DEEP LEARNING-BASED ANALYSIS OF ELECTROCHEMICAL, BIOMEDICAL, AND OPTICAL SIGNALS

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Biotechnology

by

Muhammet Çağrı YEKE

June 2024
İZMİR

We approve the thesis of **Muhammet Çağrı YEKE**

Examining Committee Members:

Asst. Prof. Dr. Abdurrahman GÜMÜŞ
Department of Electrical and Electronics Engineering, İzmir Institute of Technology

Prof. Dr. Dilek ODACI
Department of Biochemistry, Ege University

Assoc. Prof. Kıvılcım YÜKSEL ALDOĞAN
Department of Electrical and Electronics Engineering, İzmir Institute of Technology

Prof. Dr. Bilge KARAÇALI
Department of Electrical and Electronics Engineering, İzmir Institute of Technology

Asst. Prof. Dr. Başak Esin KÖKTÜRK GÜZEL
Department of Electrical and Electronics Engineering, İzmir Democracy University

26 June 2024

Asst. Prof. Dr. Abdurrahman GÜMÜŞ
Supervisor, Department of Electrical and Electronics Engineering
İzmir Institute of Technology

Prof. Dr. Dilek ODACI
Co-Supervisor, Department of Biochemistry
Ege University

Assoc. Prof. Dr. Ali Oğuz BÜYÜKKİLECİ
Head of the Department of
Biotechnology and Bioengineering

Prof. Dr. Mehtap EANES
Dean of the Graduate School of
Engineering and Science

ACKNOWLEDGMENTS

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Asst. Prof. Dr. Abdurrahman GÜMÜŞ, for his invaluable guidance and support throughout my research. His insightful comments and suggestions have been instrumental in helping me overcome challenges and achieve my goals.

I am also deeply honored to have worked with my co-advisor, Prof. Dr. Dilek ODACI. I sincerely thank her for providing laboratory facilities and her valuable contributions. My special thanks go to PhD student Sultan Sacide GELEN for her assistance in the production of biosensor data.

I would like to extend my sincere thanks to Assoc. Prof. Dr. Kıvılcım YÜKSEL ALDOĞAN for her help with my optical studies, and to PhD student Şamil ŞİRİN for his assistance in the production of optical data. Their knowledge and contributions have significantly shaped the optical aspect of my research.

Lastly, I would like to thank my family for their unwavering support.

ABSTRACT

DEEP LEARNING-BASED ANALYSIS OF ELECTROCHEMICAL, BIOMEDICAL, AND OPTICAL SIGNALS

This thesis explores the applications of deep learning (DL) techniques across various domains, demonstrating significant improvements in the detection, classification, and analysis of complex data. The study integrates DL models with different analytical methods to enhance performance in several fields. In the field of electrochemical analysis, a DL-based approach using an immunobiosensor was developed for the detection and classification of CD36. Traditional techniques often fall short in sensitivity and rapid analysis, especially at low analyte concentrations. The integration of DL models such as 1D-CNN and hybrid 1D-CNN – LSTM networks significantly improved the biosensor's sensitivity and specificity. For biomedical applications, Vision Transformers (ViT) techniques were employed to classify hand movements using surface electromyography (sEMG) signals. By analyzing sEMG data with advanced time-frequency analysis (TFA) methods and various ViT models, high accuracy was achieved. In optical sensing, DL techniques were applied to analyze Phase-Optical Time-Domain Reflectometry (Phase-OTDR) data. The use of DL methods, including 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bi-LSTM models, enhanced the efficiency of Phase-OTDR-based current sensing systems. Additionally, a method to convert optical signals into images for classification using Transfer Learning models was implemented, resulting in high classification accuracy and more efficient data storage. This thesis demonstrates the potential of integrating DL techniques with various analytical methods to achieve significant advancements. The studies show DL's versatility in enhancing data analysis performance, offering more accurate, sensitive, and efficient solutions. The methodologies developed can be extended to other biomarkers, signal types, and analytical challenges.

