper understanding of the image content.

Significance in Medical Imaging

This hierarchical approach is particularly valuable in medical image analysis, where subtle details can be crucial for accurate diagnosis. By progressively learning more complex features, CNNs can identify patterns indicative of diseases or abnormalities that might be missed by simpler methods.

Transforming Pixels into Meaningful Information

Convolution is a fundamental operation in CNNs that transforms raw pixel data into meaningful representations that computers can understand.

Pixel Processing

Convolution applies filters, which are small matrices of learned numerical values, to the input image to detect specific patterns in pixel data. This process generates feature maps highlighting the presence and location of these patterns across the image.

Applications in Medical Imaging

Convolution is essential for various medical imaging tasks, such as:

I. **Identifying anatomical structures**: Convolutional layers can help detect and delineate organs, tissues, and other anatomical features in medical images.
II. **Detecting abnormalities**: CNN can be trained to recognise patterns associated with diseases or abnormalities, aiding in diagnosis and treatment planning.
III. **Image segmentation**: Convolutional operations are used to segment medical images, separating different regions of interest for further analysis.

4.12.2 Pooling layers

Dimensionality Reduction and Feature Emphasis

Pooling layers are an essential component of CNNs, reduce the spatial dimensions of feature maps and emphasising important features. This process, known as

dimensionality reduction, helps to improve computational efficiency and prevent overfitting (Springenberg et al., 2014).

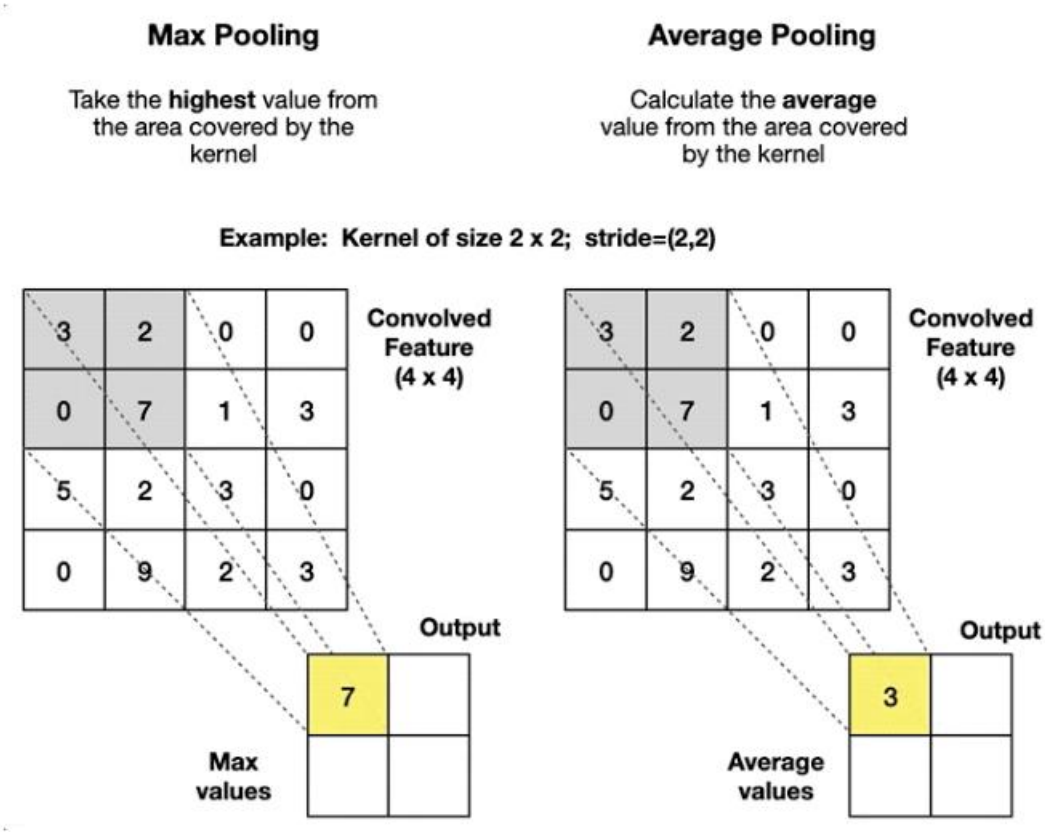


Figure 10: Illustration of CNN Max pooling vs Average pooling

Pooling operations come in various types, with two prevalent ones being:

1. **Max Pooling**: Max pooling reduces the spatial dimensions of feature maps by sliding a window over the input and taking the maximum value. This process helps decrease parameters and computational complexity while retaining important features, improving the model's focus on the most salient aspects of the input data.
2. **Average Pooling**: Average pooling reduces the spatial dimensions of feature maps by sliding a window over the input and taking the average value. This downsampling process helps decrease parameters and computational complexity while retaining important features.

Benefits in Periportal Fibrosis Detection

In the context of periportal fibrosis detection research, pooling layers help to:

- **Extract relevant features**: By reducing the dimensionality of feature maps, pooling layers emphasise the most important features, such as measurements of the periportal space, texture features, and other biomarkers indicative of the disease.
- **Create a compact representation**: Pooling layers produce a more compact and discriminative representation of the image, enabling the network to focus on essential information for accurate diagnosis.
- **Improve computational efficiency**: By decreasing the number of parameters, pooling layers contribute to good performance and faster training and inference for the network.
- **Prevent overfitting**: By reducing the complexity of the model, pooling layers help to prevent overfitting.

This approach aligns with the methods discussed by Litjens et al. (2017) in their comprehensive survey, where pooling layers are highlighted as a key component for feature extraction and dimensionality reduction in medical image applications.

4.12.3 Fully Connected Layer

Integrating Features for Classification

The Fully Connected (FC) layer is an essential part of CNNs, combining the features extracted by convolutional and pooling layers to carry out the final classification.

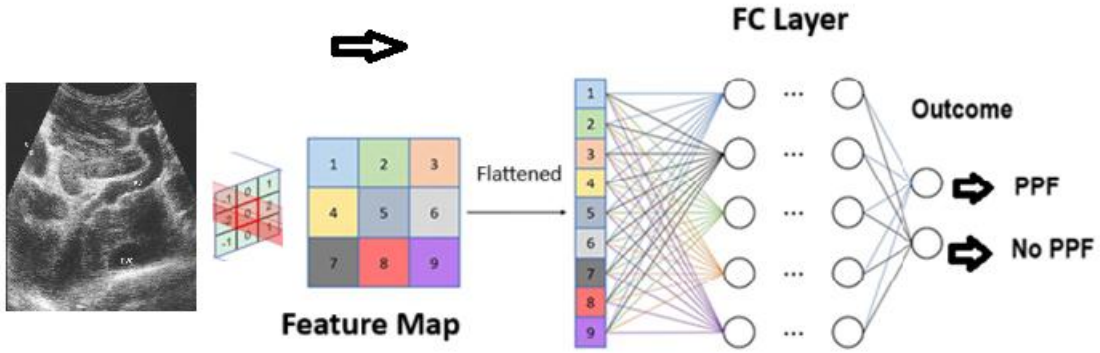


Figure 11. Illustration of a fully connected layer for CNN classification task

Functionality

In the FC layer, each neuron is connected to every neuron in the previous layer, allowing it to consider all the extracted features comprehensively. This dense connectivity enables the FC layer to learn complex relationships between features and their corresponding classes.

Role in Periportal Fibrosis Detection

In the context of periportal fibrosis detection, the FC layer performs the following functions:

- **Feature Integration**: It combines the features extracted by earlier layers, such as measurements of the periportal space, texture features, and other

biomarkers, to create a comprehensive representation of the input ultrasound image.

- **Classification**: Based on the integrated features, the FC layer determines the likelihood of the image belonging to different classes, such as the presence or absence of fibrosis or different stages of the disease.
- **Output Probabilities**: The output of the FC layer is typically passed through a SoftMax or sigmoid activation function to produce probabilities for each class, indicating the network's confidence in its prediction.

Research by Han et al. (2023) highlights the importance of the FC layer in the final stages of CNNs, where it transforms extracted features into a format suitable for classification. Basha et al. (2020) further emphasise the role of the FC layer in categorising medical images for diagnostic purposes, including the detection of periportal fibrosis.

5 Conclusion

The rise of CNNs has revolutionised medical image analysis, offering automated, accurate, and efficient diagnostic solutions. This technological advancement holds great potential for transforming disease detection and management by identifying complex patterns within medical images.

Key Findings and Contributions

This research has demonstrated the efficacy of CNNs in automating the detection and classification of medical conditions. By harnessing the power of deep learning, these models have achieved promising results in accurately identifying and staging diseases, highlighting their potential to reshape diagnostic workflows.

Future Directions

Deep learning research continues to advance in areas such as:

I. **Multi-modal data integration**: Combining various imaging modalities (e.g., ultrasound, CT, MRI) to improve diagnostic accuracy.
II. **Robust and explainable models**: Creating CNN models that are resilient to data variations and offer insights into their decision-making.
III. **Personalised treatment strategies**: Using CNNs to tailor treatment plans based on individual patient characteristics and disease progression.

These advancements enhance the detection, management, and potential prevention of diseases like periportal fibrosis, leading to better patient care and health outcomes.

Research Impact

This comprehensive review has revealed several crucial points:

1. CNNs represent a fundamental shift in medical image analysis, offering unprecedented capabilities for automated diagnosis and disease management.
2. Research has consistently demonstrated the effectiveness of CNNs in accurately detecting and classifying medical conditions.
3. Deep learning research continues to push the boundaries, exploring new applications and improvements in CNN-based medical image analysis.

These insights have directly informed the development of the machine learning model presented in this research and guided the overall research direction.

Bridging Knowledge to Methodology

Building on this foundation, the next chapter will outline the development and evaluation of a CNN model for automated periportal fibrosis detection.

6 Chapter 3 – Methodology

6.1 Dataset Source

This research employs a comprehensive liver ultrasound image dataset from a study conducted by the Uganda Schistosomiasis Multidisciplinary Research Centre (U-SMRC), a leading Ugandan institution in schistosomiasis research (U-SMRC, 2022). The dataset originates from an adult case-control study investigating risk factors for severe schistosomal morbidity in communities near Lake Victoria and Lake Albert, two distinct epidemiological settings. The study individuals were aged between 18 to 50 years. Ultrasound sonography, a common diagnostic technique, was performed using the Logiq e ultrasound system.

Figure 12: Sample Ultrasound image of the liver captured using the Logiq e ultrasound system.