ÖZET

ELEKTROKİMYASAL, BİYOMEDİKAL VE OPTİK SİNYALLERİN DERİN ÖĞRENME TABANLI ANALİZİ

Bu tez, derin öğrenme (DÖ) tekniklerinin çeşitli alanlardaki uygulamalarını inceleyerek, karmaşık verilerin tespiti, sınıflandırılması ve analizi konularında önemli iyileştirmeler sağlamaktadır. Çalışma, DÖ modellerini farklı analitik yöntemlerle entegre ederek performansı artırmayı amaçlamaktadır. Elektrokimyasal analiz alanında, CD36'nın tespiti ve sınıflandırılması için bir immünobiyosensör kullanılarak DÖ tabanlı bir yaklaşım geliştirilmiştir. Geleneksel teknikler, özellikle düşük analit konsantrasyonlarında duyarlılık ve hızlı analizde yetersiz kalmaktadır. Tek boyutlu evrişimli sinir ağı (1B-ESA) ve hibrit 1B-ESA – uzun kısa süreli bellek (UKSB) ağları gibi DÖ modellerinin entegrasyonu, biyosensörün duyarlılığını ve özgüllüğünü önemli ölçüde artırmıştır. Biyomedikal uygulamalarda, yüzey elektromiyografi (yEMG) sinyalleri kullanılarak el hareketlerinin sınıflandırılması için Vision Transformer (ViT) teknikleri kullanılmıştır. sEMG verileri, ileri zaman-frekans analizi (TFA) yöntemleri ve çeşitli ViT modelleri ile analiz edilerek yüksek doğruluk elde edilmiştir. Optik algılama alanında, Faza Duyarlı - Zaman Bölgesinde Optik Geriyansımalı Ölçüm Tekniği (Faz-OTDR) verilerinin analizi için DÖ teknikleri kullanılmıştır. DÖ yöntemlerinin Faz-OTDR tabanlı akım algılama sistemlerinin verimliliğini artırdığı gösterilmiştir. 1B-ESA, 1B-ESA – UKSB ve 1B-ESA – Çift yönlü UKSB modelleri kullanılarak, akım değerlerinin doğru bir şekilde sınıflandırılması sağlanmıştır. Ayrıca, optik sinyalleri görüntüye çevirme metodu uygulanarak, aktarımlı öğrenme modelleri ile yüksek sınıflandırma doğruluğu elde edilmiş ve veri depolama daha verimli hale getirilmiştir. Bu tez, DÖ tekniklerinin çeşitli analitik yöntemlerle entegrasyonunun önemli ilerlemeler sağlama potansiyelini göstermektedir. Çalışmalar, DÖ'nün veri analizi performansını artırmadaki çok yönlülüğünü, daha doğru, hassas ve verimli çözümler sunarak ortaya koymaktadır. Geliştirilen metodolojiler, diğer biyomarkerlar, sinyal türleri ve analitik zorluklara genişletilebilir.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

Figure **Page**

Figure **Page**

Figure **Page**

Figure **Page**

Figure **Page**

Figure **Page**

Figure **Page**

LIST OF TABLES

CHAPTER 1

INTRODUCTION

In nature, learning and decision-making are not ends in themselves; the primary goal is survival. Making the right decisions paves the path to this goal, and this process is a complex system that has evolved over countless generations. As humans, we learn because we live in a world full of uncertainties and there are reliable pieces of information we have yet to comprehend. We make predictions, correct our errors, and through this ongoing process, we continuously 'learn'. The constant change in the world means that our predictions must also evolve over time. In terms of computational capacity, we may fall behind a basic calculator, but when it comes to analytical skills, filtering critical information, understanding context, and predicting human behavior, we are still very adept.

Humans' ability to communicate, pass on knowledge to subsequent generations, and their curiosity have led to the invention of devices that make life easier, prompting continuous development of both themselves and their inventions. Technology has emerged as a result of these advancements, comprising tools, machines, systems, and methods that meet human needs, solve problems, or facilitate life. With the advancement of technology, it has become possible for machines to learn as well. Initially, systems operating on simple rules were developed, which over time have become more complex and sophisticated.

As a result of these advancements, machine learning (ML), a subset of artificial intelligence (AI), has emerged. ML enables machines to learn from data and improve their performance. The most advanced step in this field so far is deep learning (DL). DL, inspired by the neural networks of the human brain, uses algorithms to tackle much more complex tasks. By using layered neural networks on large datasets, DL has achieved unparalleled success in recognizing patterns and extracting meaning. Consequently, machines can now perform at or even surpass human levels in areas such as image recognition, language processing, gaming, and autonomous driving. The advent of deep learning has been a milestone in artificial intelligence, enabling machines to acquire human-like learning and adaptation processes. These developments have made it possible for machines to make more effective decisions and adapt to changing conditions in our uncertain world.

DL's versatility and robustness have made it an invaluable tool across various fields, including electrochemical, biomedical, and optical signal analysis. This thesis explores the application of deep learning techniques to analyze and interpret complex data from these diverse domains. By leveraging the power of deep learning, we can uncover hidden

patterns, enhance signal detection, and improve classification accuracy in ways that were previously unattainable with traditional methods.

In the realm of electrochemical analysis, DL models can significantly enhance the detection and classification of analytes at various concentrations, including levels below the limit of detection. This capability is crucial for applications such as medical diagnostics, environmental monitoring, and biochemical research, where precise and sensitive detection is paramount. We developed a DL-based approach for the detection and classification of CD36 using an immuno-biosensor. By integrating advanced neural network architectures like One Dimensional Convolutional Neural Network (1D-CNN) and hybrid 1D-CNN + Long Short-Term Memory (LSTM) networks, we improved the sensitivity and specificity of electrochemical sensors, leading to more reliable and accurate results.

In biomedical applications, deep learning facilitates the classification of hand gestures using surface electromyography (sEMG) signals. This technology is particularly beneficial for the development of advanced prosthetic devices and human-computer interaction systems. By transforming sEMG data into a format suitable for DL models, we can achieve high accuracy in gesture recognition. Combining sEMG with advanced time-frequency analysis (TFA) methods and various Vision Transformer (ViT) models, including base ViT, Shifted Window Transformer (Swin), and Multi-Axis Vision Transformer (MaxViT), the results demonstrated that ViT models achieved near-perfect accuracy, underscoring the potential of combining sEMG with ViT models.

In the field of optical sensing, deep learning techniques are employed to analyze Phase-Optical Time-Domain Reflectometry (OTDR) data for precise electrical current sensing. By integrating deep learning with phase-OTDR technology, we can enhance the accuracy and reliability of current measurements. Our study introduced DL methods to improve the efficiency of Phase-OTDR-based current sensing systems, utilizing models such as 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bidirectional-LSTM. The results highlighted the effectiveness of these models in accurately classifying current values, even in the presence of noise.

Furthermore, transforming 1D phase-OTDR base measured event data into images and classifying them using transfer learning models represents a novel approach to optical signal analysis. By converting the data into a visual format, we can leverage the power of pre-trained convolutional neural networks (CNNs) to classify and interpret the signals. Our study utilized transfer learning models, achieving high accuracies in classifying various disturbance events along optical fibers. This image-based classification approach not only improved classification accuracy but also facilitated more efficient data storage and analysis, paving the way for future advancements in fiber optic sensing data analysis.

1.1. Deep Learning Models Utilized in This Thesis

1.1.1. One-Dimensional Convolutional Neural Network (1D-CNN)

One-dimensional Convolutional Neural Network (1D-CNN) are highly effective at identifying specific patterns and unique features within datasets (Kiranyaz et al., 2019; Meliboev, Alikhanov, and Kim, 2020). By leveraging convolution operations, 1D-CNN can exploit spatial relationships in the data, enabling the automatic learning of discriminative features and patterns. This capability makes 1D-CNN suitable for a range of tasks, including signal classification, anomaly detection, and natural language processing (Krohling and Krohling, 2023; Mo, Han, and Wang, 2023; W. Wang et al., 2017).

Figure 1.1 shows a 1D-CNN architecture with one convolutional layer. This network applies a set of filters to the input data using convolutional layers. These filters slide over the input and extract features at different spatial locations. Each filter produces a feature map, passing through a non-linear activation function. The network also uses pooling layers to reduce the size of the feature maps and extract the most significant features. Finally, the flattened output from the last pooling layer is passed through fully connected layers for classification or regression.

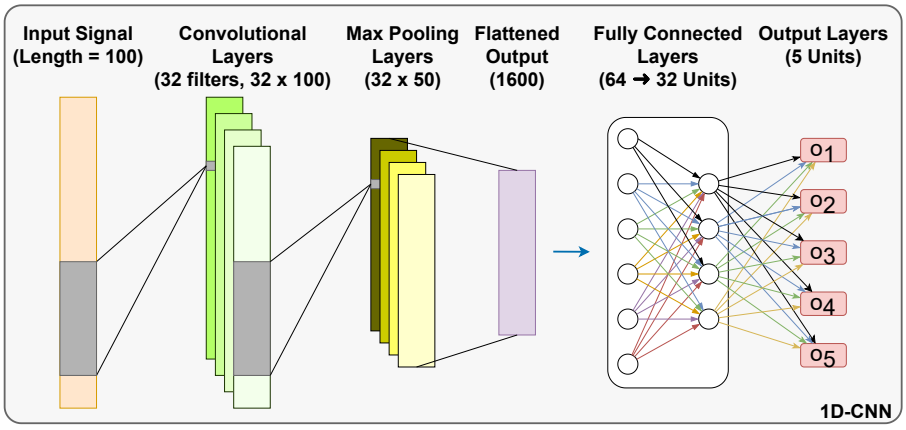

Figure 1.1. Representative 1D-CNN architecture.