6.2 Ultrasound Image Annotation and Case Selection

The liver ultrasound images were meticulously annotated by an experienced study sonographer using the Niamey protocol (Akpata et al., 2015).

a) The Niamey Protocol

The World Health Organisation's (WHO) Niamey Protocol, a standardised ultrasonography protocol, is fundamental for assessing schistosomiasis-related morbidity, particularly hepatic morbidity caused by Schistosoma mansoni (WHO, 2000). Widely accepted and used in field studies, it facilitates data comparison across endemic areas (El Scheich et al., 2014). The protocol uses ultrasound to detect liver and portal vein abnormalities, ensuring consistency and comparability. Moreover, it has significantly improved diagnosis and management of schistosomiasis-related liver disease, even in resource-limited settings (Akpata et al., 2015).

b) Liver Image Pattern Score

The Liver Image Pattern Score is a key part of the Niamey protocol, a standardised method for assessing liver damage in people with schistosomiasis. This score, as explained by Ockenden et al. (2024), helps doctors determine how severe the periportal fibrosis (scarring around the liver's blood vessels) is.

How it works:

Doctors use ultrasound images of the liver to determine the liver image pattern score. They look for specific signs of damage, such as:

- **Bumpy liver surface (nodularity)**: A healthy liver has a smooth surface. Nodularity indicates the presence of abnormal bumps or lumps.
- **Thickening around blood vessels (periportal thickening)**: Fibrosis causes the tissues around the liver's blood vessels to become thicker.
- **Increased brightness of liver tissue (parenchymal echogenicity)**: Changes in the tissue's ability to reflect sound waves can indicate damage.
- **Weakening of ultrasound signals (attenuation)**: Scar tissue can block or weaken the ultrasound signals, making it harder to see the liver's structures.

- **Nodules or masses**: These abnormal growths within the liver can also be a sign of advanced fibrosis.

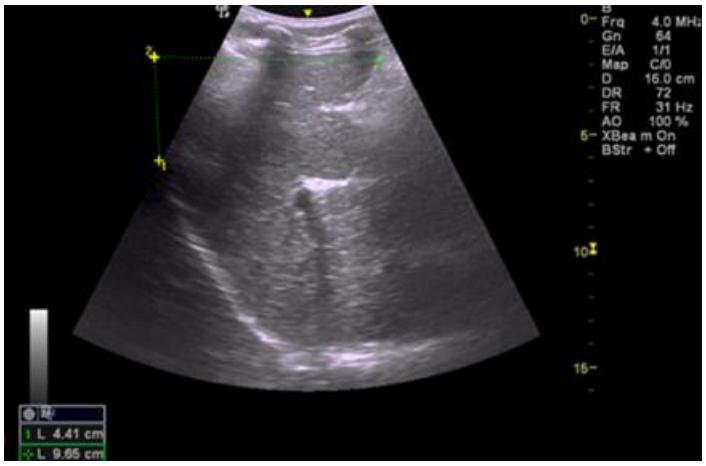


Figure 13. Sample liver ultrasound scan measurements using Logiq e system.

Systematic evaluation of these features enables sonographers to assign a Liver Image Pattern (LIP) score, quantifying liver condition. These scores are then recorded in an electronic database using REDCap, facilitating diagnosis, progression tracking, and treatment decisions for schistosomiasis-related liver damage.

Detection of periportal fibrosis using image pattern score

An image pattern score of 2 or greater, as defined by the Niamey protocol (Ockenden et al., 2024), was used to indicate the presence of specific ultrasound patterns characteristic of periportal fibrosis.

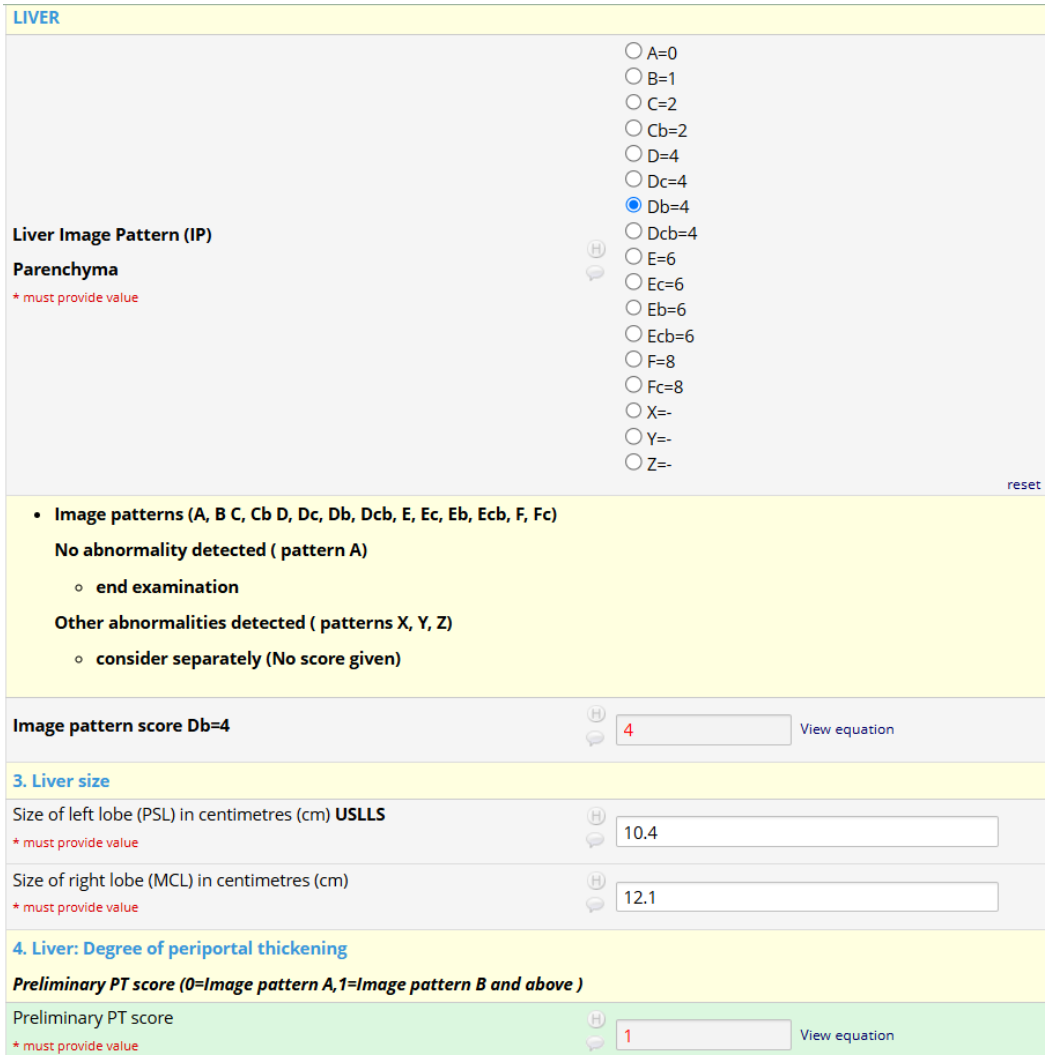


Figure 14:Screenshot illustrating data entry for the liver image pattern score within the REDCap system

Sample Size

This research utilised a total of 200 randomly selected ultrasound images of adult participants aged 18 to 50 years, comprising a balanced dataset of 100 images representing cases (individuals with periportal fibrosis) and 100 images representing controls (individuals without severe liver disease).

6.3 Ethical Considerations, Data Compliance, and Confidentiality

This research was conducted with meticulous attention to ethical principles, data compliance, and the confidentiality of USMRC study participants.

1. **Data De-identification**

Safeguarding participant privacy was paramount. All ultrasound images used in this study underwent de-identification, ensuring the removal of any information that could potentially identify individuals, such as names, dates of birth, or medical record numbers. This practice aligns with ethical standards and regulatory requirements, including the Uganda Data Privacy and Protection Act (Government of Uganda, 2019) and the General Data Protection Regulation (GDPR, 2016).

2. **Informed consent**

This project employed a retrospective study design, utilising secondary data obtained from the USMRC research study. All participants in the original USMRC study provided informed consent for the use of their ultrasound images for research purposes (Lee et al., 2020). This prior consent provided an ethical foundation for secondary data analysis.

3. **Ethical Approval and Proposal**

To further ensure ethical rigour, approval for this study was sought and obtained from the University of Essex's ethics committee. A comprehensive proposal detailing the study's aims, methodology, data management procedures, and potential ethical considerations was submitted for review. This proposal included formal consent from the USMRC principal investigators, granting permission to use the de-identified ultrasound images for this project.

6.4 Data extraction and Pre-processing

Data acquired from the ultrasound scanner were exported in the Digital Imaging and Communications in Medicine (DICOM) format, a widely adopted standard for storing,

transmitting, and handling medical imaging information (Larobina, 2023). DICOM files contain both image data and essential metadata, including patient demographics, imaging parameters, and other relevant details, ensuring comprehensive information is stored alongside the images.

1. DICOM File Retrieval and Processing

After exporting the DICOM files from the ultrasound system, we used 3D Slicer, a powerful, free, and open-source software platform designed for visualising, processing, and analysing medical images (Fedorov, 2012; 3D Slicer, 2024). 3D Slicer provided the necessary tools for accessing, viewing, and manipulating the DICOM data, enabling us to extract the relevant image features for our machine learning models.

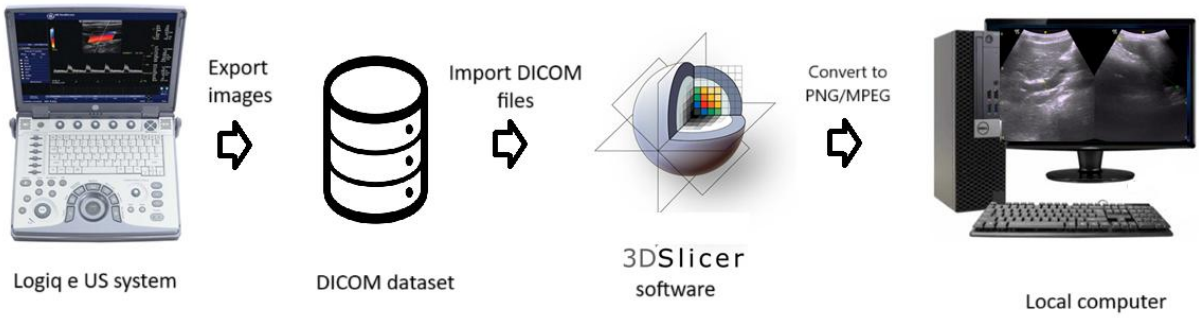


Figure 15: Pipeline illustrating the extraction and storage of ultrasonography (US) images using 3D Slicer software.

2. Image Exportation and De-identification

Using 3D Slicer, the retrieved ultrasound (US) scans were exported and saved in PNG format on a local computer. Crucially, this exportation process included the removal of metadata or embedded information containing any identifiable details, such as patient names, birth dates, and study record numbers. This de-identification process ensured

compliance with ethical standards and data privacy regulations, safeguarding the anonymity of USMRC study participants.

3. Image Pre-processing

The image pre-processing stage encompassed several crucial steps to prepare the data for model training and enhance model performance. These steps included image renaming, loading, resizing, normalisation, and augmentation, each contributing to data quality and generalisation.

4. Image Renaming and Anonymization

To ensure participant confidentiality, original study IDs, which could potentially be linked to individual identities, were replaced with new, anonymized IDs. This renaming process ensured that the images could not be traced back to specific individuals, safeguarding their privacy. The de-identified images were then securely stored on a local machine with restricted access to authorised personnel only.

```
# Define the folder path
folder_path = 'D:/SCHULE/MSc Research Project/DEVELOPMENT/Data'

class_name = 'fibrosis'

# Get list of image files in the folder (filtering for common image types)
image_files = [f for f in os.listdir(folder_path) if f.endswith(('.png', '.jpg', '.jpeg'))]

# Rename each image by appending the class name at the beginning
for image_file in image_files:
    old_path = os.path.join(folder_path, image_file)

    # Create new file name with class name prepended
    new_file_name = f"{class_name}_{image_file}"
    new_path = os.path.join(folder_path, new_file_name)

    # Rename the file
    os.rename(old_path, new_path)
```

Code Renaming images to categorise images for machine learning.

Figure 17 illustrates the code snippet used to rename image files, adding prefixes ("fibrosis_" or "nofibrosis_") to their filenames to categorise them for machine learning tasks. This organisation is crucial for efficient model training and data management.

5. Image Loading and Resizing

Images were initially loaded from their respective subfolders, representing the different classes (fibrosis and no fibrosis). To maintain aspect ratios and prevent distortion during resizing, a technique was employed where images were resized to a target dimension (initially set to 32x32 pixels) while preserving their original aspect ratios. If the resized image was smaller than the target size, padding was applied to ensure uniformity and consistency across the dataset (Russ, 2011).

```python
def load_images_from_folder(folder,image_size):
    images = []
    labels = []
    for subfolder in os.listdir(folder):
        subfolder_path = os.path.join(folder, subfolder)
        if os.path.isdir(subfolder_path):  # Check if it's a directory
            label = subfolder
            for filename in os.listdir(subfolder_path):
                img_path = os.path.join(subfolder_path, filename)
                try:
                    img = cv2.imread(img_path)

                    # Get original image dimensions
                    height, width = img.shape[:2]

                    # Calculate aspect ratio
                    aspect_ratio = width / height

                    # Resize while maintaining aspect ratio
                    if aspect_ratio > 1:  # Width > Height
                        new_width = image_size[0]
                        new_height = int(new_width / aspect_ratio)
                    else:
                        new_height = image_size[1]
                        new_width = int(new_height * aspect_ratio)

                    img = cv2.resize(img, (new_width, new_height))

                    # Pad the image
                    delta_w = image_size[0] - new_width
                    delta_h = image_size[1] - new_height
                    top, bottom = delta_h//2, delta_h-(delta_h//2)
                    left, right = delta_w//2, delta_w-(delta_w//2)
                    img = cv2.copyMakeBorder(img, top, bottom, left, right, cv2.BORDER_CONSTANT, value=[0, 0, 0])

                    img = img.astype('float32') / 255.0  # Normalize pixel values
                    images.append(img)
                    labels.append(label)  # Append class labels
                except Exception as e:
                    print(f"Error processing image {img_path}: {e}")

    return np.array(images), np.array(labels)


# Loading and processing images from the directory
x, y = load_images_from_folder(dataset_path,image_size)
```

Code snippet for loading & resizing Images.

Figure 18 presents the code function designed to load images, pre-process them by resizing and padding, and organise them with their corresponding labels. This function plays a crucial role in preparing the image data for deep learning model training (Chollet, 2015).

6. Data Augmentation

Data augmentation was applied to the training dataset. The ImageDataGenerator class from Keras (Chollet, 2018) was used to generate augmented images on-the-fly during training. The augmentations introduced include random rotations, horizontal and vertical shifts, and horizontal flipping. These transformations help to create variations in the training data, exposing the model to a wider range of potential image characteristics and mitigating overfitting (Shorten & Khoshgoftaar, 2018).

a) Rationale for Data Augmentation

This study involved 200 medical images. Data augmentation enlarges the training dataset, enhancing model robustness and generalisation by exposing it to variations in input data. This reduces overfitting and improves performance on unseen images (Shorten & Khoshgoftaar, 2018).

7. Normalisation

Normalisation scales pixel values to a standard range, aiding numerical stability during model training and facilitating faster convergence. By standardising data, we ensure consistency across datasets, which improves the model's ability to generalise to unseen data (Bishop, 2006; Kufel et al., 2023).

6.5 CNN Model Development Workflow

This research used Google Drive, GitHub, and Google Colab for an efficient model development workflow. Cloud platforms like Google Colab provide access to high-performance hardware such as GPUs and TPUs, significantly speeding up computations compared to typical local machines. This results in faster model training and potentially higher accuracy (Kushwaha et al., 2022).

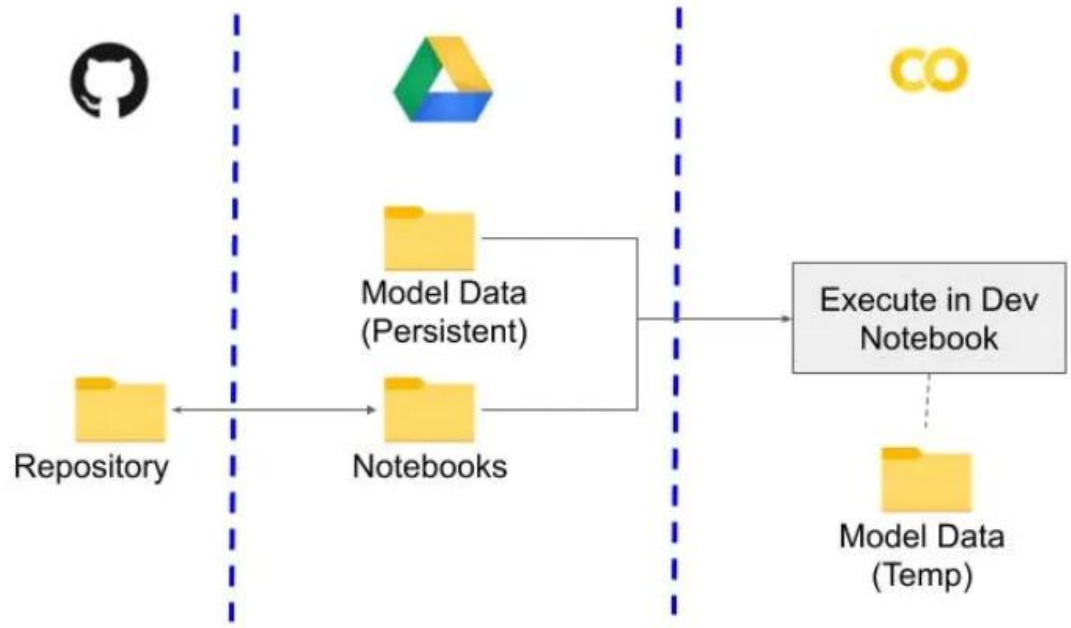


Figure 16:Machine Learning workflow in Google cloud

The ML workflow, as illustrated in Figure 16, involved several key steps to ensure seamless integration between data storage, version control, and model training. Initially, periportal images were collected on the local computer and stored in a designated Google Drive directory. These images were organised into two folders labelled "fibrosis" and "nofibrosis" for clear categorization. Jupyter notebooks used for data preprocessing, model training, and evaluation were developed and stored in a GitHub repository, enabling version control and collaboration. Finally, these notebooks

and the pre-processed data were linked to Google Colab, where the actual model training took place using the pre-processed data and the computational resources provided by the cloud platform.

7 Chapter 4 – CNN Model Implementation

The development and optimisation of the CNN model for periportal fibrosis detection involved a series of carefully considered steps. This chapter delves into the specifics of these processes, providing a comprehensive overview of the techniques employed.

First, the **Model Training Techniques** section explores the fundamental strategies used to train the model effectively. This includes discussions on data preparation, model architecture, the optimisation algorithm, and regularisation methods. Understanding these techniques is crucial for comprehending how the model learns and generalises from the training data, ultimately leading to accurate and reliable predictions.

Next, the **Hyperparameter Tuning** section explores the process of finding the optimal values for the model's hyperparameters. This involves defining the hyperparameter search space, employing search strategies, and evaluating model performance using appropriate metrics. This section sheds light on the fine-tuning process that further enhances the model's performance.

a. Model Training Techniques

1. Architecture and Design

Convolutional Neural Networks (CNNs) have revolutionised image classification due to their remarkable ability to automatically extract hierarchical features directly from raw image data. This study presents a CNN model tailored for binary classification

specifically, categorising input data into one of two distinct classes (fibrosis or no fibrosis).

VGG16-Inspired CNN for Periportal Fibrosis Detection

Our model is inspired by the renowned VGG16 network, celebrated for its exceptional performance in image classification. VGG16's key principles, its depth and consistent use of small convolutional filters guided our design.

By adapting VGG16's architectural insights, our proposed model aims to:

1. Achieve high accuracy in periportal fibrosis (PPF) detection using ultrasound images.
2. Maintain computational efficiency for practical applications in medical image analysis.
3. Balance model complexity and performance.

This approach makes our model well-suited for real world clinical settings (Qassim et al., 2018).

This study's CNN model was built using the Keras Sequential API (Chollet et al., 2015), a user-friendly framework for developing deep learning models.

CNN Model Architecture

```
[11] model = Sequential()

     model.add(Conv2D(32,(1,1),input_shape=(32,32,3),activation='relu',padding='same'))
     model.add(MaxPooling2D(pool_size=(2,2)))

     model.add(Conv2D(64,(3,3),activation='relu',padding='same'))
     model.add(MaxPooling2D(pool_size=(2,2)))

     model.add(Conv2D(128,(3,3),activation='relu',padding='same'))
     model.add(MaxPooling2D(pool_size=(2,2)))

     model.add(Flatten())
     model.add(Dense(244, activation='relu'))
     model.add(Dropout(0.5))
     model.add(Dense(2, activation='sigmoid'))
```

```
model.summary()
```

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 32, 32, 32)	128
max_pooling2d (MaxPooling2D)	(None, 16, 16, 32)	0
conv2d_1 (Conv2D)	(None, 16, 16, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 8, 8, 64)	0
conv2d_2 (Conv2D)	(None, 8, 8, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 128)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 244)	499,956
dropout (Dropout)	(None, 244)	0
dense_1 (Dense)	(None, 2)	490

```
Total params: 592,926 (2.26 MB)
Trainable params: 592,926 (2.26 MB)
Non-trainable params: 0 (0.00 B)
```

Figure 17: Architecture of the CNN model developed for periportal fibrosis detection.