The convolution operation in 1D-CNN can be mathematically expressed as follows:

$$x_l^{(m)} = \sigma\left(\sum_{c=1}^{C} w_l^{(c,m)} * x_{l-1}^{(c)} + b_l^{(m)}\right) \tag{1.1}$$

In this equation, $x_l^{(m)}$ represents the m-th channel of the input in the l-th layer. σ denotes the activation function, $w_l^{(c,m)}$ represents the filter weights from the c-th channel in the $(l-1)$-th layer to the m-th channel in the l-th layer, $x_{l-1}^{(c)}$ denotes the c-th channel of the input in the $(l-1)$-th layer, and $b_l^{(m)}$ is the bias term in the m-th channel of the l-th layer. Here, C represents the total number of channels in the previous layer, and the equation performs a convolution operation over all the channels to generate the output channel $x_l^{(m)}$ in each layer. The convolution operation $(*)$ is applied between $w_l^{(c,m)}$ and $x_{l-1}^{(c)}$.

ReLU is a commonly used non-linear activation function in 1D-CNN:

$$f(x) = \max(0, x) \tag{1.2}$$

Max pooling is used for subsampling the feature maps. This operation can be mathematically expressed as follows:

$$y_i = \max_{j=0}^{p-1} x_{l-1,ip+j}^{(c)} \tag{1.3}$$

$$y_i = \max\left(x_{l-1,ip}^{(c)}, x_{l-1,ip+1}^{(c)}, \ldots, x_{l-1,ip+p-1}^{(c)}\right) \tag{1.4}$$

Here, y_i represents the i-th element of the feature map obtained after subsampling. $x_{l-1,ip+j}^{(c)}$ represents the j-th element of the i-th segment of the c-th channel in the $l-1$-th layer, and p denotes the size of the pooling window.

The fully connected layer can be mathematically expressed as follows:

$$o_i = \sigma\left(\sum_{j=1}^{n} w_{ij} y_j + b_i\right) \tag{1.5}$$

Here, o_i represents the i-th output element, σ denotes the activation function, w_{ij} represents the weight from the j-th input element to the i-th output element, y_j represents the j-th input element, b_i represents the bias term of the i-th output element, and n represents the number of input elements.

Finally, the output layer in 1D-CNN uses the softmax activation function to generate classification probabilities:

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{1.6}$$

Here, y_i represents the probability of the i-th class, z_i represents the output of the previous layer for the i-th class, and k represents the number of classes.

1.1.2. Long Short-Term Memory (LSTM)

LSTM is a type of Recurrent Neural Network (RNN) that is designed to handle the vanishing gradient problem of conventional RNNs (Graves, 2012; Yong Yu et al., 2019). LSTMs have been widely used in various applications, such as speech recognition, natural language processing, and time series analysis, due to their ability to capture long-term dependencies in sequential data (Jelodar et al., 2020; Shewalkar, Nyavanandi, and Ludwig, 2019; Yantao Yu et al., 2022).

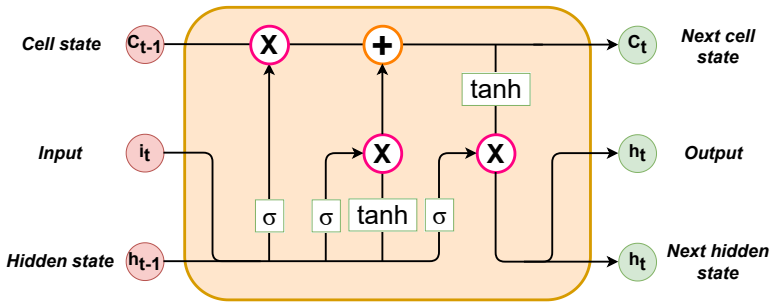

Figure 1.2. Diagram and mathematical model of LSTM cell structure and gates.

LSTM maintains a cell state that can store information over long periods, selectively forget or add information to the cell state, and output the current state to the next cell. The LSTM cell consists of three gates: forget, input, and output gates. The forget gate decides which information to forget from the cell state, while the input gate decides which information to store in the cell state. The output gate decides which information to output to the next cell. Figure 1.2 illustrates the mathematical model of an LSTM. The mathematical equation of an LSTM cell can be represented as follows:

$$f_t = \sigma_g(W_f[x_t, h_{t-1}] + b_f) \tag{1.7}$$

$$i_t = \sigma_g(W_i[x_t, h_{t-1}] + b_i) \tag{1.8}$$

$$\tilde{C}_t = \tanh(W_c[x_t, h_{t-1}] + b_c) \tag{1.9}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{1.10}$$

$$o_t = \sigma_g(W_o[x_t, h_{t-1}] + b_o) \tag{1.11}$$

$$h_t = o_t * \tanh(C_t) \tag{1.12}$$

Where f_t, i_t, and o_t are the values of the forget, input, and output gates at time step t, respectively. C_t is the cell state at time step t, which is updated by the forget gate f_t, the input gate i_t, and the cell candidate value $\tilde{C}_t$. h_t is the output of the LSTM cell at time step t, which is the element-wise product of the output gate o_t and the hyperbolic tangent of the cell state C_t.