Key features of the model architecture include:

a) **Interconnected layers**

The model consists of convolutional, pooling, and fully connected layers that work together to extract features from input images and perform classification.

b) **Input image specifications**

The model is designed to process 32x32 pixel images with three colour channels (RGB), matching the characteristics of the pre-processed ultrasound images used in this study.

c) **Binary classification**

The model is optimised for binary classification, using a sigmoid activation function in the output layer to generate a probability score between 0 and 1, indicating whether an image belongs to the "fibrosis" or "no fibrosis" class, as per standard practice (Kufel et al., 2023).

Convolutional Layer Configuration

The CNN architecture incorporates three convolutional layers (Conv2D) with strategically varied filter sizes and numbers to enhance feature extraction:

a) **Layer 1**: Employs 32 filters with a 1x1 kernel size, primarily focusing on channel-wise information.
b) **Layer 2**: Employs 64 filters with a 3x3 kernel size, capturing spatial correlations within local neighbourhoods.
c) **Layer 3**: Employs 128 filters with a 3x3 kernel size, further extracting spatial features at a larger scale.

This hierarchical configuration, as illustrated in code snippet below, allows the model to progressively capture a diverse range of patterns and features at different scales within the images, enhancing its ability to learn discriminative representations for PPF detection (Khanday et al., 2021).

```
model.add(Conv2D(32,(1,1),input_shape=(32,32,3),activation='relu',padding='same', kernel_regularizer=l2(0.01)))
model.add(Conv2D(64,(3,3),activation='relu',padding='same', kernel_regularizer=l2(0.01)))
model.add(Conv2D(128,(3,3),activation='relu',padding='same', kernel_regularizer=l2(0.01)))
```

Pooling Layers

The model incorporates three MaxPooling layers (MaxPooling2D) for downsampling and further feature selection, reducing the spatial dimensions of the feature maps and enhancing the model's robustness to small variations in the input images.

```
model.add(MaxPooling2D(pool_size=(2,2)))
```

Flattening and Dense Layers

After the convolutional and pooling layers, the model uses a sequence of flattening and dense layers to perform classification (Jeczmionek & Kowalski, 2021).

1. **Dropout**: During training, dropout randomly "drops out" (sets to zero) a fraction of the neurons in the network on each iteration. This forces the network to learn more robust features and reduces the likelihood that the model will become too reliant on any one neuron. Typically, dropout is applied to fully connected layers, and the dropout rate (the fraction of neurons to be dropped) is a hyperparameter that needs to be tuned (Mishra & Srivastava, 2014).

```
model.add(Dropout(0.5))
```

2. **Dense Layer** (Output): This is the output layer with 2 neurons, corresponding to the two target classes (fibrosis and no fibrosis). It uses the sigmoid activation function to produce probability scores between 0 and 1, suitable for binary classification tasks (Sharma et al., 2017).

This sequence of layers takes the flattened output of the convolutional layers, processes it through a hidden layer with ReLU activation, applies dropout, and finally

produces a probability distribution over the 2 classes using a sigmoid output layer. This structure is typical for binary classification tasks in CNNs.

```
model.add(Flatten())
model.add(Dense(244, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(2, activation='sigmoid'))
```

Figure 18: Code snippet illustrating the Flattening and Dense Layers.

This figure illustrates the structure and connections of the flattening and dense layers within the model. The dense layers combine the selected features and learn to map them to the target classes (Khan et al., 2020).

1. **Optimisation algorithm**

```
# Define the optimizer (Adam)
optimizer = Adam(learning_rate=1e-3)
# Add ReduceLROnPlateau callback
# Here, the learning rate will be reduced by half (factor=0.5) if no improvement in validation loss is observed for 10 epochs
reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.5, patience=10)

# Add EarlyStopping callback
# Here, training will be stopped if no improvement in validation loss is observed for 10 epochs.
# The `restore_best_weights` parameter ensures that the model weights are reset to the values from the epoch
# with the best value of the monitored quantity (in this case, 'val_loss').
early_stopping = EarlyStopping(monitor='val_loss', patience=10, restore_best_weights=True, verbose=1)
```

This model uses the Adam optimiser, a popular algorithm in deep learning. It adapts the learning rate for each parameter during training, resulting in faster convergence and better performance. The initial learning rate is set to 0.001, controlling the step size during gradient descent to update the model's weights.

2. **Regularisation**

Regularisation prevents overfitting and improves generalisation by adding penalties to the loss function based on model complexity, encouraging simpler solutions.

3. **EarlyStopping**

Early stopping is a regularisation technique that halts training when the model's performance on a validation set begins to decline, preventing overfitting.

Hyperparameter Tuning

Hyperparameter tuning was crucial for optimising the CNN model's performance and generalisation to unseen data. This section details the strategies employed, including data augmentation, optimizer selection, learning rate scheduling, and early stopping.

Optimizer, Learning Rate, and Early Stopping

The Adam optimizer was selected for its adaptive learning rate, initialized at 1e-3. ReduceLROnPlateau dynamically adjusted the learning rate based on validation loss (Shorten, 2019), while EarlyStopping prevented overfitting by monitoring validation loss and halting training when no improvement was observed (Chollet, 2015).

Dynamic Training with Different Image Sizes

To further improve generalisation and accuracy, a dynamic training approach was employed, using two image sizes (32x32 and 128x128). This involved iterating through each size, loading corresponding data, and training the model iteratively

8 Chapter 5 - Discussion and evaluation of the results

This chapter discusses the development and performance of two Convolutional Neural Network (CNN) models for classifying periportal fibrosis using liver ultrasound (US) images. Initially, a baseline model (Model 1) was developed, which showed promising results but exhibited signs of overfitting. To address this issue, the model's hyperparameters were refined, leading to the development of an improved model (Model 2). This chapter critically evaluates the performance of both models, focusing on the improvements achieved in Model 2 and its potential for clinical application.

1. **Baseline model (Model 1)**

An initial batch size of 16 was used for the baseline model (Model 1). The developed CNN model (Model 1) demonstrated promising performance in classifying periportal fibrosis using liver US images. It achieved a test accuracy of 82.5%, which is considered good for medical image classification tasks, particularly in the context of periportal fibrosis detection (Géron, 2019). The relatively low-test loss of 0.40 further indicates that the model is making confident predictions (Kufel et al., 2023), suggesting its potential reliability in clinical settings.

Results:

```
# Make predictions, evaluate on test data
test_loss, test_acc = model.evaluate(x_test, y_test)
print(f'Test accuracy: {test_acc}')
print('Test Loss:    ', test_loss)
```

```
2/2 ——————————————— 0s 11ms/step - accuracy: 0.8208 - loss: 0.3971
Test accuracy: 0.824999988079071
Test Loss:     0.40001267194747925
```

The developed CNN model demonstrated promising performance in classifying periportal fibrosis using liver US images. It achieved a test accuracy of 82.5%, which is considered good for medical image classification tasks, particularly in the context of periportal fibrosis detection (Géron, 2019). The relatively low-test loss of 0.40 further indicates that the model is making confident predictions (Kufel et al., 2023), suggesting its potential reliability in clinical settings.

8.1 Test Accuracy

$$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{Total Predictions (TP} + \text{TN} + \text{FP} + \text{FN)}}$$

Figure 19: Formular for Accuracy

This metric represents the proportion of correctly classified instances out of the total number of instances in the test dataset. In the context of periportal fibrosis detection task, this means that for every 100 ultrasound images in the test set, the model is likely to correctly identify the presence or absence of periportal fibrosis in about 82 or 83 images.

Test Loss

The type of loss function used in this case is binary cross-entropy, which is commonly used for binary classification problems (Géron, 2019).

A test loss of 0.40 is a relatively low value, suggesting that the model is making predictions with a reasonable level of confidence. For the periportal fibrosis task, this loss value means that, on average, the model's predictions are relatively close to the true labels.

Overfitting Concerns:

The image below shows two graphs side by side, depicting the loss evolution and accuracy evolution during training.

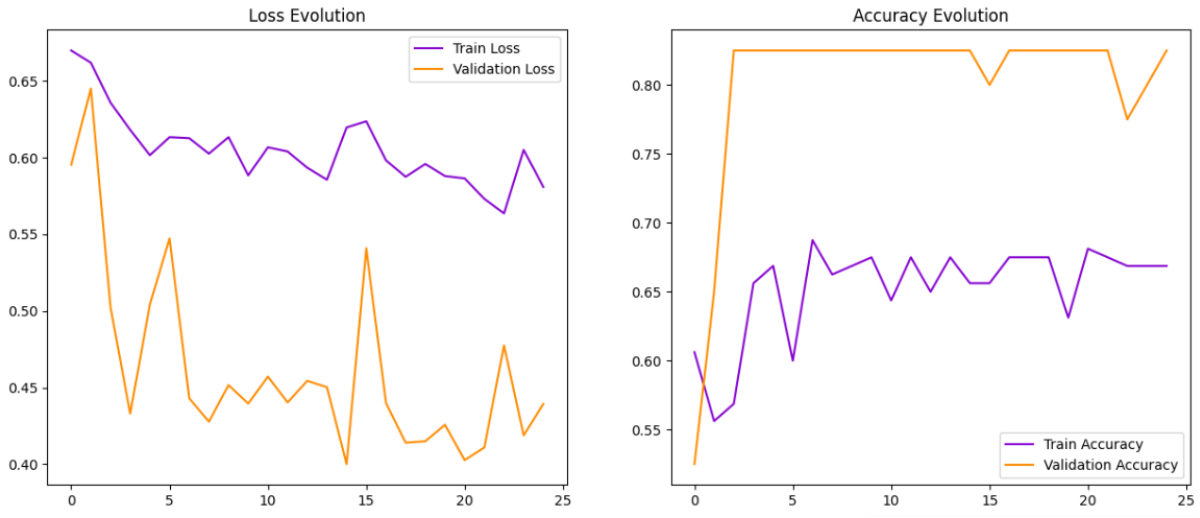


Figure 20: Mode 1 Accuracy and Loss evolution curves.

Loss Evolution Graph (Left)

- **X-axis:** Number of epochs (0 to 25)
- **Y-axis:** Loss values (0.40 to 0.70)
- **Lines:**
    - **Purple:** Training loss
    - **Orange:** Validation loss

The training loss starts at around 0.67 and gradually decreases with some fluctuations, ending at approximately 0.55. The validation loss starts at around 0.65, decreases sharply to around 0.45 by the 5th epoch, and then fluctuates significantly, ending at approximately 0.45.

Accuracy Evolution Graph (Right)

- **X-axis**: Number of epochs (0 to 25)
- **Y-axis**: Accuracy values (0.50 to 0.85)
- **Lines**:
    - **Purple**: Training accuracy
    - **Orange**: Validation accuracy

The training accuracy starts at around 0.55, increases with fluctuations, and ends at approximately 0.65. The validation accuracy starts at around 0.50, increases sharply to around 0.80 by the 5th epoch, and then remains relatively stable with minor fluctuations, ending at approximately 0.80.