LSTMs have been chosen over other types of RNNs due to their ability to handle the vanishing gradient problem and capture long-term dependencies in sequential data. Additionally, LSTMs have been shown to outperform other types of RNNs in various applications, such as language modeling and speech recognition. A bidirectional LSTM (Bi-LSTM) is a variant of the LSTM that processes the input sequence in both forward and backward directions and concatenates the outputs of both directions (Fang, Chen, and Xue, 2021; Li and Shen, 2017). This allows the Bi-LSTM to capture both past and future context, which can be useful in applications such as machine translation and speech recognition (Atila and Sabaz, 2022; G. Xu et al., 2019). The mathematical equation of a Bi-LSTM is similar to a regular LSTM, but it has two sets of gates and cell states, one for the forward direction and one for the backward direction. Overall, LSTMs and Bi-LSTMs have become popular choices in DL for sequence modeling tasks, and their effectiveness has been demonstrated in various applications (S. Shi et al., 2022; Tatsunami and Taki, 2022).

1.1.3. Hybrid Neural Network Models

The hybrid 1D-CNN – LSTM model combines two neural networks to process sequential data. The 1D-CNN can extract relevant spatial features from the input data, while the LSTM can model the temporal dependencies between these features. By combining these two types of networks, the 1D-CNN – LSTM model can capture both the spatial and temporal dependencies in the input data. The model merges the strengths of 1D-CNN and LSTM to extract relevant features and regulate information flow through the network (Huan et al., 2023; Rehman et al., 2019).

The 1D-CNN – LSTM model starts by applying a 1D convolutional layer to the input data, which applies a set of filters to extract relevant features from the time-series data. The resulting feature maps are then fed into a series of max-pooling layers that reduce the dimensionality of the output. Next, the output of the max-pooling layers is fed into an LSTM layer, which processes the sequential data and uses memory cells to store information about past inputs. A dropout layer follows the LSTM layer, randomly dropping some connections between the neurons to prevent overfitting. Finally, the output of the dropout layer is fed into a fully connected neural network with several dense layers, which performs the classification task.

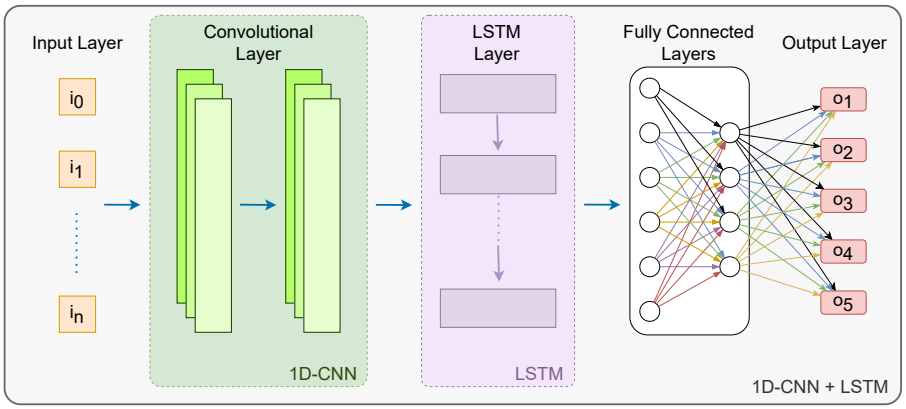

Figure 1.3. Architecture of the Hybrid 1D-CNN – LSTM model with 5 classes at the end.

The 1D-CNN – Bi-LSTM version of the hybrid model is similar to the 1D-CNN – LSTM model, but it uses a Bi-LSTM instead of a regular LSTM. The Bi-LSTM processes the input data in both forward and backward directions, allowing it to capture information about past and future inputs. This can be particularly useful in some time-series applications where future inputs may be important for making accurate predictions (Sathi et al., 2023; Yin et al., 2022).

1.1.4. Transfer Learning Models

Transfer Learning is a revolutionary concept in machine learning that enhances task-solving efficiency by transferring knowledge from one domain to another. This method is particularly beneficial for reducing the need for extensive resources, such as large datasets and high computational power, typically required for training deep learning

models. By applying knowledge learned from one task to another, Transfer Learning accelerates the training process and reduces data requirements (Z. Zhao et al., 2024; Zhuang et al., 2021).

Transfer Learning is widely applicable across various fields, including computer vision, natural language processing, and healthcare (Y. Chen et al., 2020; Taiar et al., 2022; M. Wang et al., 2023). It offers adaptability, allowing model architectures to be fine-tuned with both trainable and untrainable layers to fit specific tasks and data availability.

When implementing transfer learning, the decision to set layers as trainable or untrainable is crucial. Setting layers as "Trainable: True" retains their initial weights from pre-training but allows further fine-tuning with the new dataset for the specific task. However, this increases the number of trainable parameters, potentially leading to a longer learning process and higher computational load. This approach is advantageous when the new dataset is small and closely related to the original dataset. Conversely, setting layers as "Trainable: False" freezes these layers, preventing any updates to their weights during training for the new task. This reduces the number of trainable parameters, expediting the training process but potentially limiting adaptability to the new task. This option is beneficial when the new dataset is significantly different or when the original dataset is much larger than the new one.

In conclusion, Transfer Learning is pivotal in machine learning, making data-driven solutions more accessible and accelerating the development of intelligent systems. Notable Transfer Learning architectures include:

VGG-16 and VGG-19: Developed by the Visual Geometry Group (VGG) at the University of Oxford, these models are known for their straightforward yet effective design (Simonyan and Zisserman, 2014). VGG-16 has 16 convolutional layers, and VGG-19 has 19, both ending with three fully connected layers. They are characterized by their uniform filter size (3x3) and deep architecture, making them suitable for image classification and feature extraction.

ResNet50, InceptionV3, and InceptionResnetV2: ResNet, introduced by Kaiming He and his team at Microsoft Research Asia, uses residual blocks with skip connections to address the vanishing gradient problem (He et al., 2016). ResNet-50 has 50 convolutional layers. InceptionV3, developed by Google Research, uses inception blocks with parallel convolutional layers of varying kernel sizes. InceptionResNetV2 combines the strengths of both ResNet and Inception, featuring residual inception blocks. These architectures are crucial for tasks like image recognition, object detection, and image segmentation.

MobileNet and EfficientNet-B0: MobileNet, introduced by Google Research, features depthwise separable convolutions to reduce computation and model size while

maintaining performance (Howard et al., 2017). It is designed for mobile and embedded applications. EfficientNet, developed by Mingxing Tan and Quoc V. Le at Google, balances network depth, width, and resolution using a compound scaling technique (Tan and Le, 2019). EfficientNet-B0, the base model, offers excellent performance with a relatively small number of parameters, making it ideal for mobile and edge computing scenarios.

DenseNet: Introduced by Gao Huang et al., DenseNet features dense blocks where each layer is connected to every other layer within the block (Huang et al., 2022). This connectivity improves gradient flow, enhances feature reuse, and reduces the number of parameters. DenseNet-121, a popular variant, has 121 convolutional layers and is particularly powerful in medical image analysis and similar tasks requiring fine-grained pattern recognition.

1.1.5. Base Vision Transformer (ViT)

The Vision Transformer (ViT) is a groundbreaking model architecture introduced by Google AI, building upon the Transformer framework initially presented in the seminal paper "Attention Is All You Need" by Vaswani et al. (Vaswani et al., 2017). Unlike traditional convolutional neural networks (CNNs) that rely on convolutional operations to process images, ViT leverages the Transformer architecture and self-attention mechanisms to capture long-range dependencies within images.

ViT begins by dividing an input image into a grid of fixed-size patches. Each patch is then flattened into a vector and linearly projected into a lower-dimensional embedding space. Positional embeddings are added to these patch embeddings to retain spatial information, which is crucial for understanding the structure of the image.

The sequence of patch embeddings, along with an additional learnable class token, is fed into a series of Transformer Encoder blocks. Each block consists of several components. Layer normalization is applied to stabilize and accelerate the training process. The multi-head self-attention mechanism allows the model to focus on different parts of the image simultaneously, capturing intricate patterns and relationships. Residual connections help in mitigating the vanishing gradient problem and enable the training of deeper networks. Finally, a fully connected feed-forward neural network (MLP) further processes the output of the attention mechanism.

The final output from the Transformer Encoder is then passed through a Multi-Layer Perceptron (MLP) Head, which classifies the image into one of the predefined classes.

ViT has demonstrated impressive performance across a variety of visual tasks,

including image classification, object detection, and image generation (Dong, Zhang, and Zou, 2022; Huo et al., 2023). Achieving state-of-the-art results on benchmark datasets, ViT effectively identifies and localizes objects within images and generates high-quality images from textual descriptions or other inputs. The ability of ViT to scale to larger image sizes and learn global context information makes it a versatile and powerful tool in the field of computer vision.