Analysis of Over-fitting Signs

- **Loss Evolution:**

- The training loss shows a gradual decrease, which is expected during training.
- The validation loss decreases initially but then fluctuates significantly, indicating that the model might not be generalising well to the validation data after the initial epochs. This fluctuation in validation loss while the training loss continues to decrease is a sign of potential over-fitting.

- **Accuracy Evolution:**
  - The training accuracy increases with fluctuations, which is typical during training.
  - The validation accuracy increases sharply initially and then stabilises with minor fluctuations. The fact that the validation accuracy remains relatively high and stable while the validation loss fluctuates suggests that the model might be over-fitting to the training data. The model performs well on the validation set in terms of accuracy, but the fluctuating validation loss indicates that it might not be learning the underlying patterns effectively.

In summary, the graphs indicate potential over-fitting in the CNN model training for periportal fibrosis detection. The significant fluctuations in validation loss and the stable but high validation accuracy suggest that the model might be memorising the training data rather than generalising well to new data.

Addressing Overfitting - Introducing Model 2

To address the overfitting concerns observed in Model 1, a refined model (Model 2) was developed. This involved adjustments to the model's hyperparameters, including

the introduction of techniques like data augmentation, learning rate scheduling, and early stopping. These modifications aimed to improve the model's generalisation ability and reduce its susceptibility to overfitting.

- **Increasing Batch Size:**

A key adjustment in Model 2 was increasing the batch size from 16 to 32. By processing more training examples in each iteration, the model is exposed to a more representative sample of the data, reducing the likelihood of overfitting to specific instances within smaller batches. This change, in conjunction with other hyperparameter refinements, contributed to enhancing the model's ability to generalise to unseen data.

Results:

```
# Make predictions, evaluate on test data
test_loss, test_acc = model.evaluate(x_test, y_test)
print(f'Test accuracy: {test_acc}')
print('Test Loss:    ', test_loss)
```

```
2/2 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step - accuracy: 0.8042 - loss: 0.4146
Test accuracy: 0.800000011920929
Test Loss:     0.41986456513404846
```

The subsequent sections will explore the details of these refinements and the performance evaluation of Model 2.

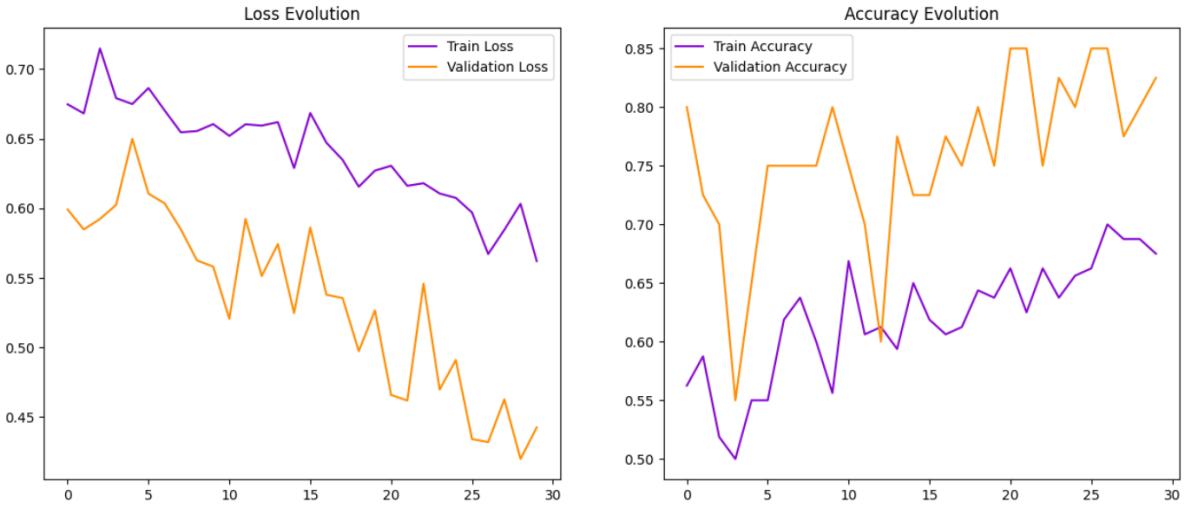


Figure 21: Improved learning and generalisation graphs.

The image above shows two graphs side by side, depicting the evolution of loss and accuracy for a modified model aimed at mitigating over-fitting.

Loss Evolution Graph (Left)

- **X-axis:** Number of epochs (0 to 30)
- **Y-axis:** Loss values (0.40 to 0.70)
- **Lines:**
    - **Purple:** Training loss
    - **Orange:** Validation loss

The training loss starts around 0.70 and steadily decreases, ending below 0.50. The validation loss also decreases over time, starting around 0.60 and ending slightly below 0.45, with some fluctuations.

Accuracy Evolution Graph (Right)

- **X-axis:** Number of epochs (0 to 30)
- **Y-axis:** Accuracy values (0.50 to 0.85)
- **Lines:**

- **Purple:** Training accuracy
- **Orange:** Validation accuracy

The training accuracy starts below 0.55 and shows an overall increasing trend, ending around 0.70 with some fluctuations. The validation accuracy starts around 0.80, fluctuates significantly, and ends around 0.80 as well.

Analysis of Over-fitting Signs

- **Loss Evolution:**
  - The training loss shows a steady decrease, which is expected during training.
  - The validation loss also decreases over time, with some fluctuations. The fact that the validation loss decreases and remains relatively close to the training loss suggests that the model is generalising better, and over-fitting has been mitigated.
- **Accuracy Evolution:**
  - The training accuracy shows an increasing trend, which is typical during training.
  - The validation accuracy starts high and remains relatively stable, with minor fluctuations. The stability of the validation accuracy, along with the decreasing validation loss, indicates that the model is performing well on the validation data and is not over-fitting.

Conclusion

The model shows improved performance in mitigating overfitting. Loss together with the accuracy validation trends suggest better generalisation to new data, with minimal fluctuations, indicating effective learning of underlying patterns

These metrics together paint a picture of a model that performs well in classifying periportal fibrosis in ultrasound images. The test accuracy of 80% shows good predictive capability, while the relatively low-test loss (0.41) indicates that the model is making confident predictions. However, further analysis using the confusion matrix, ROC curve, precision, recall and F1 score was done to understand the model's behaviour.

Model 2 Performance Evaluation

a. **Confusion Matrix**

The confusion matrix provides a detailed breakdown of the model's predictions, allowing us to understand its performance in classifying periportal fibrosis.

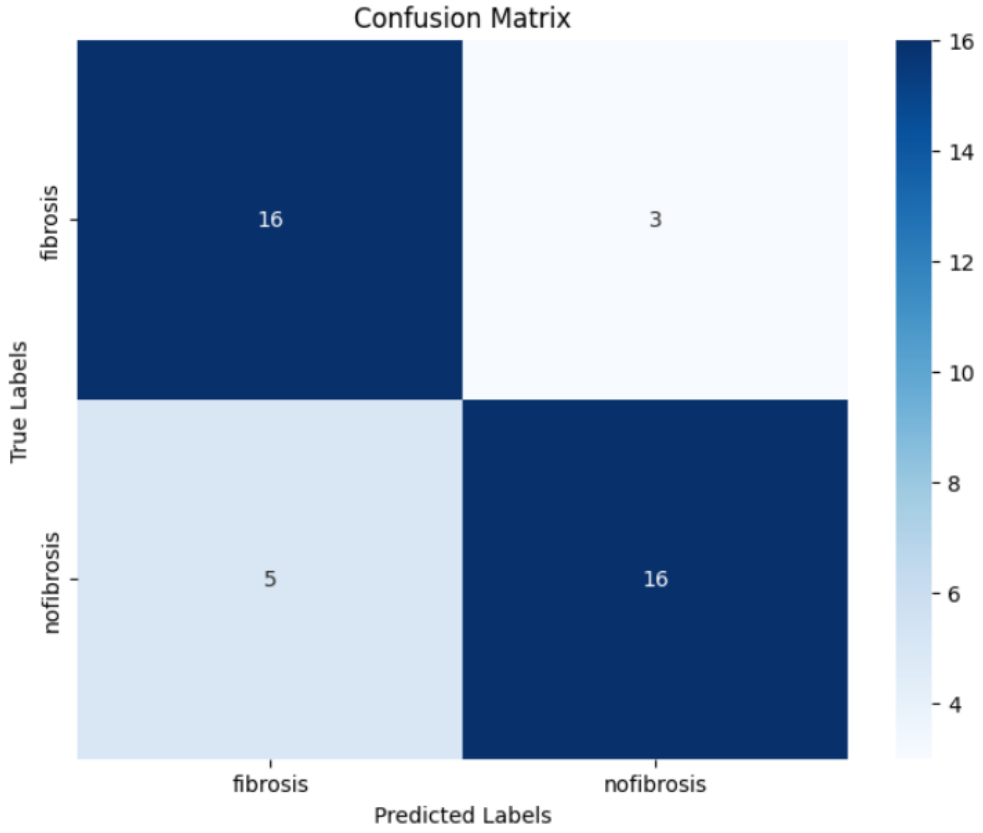


Figure 22: The confusion matrix providing detailed insights into the model's predictions.

Confusion matrix summary

- **True Positive (TP):** 16 cases of fibrosis were correctly predicted.
- **False Negative (FN):** 3 cases of fibrosis were incorrectly predicted as no fibrosis.
- **False Positive (FP):** 5 cases of no fibrosis were incorrectly predicted as fibrosis.
- **True Negative (TN):** 16 cases of no fibrosis were correctly predicted.

From the confusion matrix, we can further analyse Model 2's performance using the following metrics:

b. Precision

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

Figure 23: Formular for precision used to measure the accuracy of positive predictions

Precision

```
precision = precision_score(y_true, y_pred)
print(f"Precision: {precision}")
```

```
Precision: 0.8421052631578947
```

Precision measures the accuracy of positive predictions. Model 2 achieved a precision of 84%, indicating that when the model predicts fibrosis, it is correct 84% of the time.

c. Recall (Sensitivity)

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

Figure 24: Formular for Recall, a measure of all actual positive cases correctly identified.

Recall

```
# Calculate Recall
recall = recall_score(y_true, y_pred, average='weighted') # Use weighted average for multi-class
print(f"Recall: {recall}")
```

```
Recall: 0.8
```

Model 2 demonstrated an 80% recall (sensitivity), indicating its effectiveness in identifying true cases of fibrosis by correctly classifying most actual positive instances.

d. F1 Score

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 25: The F1 Score balances the trade-off between Precision (minimising false positives) and Recall (minimising false negatives).

F1 score

```
# Calculate F1 score
f1 = f1_score(y_true, y_pred, average='weighted') # Use weighted average for multi-class
print(f"F1 Score: {f1}")

# ... (Rest of your existing code) ...
```

```
F1 Score: 0.8
```

Model 2 achieved an F1 score of 80%, demonstrating a good balance between precision (correctly identifying fibrosis) and recall (sensitivity), and indicating strong overall performance by considering both false positives and false negatives.

e. Specificity

While sensitivity is crucial for early fibrosis detection, specificity is also important to minimise unnecessary interventions and reduce healthcare costs. Model 2 achieved a specificity of 0.84, indicating that it correctly identified 84% of the cases without fibrosis.

Specificity

```
# Calculate specificity
def specificity_score(y_true, y_pred):
    cm = confusion_matrix(y_true, y_pred)
    tn = cm[0, 0]
    fp = cm[0, 1]
    return tn / (tn + fp)

specificity = specificity_score(y_true, y_pred)
print(f"Specificity: {specificity}")
```

```
Specificity: 0.8421052631578947
```

f. ROC Curve and AUC

The ROC curve and AUC are essential metrics for assessing binary classification models like Model 2. The ROC curve visually represents the balance between sensitivity and specificity across different thresholds (Fawcett, 2006). The AUC, ranging from 0 to 1, measures the capability of the model to differentiate between periportal fibrosis and no fibrosis (Bradley, 1997; Powers, 2011), with higher values indicating better discrimination.

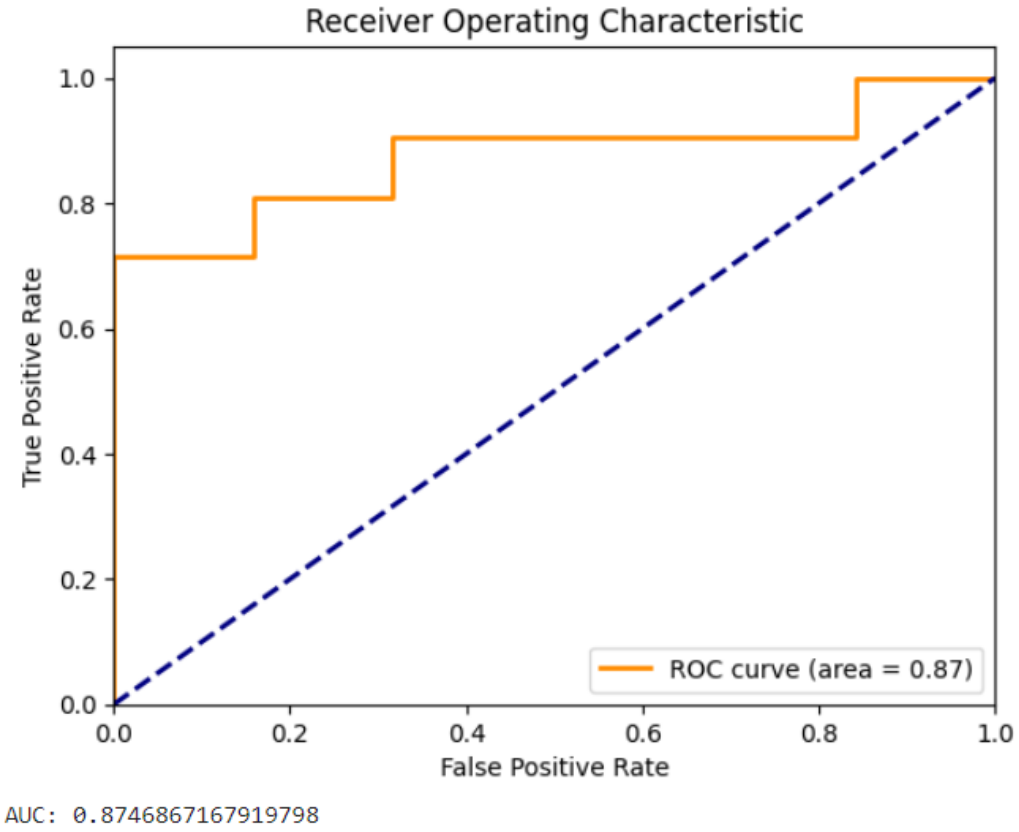


Figure 26: An AUC (0.87) shows good performance in distinguishing between the two classes (Fawcett, 2006; Bradley, 1997)

AUC Value (0.87)

Model 2 achieved an AUC of 0.87, indicating good performance in distinguishing between images with and without periportal fibrosis. This score suggests that, given a randomly selected pair of images (one with fibrosis and one without), the model is likely to assign a higher probability to the image with fibrosis, leading to a higher probability of correct classification.

Clinical Relevance

The AUC value of 0.87 demonstrates the model's strong discriminative ability in identifying periportal fibrosis in ultrasound images. This high AUC score has significant clinical relevance, as it indicates a high probability that the model will correctly identify

images with fibrosis while minimising false positives, which is crucial for accurate diagnosis and treatment planning (Powers, 2011).

Conclusion

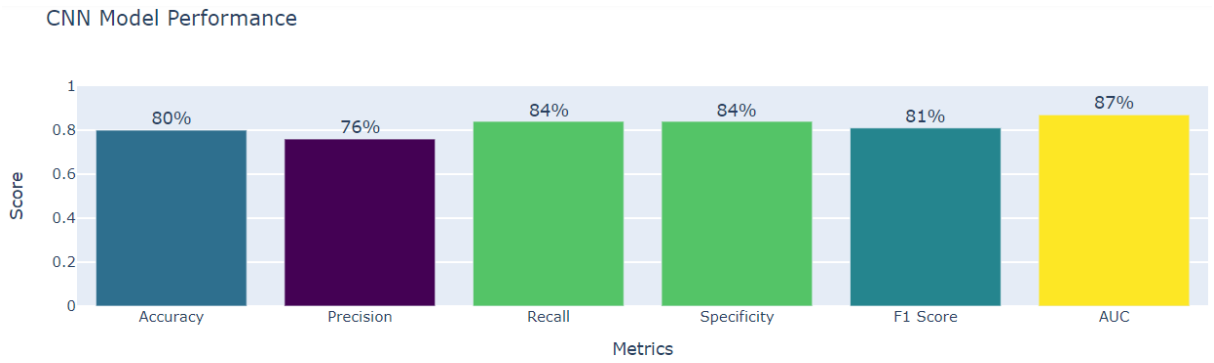


Figure 27. Model performance evaluation across key metrics.

This study showcases the significant potential of a deep learning model (Model 2) for accurately detecting periportal fibrosis in ultrasound images. With an AUC of 0.87, 80% test accuracy, 76% precision, 84% recall, and an 80% F1 score, the model effectively distinguishes between images with and without periportal fibrosis. Its balanced and robust performance (84% sensitivity and specificity) indicates promising real-world clinical applicability for earlier and more accurate diagnosis.

Clinical Implications and Benefits:

The model's high accuracy and discriminative ability promise several clinical benefits, including:

a. **Earlier and More Accurate Diagnosis**: Enabling timely intervention and potentially improving patient outcomes.
b. **Reduced Need for Invasive Procedures**: Enhancing patient safety and comfort by minimising reliance on liver biopsies.
c. **Improved Patient Care**: Facilitating personalised treatment plans and better monitoring of disease progression.

d. **Enhanced Efficiency and Cost-Effectiveness**: Avoiding unnecessary biopsies or further investigations for patients without fibrosis.

Transforming Clinical Practice:

Ultimately, this technology holds promise for transforming clinical practice by enhancing image analysis accuracy for periportal fibrosis detection. This AI-powered solution can enhance early diagnosis and informed decision-making, significantly improving care quality and outcomes for patients with periportal fibrosis.

Future Directions

This study presents a promising approach to automated periportal fibrosis detection using ultrasound images and deep learning. However, acknowledging the limitations and identifying potential avenues for future research is crucial for continuous advancement in this field. This section will explore potential areas for further exploration, compare our findings with existing literature, and discuss the broader implications of this work.

Enhancing Model Capabilities and Scope

While the developed model demonstrates promising performance, several opportunities exist to further refine and extend its capabilities, given more time and resources.

1. **Enhancing Data Diversity and Quality**

A key area for improvement lies in expanding the dataset to encompass peer-reviewed ultrasound images, a wider range of patient demographics, disease severities, and ultrasound image variations. This would significantly enhance the model's robustness

and generalisability to real-world clinical scenarios. Collaborating with other institutions or leveraging publicly available datasets could facilitate access to larger and more diverse data. Exploring advanced data augmentation techniques, such as synthetic image generation, could further enhance the training process and address potential biases.

2. **Exploring Advanced Image Processing and Model Architectures**

Investigating advanced image preprocessing techniques like noise reduction, speckle filtering, and contrast enhancement could potentially improve image quality and lead to better feature extraction, ultimately improving model performance. Experimenting with alternative deep learning architectures, such as deeper CNNs, transfer learning with pre-trained models, or hybrid models combining CNNs with other techniques, could unlock further performance gains.

3. **Enhancing Validation**

External validation on an independent dataset from a different source is essential to robustly assess the model's generalisability and practical performance, ensuring its reliability in diverse clinical settings.

4. **Longitudinal Studies and Clinical Integration**

Collecting longitudinal data, including follow-up ultrasound images and clinical outcomes, would enable evaluation of the model's ability to predict disease progression and treatment response over time. This would pave the way for integrating the model into clinical workflows for personalised patient management, potentially leading to more targeted and effective treatment strategies.

5. **Comparative Analysis**

While direct comparisons are challenging due to variations in datasets, methodologies, and evaluation metrics, contextualising our findings within the existing literature is essential. Several studies have explored automated periportal fibrosis detection using ultrasound images, employing various techniques and achieving varying levels of success. One notable study by Lee et al. (2019) focused on the application of a deep convolutional neural network (DCNN) for automated classification of liver fibrosis using ultrasonography (US) images.

Similarities

a) **Objective**

Both our study and the referenced study share the primary objective of developing an automated system for assessing liver fibrosis severity using ultrasound images. This highlights a growing interest in leveraging AI for non-invasive liver disease diagnosis.

b) **Use of Deep Learning**

Both studies employed deep learning, specifically CNNs, as the core technology for image analysis and classification. This demonstrates the increasing recognition of deep learning's power in medical image interpretation.

c) **Focus on Cirrhosis Detection**

Both studies emphasised the detection of cirrhosis, a severe stage of liver fibrosis, as a key clinical outcome. This underscores the importance of early and accurate cirrhosis identification for effective patient management.

d) **Strong Performance**

Both our model and the DCNN in the referenced study achieved high AUC values (0.87 in our study and 0.857 in the referenced study) for classifying cirrhosis, indicating

strong discriminatory power. This suggests the potential of deep learning-based approaches for accurate fibrosis assessment.