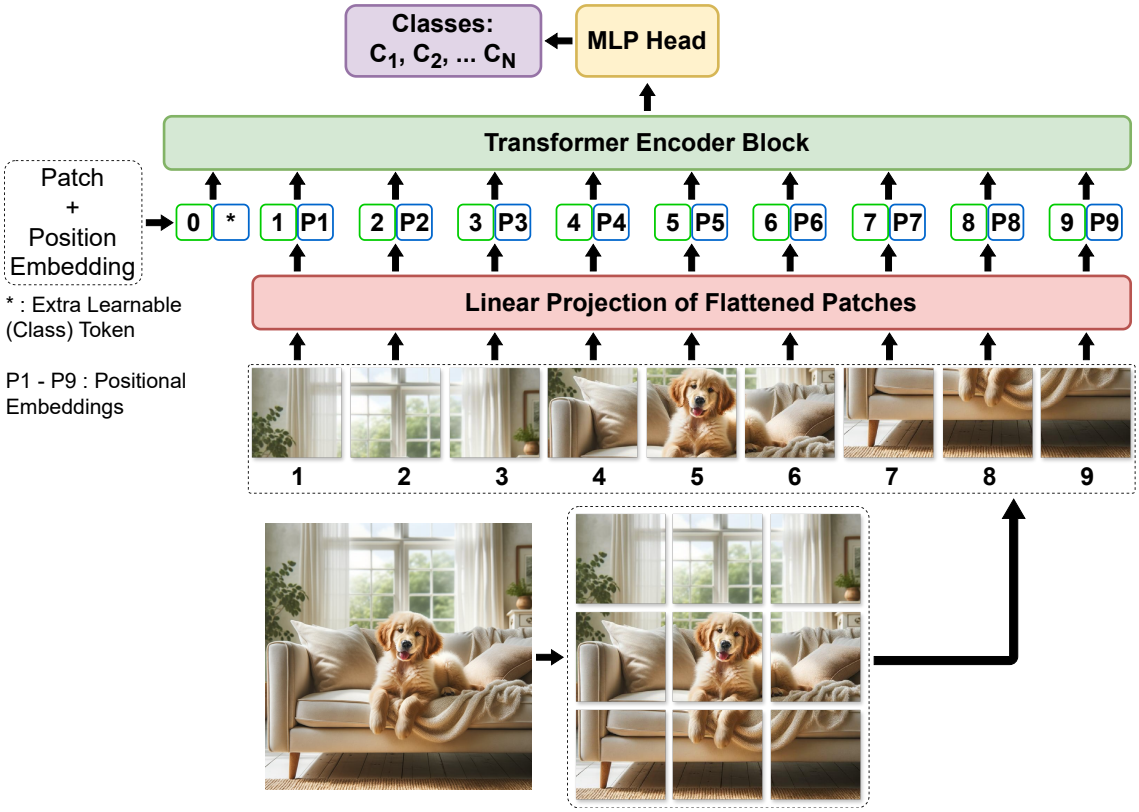

Figure 1.4. Visualizing Vision Transformer (ViT): The image is divided into patches, each of which is linearly projected and embedded with positional information. These embeddings, along with an extra learnable class token, are processed through a Transformer Encoder Block. The output is then classified by an MLP Head into one of the predefined classes.

By leveraging the power of self-attention and the Transformer architecture, ViT represents a significant advancement in the field of computer vision, offering a robust alternative to traditional convolutional approaches.

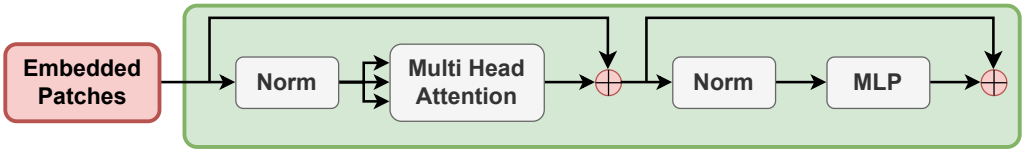

Figure 1.5. Transformer Encoder Block in Vision Transformer (ViT): The embedded patches are first normalized, then processed through a Multi-Head Attention mechanism. The output is normalized again and passed through a Multi-Layer Perceptron (MLP). The final output is obtained after residual connections are applied at each stage.

1.1.6. Shifted Window Transformer (Swin)

The Swin Transformer (Swin) is a model architecture introduced by Microsoft Research, detailed in the paper "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows" by Liu et al. (Z. Liu et al., 2021). Swin Transformers build upon the Transformer architecture, introducing shifted window-based self-attention mechanisms to efficiently capture both local and global dependencies in images.

Unlike the Vision Transformer (ViT), which processes the entire image at once or uses fixed-size patches, Swin divides the image into non-overlapping windows and computes self-attention within each window. This localized attention mechanism allows the model to focus on smaller regions of the image, making it computationally efficient. To capture more comprehensive spatial relationships, the windows are then shifted, introducing cross-window connections. This shifting mechanism ensures that the model can learn dependencies across different parts of the image, enhancing its ability to understand complex structures.

The hierarchical approach of Swin enables it to handle high-resolution images effectively. By processing images in a multi-scale manner, Swin can maintain computational efficiency while scaling to larger image sizes. This makes it particularly well-suited for tasks that require detailed spatial understanding, such as image classification, object detection, and semantic segmentation.