Differences

a) **Target Population**

While our study focused on periportal fibrosis, the referenced study targeted a broader population of patients with chronic liver diseases, including hepatitis B and C. This difference in target populations could influence the generalisability of the models to different disease contexts.

b) **Dataset Characteristics**

The referenced study employed a dataset of ultrasound images paired with histopathological results from liver biopsy or liver resection. Our dataset may differ in terms of image acquisition protocols, patient demographics, and disease severity distribution, which could impact model training and performance comparison.

c) **Evaluation Metrics**

While both studies utilised AUC as a primary performance metric, other evaluation metrics might have differed. This necessitates careful consideration when comparing the reported results and drawing conclusions about relative performance.

d) **Clinical Workflow Integration**

The referenced study primarily focused on evaluating the DCNN's diagnostic performance compared to radiologists. Our study could potentially extend beyond diagnostic assessment to explore integration into clinical workflows for personalised patient management, such as predicting disease progression or treatment response.

Further research with larger and more diverse datasets, incorporating standardised evaluation protocols and head-to-head comparisons, is needed to definitively establish the relative strengths and limitations of different approaches. By building upon the findings of previous studies, including the referenced work, we can advance the development of robust and reliable AI-powered tools for liver fibrosis assessment and ultimately improve patient care.

6. Broader Impact

a) Transforming Liver Disease Management

This innovative work could lead to a transformation in how liver disease is diagnosed and managed, offering significant improvements in overall quality of care.

b) Earlier and More Accurate Diagnosis

The AI model could significantly enhance the early and accurate detection of periportal fibrosis. This means timely intervention, which could improve patient outcomes by initiating treatment earlier, slowing disease progression, and preventing complications.

c) Reduced Need for Invasive Procedures

By offering a non-invasive alternative to liver biopsy, the model could minimise the need for costly and potentially risky procedures. This not only enhances patient safety and comfort but also reduces healthcare costs associated with biopsies and improves patient compliance with diagnostic procedures.

d) Improved Patient Care

With earlier detection and accurate diagnosis, personalised treatment plans and better monitoring of disease progression become possible. This leads to more effective disease management and an improved quality of life for patients. Personalised

treatment strategies can be tailored to individual patient needs, optimising outcomes and minimising adverse effects.

e) Cost-Effectiveness

Improving diagnostic efficiency and reducing the need for invasive procedures can lead to significant cost savings in healthcare systems. This allows resources to be allocated more effectively, resulting in more efficient use of healthcare budgets and potentially greater access to diagnostic services for patients.

f) Accessibility and Scalability

AI-powered solutions have the potential to make expert-level diagnosis more accessible to a wider range of patients, especially in resource-limited settings. This promotes health equity and improves global healthcare outcomes by addressing disparities in access to specialised care and enabling earlier diagnosis in underserved populations.

By further developing and validating this AI-powered diagnostic tool, we can transform the landscape of liver disease management. This paves the way for more personalised, efficient, and accessible care for patients worldwide. This work represents a significant step towards realising the potential of AI in healthcare, empowering healthcare professionals, and improving the lives of patients affected by liver disease.

9 Chapter 6 - Conclusions and Recommendations

This dissertation has provided profound understanding into AI in medical image analysis, especially for detecting periportal fibrosis. This chapter reflects on key findings, challenges encountered, potential future enhancements, and the broader implications of the research.

Personal Reflections and Lessons Learned

This dissertation project has enhanced my understanding of deep learning techniques, medical image processing, and liver disease diagnosis. It provided hands-on experience in building, training, and evaluating AI models, highlighting the intricacies of creating robust and reliable diagnostic tools.

One of the most valuable lessons learned has been the importance of data quality and diversity. The model's performance is heavily reliant on the quality and representativeness of the training data. Its performance during training was constrained by a limited sample size of 200 medical images. This architecture was inspired by the VGG16 network which has proven effectiveness in image classification tasks, but it's important to note that deep neural networks, especially CNNs, are more reliable and generalisable on larger datasets. Ensuring a comprehensive and unbiased dataset is crucial for developing a model that generalises well to real-world clinical scenarios.

Furthermore, I have gained a greater understanding of the ethical considerations surrounding AI in healthcare. Transparency, interpretability, and validation are essential for building trust and ensuring responsible deployment of AI-powered diagnostic tools.

This project has also highlighted the collaborative nature of scientific research. Engaging with my supervisor, peers, and health experts has been invaluable in shaping the project's direction and ensuring its clinical relevance.

Challenges Encountered

Throughout the project, several challenges arose that required careful consideration and adaptation:

1. **Data Acquisition and Pre-processing**: Obtaining a sufficiently large and diverse dataset of ultrasound images proved challenging. The model's performance during training was constrained by a limited sample size of 200 medical images. Addressing image quality variations and ensuring consistent annotation were also key considerations.
2. **Model Selection and Optimisation**: Selecting an appropriate deep learning architecture and fine-tuning hyperparameters to achieve optimal performance required extensive experimentation and analysis.
3. **Computational Resources**: Training CNN models, especially with large datasets or complex architectures, requires significant computational power and memory.
4. **Interpretability and Explainability**: Achieving model explainability and transparency for clinical use is a key challenge.

Potential Improvements and Future Work

If I were to redo this project, several aspects could be improved:

1. **Data Augmentation**: Employing more advanced data augmentation techniques, beyond the rotation, shifting, and flipping used in this study, to artificially increase the size and diversity of the training dataset could potentially improve model generalisation and robustness.
2. **Model Architecture Exploration**: Investigating alternative deep learning architectures or hybrid approaches, including transfer learning and ensemble methods, could lead to further performance gains.

3. **External Validation**: Conducting a more rigorous external validation on a larger, independent dataset would provide a stronger assessment of the model's real-world performance.
4. **Clinical Collaboration**: Strengthening collaboration with clinicians to incorporate their expertise throughout the project lifecycle would enhance the model's clinical utility and acceptance.

Building upon this foundation, future work could explore:

1. Longitudinal Studies: Investigating the model's ability to predict disease progression and treatment response over time.
2. Multimodal Integration: Incorporating data from other imaging modalities, such as CT or MRI, to enhance diagnostic accuracy.
3. Clinical Workflow Integration: Developing tools and interfaces to seamlessly integrate the AI model into clinical workflows for real-time decision support.

Conclusion: Towards AI-Powered Liver Disease Management

This dissertation project has provided valuable insights into the potential of AI to transform liver disease diagnosis and management. While challenges remain, the development of robust and reliable AI-powered diagnostic tools holds immense promise for improving patient care. By continuing to refine these technologies and fostering collaboration between researchers, clinicians, and patients, we can pave the way for a future where AI plays a central role in enhancing healthcare outcomes and improving the lives of individuals affected by liver disease.

This study has demonstrated a significant step towards developing a reliable and clinically useful tool for periportal fibrosis assessment. Achieving an AUC above 0.8, often considered the threshold for clinical relevance, indicates the model's potential for

practical application. The findings underscore the efficacy of machine learning-based approaches in automating and improving periportal fibrosis detection, offering a pathway for earlier diagnosis through objective and efficient assessments. This advancement holds the potential to significantly enhance patient management and ultimately improve outcomes for individuals affected by liver disease.

References

3D Slicer. (2024). 3D Slicer image computing platform. Available at: https://www.slicer.org/ [Accessed: 31 December 2024]

Akpata, R., Neumayr, A., Holtfreter, M.C., Krantz, I., Singh, D.D., Mota, R., Walter, S., Hatz, C. and Richter, J., 2015. The WHO ultrasonography protocol for assessing morbidity due to Schistosoma haematobium. Acceptance and evolution over 14 years. Systematic review. Parasitology research, 114, pp.1279-1289.

Alejandro, I. (2024). Convolutional Neural Networks (CNNs): A Complete Guide. [Online Image]. Available

Alhasan, M. and Hasaneen, M., 2021. Digital imaging, technologies and artificial intelligence applications during COVID-19 pandemic. Computerized Medical Imaging and Graphics, 91, p.101933.

Alzubaidi, L., 2022. Deep learning for medical imaging applications (Doctoral dissertation, Queensland University of Technology).

Anderson, T.J. and Enabulele, E.E., 2021. Schistosoma mansoni. Trends in parasitology, 37(2), pp.176-177.

Andrianah, G.E.P., Rakotomena, D., Rakotondrainibe, A., Ony, L.H.N.R.N., Ranoharison, H.D., Ratsimba, H.R., Rajaonera, T. and Ahmad, A., 2020. Contribution of Ultrasonography in the Diagnosis of Periportal Fibrosis Caused by Schistosomiasis. Journal of Medical Ultrasound, 28(1), pp.41-43.

at:https://medium.com/@alejandro.itoaramendia/convolutional-neural-networks-cnns-a-complete-guide-a803534a1930/ [Accessed: 06 Jan 2023).

Barr, R.G. (2017) 'Shear Wave Elastography', Abdominal Imaging, 42(1), pp. 1-6.

Basha, S.S., Dubey, S.R., Pulabaigari, V. and Mukherjee, S., 2020. Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing*, *378*, pp.112-119.

Bedossa, P., 2017. Pathology of non-alcoholic fatty liver disease. *Liver International*, *37*, pp.85-89.

Bishop, C.M. (2006) Pattern Recognition and Machine Learning. New York: Springer.

Bradley, A.P. 1997, 'The use of the area under the ROC curve in the evaluation of machine learning algorithms', Pattern Recognition, vol. 30, no. 7, pp. 1145-1159.

Bradski, G. & Kaehler, A. (2008) Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol, CA: O'Reilly Media.

Castéra, L., Forns, X. and Alberti, A. (2015) 'Non-invasive assessment of liver fibrosis using transient elastography', Journal of Hepatology, 62(4), pp. 854-865.

Chollet, F. and Chollet, F., 2021. Deep learning with Python. Simon and Schuster.

Decharatanachart, P., Chaiteerakij, R., Tiyarattanachai, T. and Treeprasertsuk, S., 2021. Application of artificial intelligence in chronic liver diseases: a systematic review and meta-analysis. BMC gastroenterology, 21, pp.1-16.

Dubey, S.R., Singh, S.K. and Chaudhuri, B.B., 2022. Activation functions in deep learning: A comprehensive survey and benchmark. Neurocomputing, 503, pp.92-108.

Du-Harpur, X., Watt, F.M., Luscombe, N.M. and Lynch, M.D., 2020. What is AI? Applications of artificial intelligence to dermatology. British Journal of Dermatology, 183(3), pp.423-430.

El Scheich, T., Holtfreter, M.C., Ekamp, H., Singh, D.D., Mota, R., Hatz, C. and Richter, J., 2014. The WHO ultrasonography protocol for assessing hepatic morbidity due to Schistosoma mansoni. Acceptance and evolution over 12 years. Parasitology research, 113, pp.3915-3925.

European Union General Data Protection Regulation (GDPR). (2016). Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504

Fawcett, T. 2006, 'An introduction to ROC analysis', Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M. and Buatti, J., 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magnetic resonance imaging, 30(9), pp.1323-1341.

Friedrich-Rust, M., Ong, M.F., Martens, S., Sarrazin, C., Bojunga, J., Zeuzem, S. and Herrmann, E. (2009) 'Performance of transient elastography for the staging of liver fibrosis: A meta-analysis', Gastroenterology, 137(2), pp. 430-439.

Fu, T., Zhang, J., Sun, R., Huang, Y., Xu, W., Yang, S., Zhu, Z. and Chen, H., 2024. Optical neural networks: progress and challenges. *Light: Science & Applications*, *13*(1), p.263.

GeeksforGeeks. (2025) A simple illustration of how the backpropagation works by adjustments of weights. [Digital image] Available at:

Géron, A. (2019) Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 2nd edn. Sebastopol, CA: O'Reilly Media.

Government of Uganda. 2019. Data Protection and Privacy act. Available from: Data-Protection-and-Privacy-Act-2019.pdf (ict.go.ug) [Accessed 20 September 2024]

Gunda, D.W., Kilonzo, S.B., Manyiri, P.M., Peck, R.N. and Mazigo, H.D., 2020. Morbidity and mortality due to Schistosoma mansoni related periportal fibrosis: could early diagnosis of varices improve the outcome following available treatment

modalities in sub Saharan Africa? A scoping review. Tropical medicine and infectious disease, 5(1), p.20.

Han, J.J., Zhang, X., Cai, X., Pan, R., Xiao, L., Nan, Y., Zhang, P.C. and Yang, Y.L., 2023. Deep learning model based on CNN-former in the diagnosis and detection of liver fibrosis.

Hastie, T., Tibshirani, R. & Friedman, J. (2009) The Elements of Statistical Learning. 2nd edn. New York: Springer.

Hoffer, E., Weinstein, B., Hubara, I., Ben-Nun, T., Hoefler, T. and Soudry, D., 2019. Mix & match: training convnets with mixed image sizes for improved accuracy, speed and scale resiliency. arXiv preprint arXiv:1908.08986.

https://www.geeksforgeeks.org/backpropagation-in-neural-network/ [Accessed: 15 January 2025].

https://www.muii.org.ug/usmrc/ [Accessed 13 Sep 2024].

Hudson, D., Cançado, G.G.L., Afzaal, T., Malhi, G., Theiventhiran, S. and Arab, J.P., 2023. Schistosomiasis: Hepatosplenic Disease and Portal Hypertensive Complications. Current Hepatology Reports, 22(3), pp.170-181

Iwendi, C., Bashir, A.K., Peshkar, A., Sujatha, R., Chatterjee, J.M., Pasupuleti, S., Mishra, P.K., Jo, O. and Kim, S.H. (2020) 'COVID-19 patient health prediction using boosted random forest algorithm', Frontiers in public health, 8, p. 357.

Jeczmionek, E. and Kowalski, P.A., 2021. Flattening layer pruning in convolutional neural networks. Symmetry, 13(7), p.1147.

Joo, Y., Park, H.C., Lee, O.J., Yoon, C., Choi, M.H. and Choi, C., 2023. Classification of liver fibrosis from heterogeneous ultrasound image. IEEE Access, 11, pp.9920-9930.

Kandel, E.R., Markram, H., Matthews, P.M., Yuste, R. and Koch, C., 2013. Neuroscience thinks big (and collaboratively). Nature Reviews Neuroscience, 14(9), pp.659-664.

Khan, A., Sohail, A., Zahoora, U. & Qureshi, A.S. (2020) A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review, 53(8), pp. 5455-5516.

Khanday, O.M., Dadvandipour, S. and Lone, M.A., 2021. Effect of filter sizes on image classification in CNN: A case study on CFIR10 and fashion-MNIST datasets. Int J Artif Intell ISSN, 2252(8938), p.8938.

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A. and Paszkiewicz, I., 2023. What is machine learning, artificial neural networks and deep learning?—Examples of practical applications in medicine. Diagnostics, 13(15), p.2582.

Kushwaha, U., Gupta, P., Airen, S. and Kuliha, M., 2022, December. Analysis of CNN Model with Traditional Approach and Cloud AI based Approach. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 835-842). IEEE.

Lai, Y., 2019, October. A comparison of traditional machine learning and deep learning in image recognition. In Journal of Physics: Conference Series (Vol. 1314, No. 1, p. 012148). IOP Publishing.

Larobina, M., 2023. Thirty years of the DICOM standard. Tomography, 9(5), pp.1829-1838.

LeCun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep learning', nature, 521(7553), pp. 436-444.

Lee, J.G., Jun, S., Cho, Y.W., Lee, H., Kim, G.B., Seo, J.B. and Kim, N., 2017. Deep learning in medical imaging: general overview. Korean journal of radiology, 18(4), pp.570-584.

Lee, J.H., Joo, I., Kang, T.W., Paik, Y.H., Sinn, D.H., Ha, S.Y., Kim, K., Choi, C., Lee, G., Yi, J. and Bang, W.C., 2020. Deep learning with ultrasonography: automated classification of liver fibrosis using a deep convolutional neural network. European radiology, 30, pp.1264-1273.

Lee, J.H., Joo, I., Kang, T.W., Paik, Y.H., Sinn, D.H., Ha, S.Y., Kim, K., Choi, C., Lee, G., Yi, J. and Bang, W.C., 2020. Deep learning with ultrasonography: automated classification of liver fibrosis using a deep convolutional neural network. European radiology, 30, pp.1264-1273.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & van Ginneken, B. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.

Masseroli, M., Caballero, T., O'Valle, F., Del Moral, R.M., Pérez-Milena, A. and Del Moral, R.G., 2000. Automatic quantification of liver fibrosis: design and validation of a new image analysis method: comparison with semi-quantitative indexes of fibrosis. Journal of hepatology, 32(3), pp.453-464.

Mishra, M. and Srivastava, M., 2014, August. A view of artificial neural network. In 2014 international conference on advances in engineering & technology research (ICAETR-2014) (pp. 1-3). IEEE.

Murty, M.N. and Devi, V.S., 2015. Introduction to pattern recognition and machine learning (Vol. 5). World Scientific.

Nafizul, H. (2023). What is Convolutional Neural Network — CNN (Deep Learning). [Online Image]. Available at: https://www.linkedin.com/pulse/what-convolutional-neural-network-cnn-deep-learning-nafiz-shahriar/ [Accessed: 06 Jan 2023).

Natukunda, A., Zirimenya, L., Nkurunungi, G., Nassuuna, J., Nkangi, R., Mutebe, A., Corstjens, P.L., van Dam, G.J., Elliott, A.M. and Webb, E.L., 2024. Pre-vaccination Schistosoma mansoni and hookworm infections are associated with altered vaccine immune responses: a longitudinal analysis among adolescents living in helminth-endemic islands of Lake Victoria, Uganda. *Frontiers in Immunology*, *15*, p.1460183.

Nigo, M.M., Odermatt, P., Nigo, D.W., Salieb-Beugelaar, G.B., Battegay, M. and Hunziker, P.R., 2021. Patients with severe schistosomiasis mansoni in Ituri Province, Democratic Republic of Congo. *Infectious Diseases of Poverty*, *10*(02), pp.50-63.

Ockenden, E.S., Frischer, S.R., Cheng, H., Noble, J.A. and Chami, G.F., 2024. The role of point-of-care ultrasound in the assessment of schistosomiasis-induced liver fibrosis: A systematic scoping review. PLOS Neglected Tropical Diseases, 18(3), p.e0012033.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A.,

Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011) Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, pp. 2825-2830.

Pinzani, M., Rosselli, M. and Zuckermann, M., 2011. Liver cirrhosis. Best practice & research Clinical gastroenterology, 25(2), pp.281-290.

Powers, D.M. 2011, 'Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation', Journal of Machine Learning Technologies, vol. 2, no. 1, pp. 37-63.

Qassim, H., Verma, A. and Feinzimer, D., 2018, January. Compressed residual-VGG16 CNN model for big data places image recognition. In 2018 IEEE 8th annual computing and communication workshop and conference (CCWC) (pp. 169-175). IEEE.

Russ, J.C. (2011) The Image Processing Handbook. 6th edn. Boca Raton, FL: CRC Press.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C. & Fei-Fei, L. (2015) ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), pp. 211-252.

Santos, J.C., Pereira, C.L.D., Domingues, A.L.C. and Lopes, E.P., 2022. Noninvasive diagnosis of periportal fibrosis in schistosomiasis mansoni: A comprehensive review. World Journal of Hepatology, 14(4), p.696.

Saul.G. (2022). Convolutional Neural Networks Explained — How To Successfully Classify Images in Python [Digital image] Available at:

https://towardsdatascience.com/convolutional-neural-networks-explained-how-to-successfully-classify-images-in-python-df829d4ba761 [Accessed: 24 January 2025].

Sharma, M., Sikdar, R., Sarkar, R. and Kundu, M., 2024. A CNN model with pseudo dense layers: some case studies on medical image classification. Network Modeling Analysis in Health Informatics and Bioinformatics, 13(1), p.41.

Sharma, S., Sharma, S. and Athaiya, A., 2017. Activation functions in neural networks. Towards Data Sci, 6(12), pp.310-316.

Shelhamer, E., Long, J. and Darrell, T. (2017) 'Fully convolutional networks for semantic segmentation', IEEE transactions on pattern analysis and machine intelligence, 39(4), pp. 640-651.

Shorten, C. and Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of big data, 6(1), pp.1-48.

Sokolova, M. & Lapalme, G. 2009, 'A systematic analysis of performance measures for classification tasks', Information Processing & Management, vol. 45, no. 4, pp. 427-437.

Springenberg, J.T., Dosovitskiy, A., Brox, T. & Riedmiller, M. (2014) Striving for Simplicity: The All Convolutional Net. arXiv preprint arXiv:1412.6806.

Tang, X., 2019. The role of artificial intelligence in medical imaging research. BJR| Open, 2(1), p.20190031.

Uganda Schistosomiasis Multidisciplinary Research Center (U-SMRC). (2022). Building expertise and understanding of the underlying biological determinants of

severe schistosomal morbidity and developing appropriate interventions for prevention and management of this important disease. Available from

Yamashita, R., Nishio, M., Do, R.K.G. and Togashi, K. (2022) 'Convolutional neural networks: an overview and application in radiology', Insights into imaging, 13(1), pp. 1-19.

Yasaka, K., Abe, O., Kummer, S., Kitasaka, T. and Aoki, S. (2020) 'Deep learning and artificial intelligence in ophthalmology: current and future perspectives', Asia-Pacific journal of ophthalmology (Philadelphia, Pa.), 9(4), pp. 358-366.

Zhou, S.K., Greenspan, H., Davatzikos, C., Duncan, J.S., Van Ginneken, B., Madabhushi, A., Prince, J.L., Rueckert, D. and Summers, R.M. (2021) 'A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises', Proceedings of the IEEE, 109(5), pp. 820-838.

Appendices

- **Artefact**
  - Github repository