The Figure 1.6 above illustrates the Swin Transformer architecture. The process begins with the image being divided into patches, which are then linearly embedded. The model proceeds through multiple stages, each consisting of Swin Transformer blocks and patch merging operations. This hierarchical structure allows the model to progressively reduce the spatial dimensions while increasing the feature dimensions, enabling efficient

processing of high-resolution images.

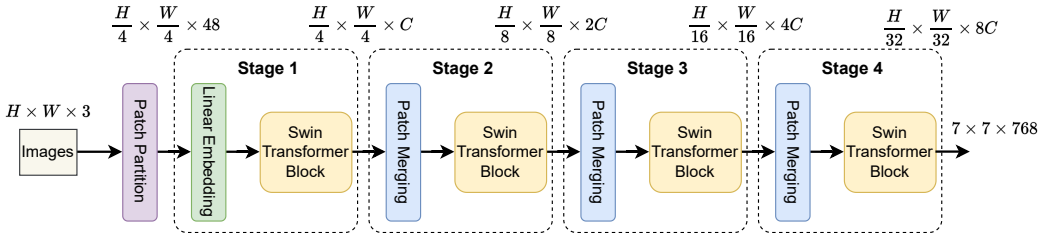

Figure 1.6. Depiction of the Swin Transformer used for image classification, explaining the process of shifted windows and embedded patch processing for feature extraction, as applied to the classification of intricate image data.

By leveraging shifted windows and hierarchical processing, the Swin Transformer represents a significant advancement in the field of computer vision. It offers a robust and efficient alternative to traditional convolutional approaches and other Transformer-based models, providing state-of-the-art performance across a variety of visual tasks.

1.1.7. Multi-Axis Vision Transformer (MaxViT)

The Multi-Axis Vision Transformer (MaxViT) is a cutting-edge vision processing model developed by Tu et al. (Tu et al., 2022). This model introduces a multi-axis attention mechanism that enhances the scalability of attention mechanisms based on image sizes, enabling both global and local spatial interactions. Central to MaxViT is the multi-axis self-attention (Max-SA) module, which efficiently integrates local and global spatial interactions within a single block while maintaining linear computational complexity. This innovative approach allows MaxViT to capture comprehensive spatial dependencies without the heavy computational cost typically associated with full self-attention mechanisms.

The MaxViT model processes input images through a series of stages, each designed to capture different levels of spatial information. Initially, the input image of size 224x224 is processed by a convolutional layer (Conv 3x3) with a stride of 2, reducing the spatial dimensions to 112x112. This initial stage is referred to as the Stem (S0). Following this, the model consists of three main stages (S1, S2, S3), each containing multiple MaxViT blocks. Each stage processes features at different spatial resolutions: S1 processes features of size 56x56, repeated L_1 times; S2 processes features of size 28x28,

repeated L_2 times; and S3 processes features of size 14x14, repeated L_3 times.

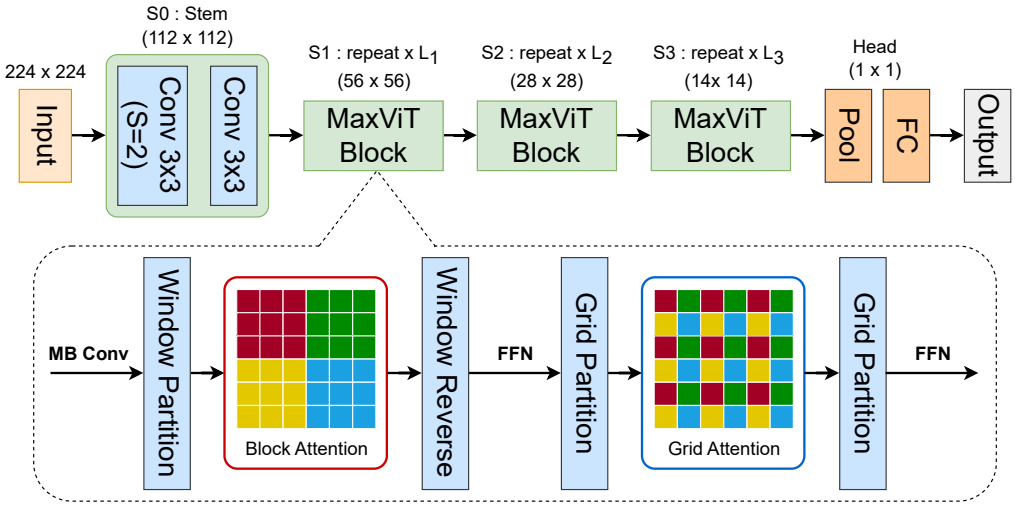

Figure 1.7. The structure of the MaxViT model.

Each MaxViT block includes a combination of window partitioning and grid partitioning mechanisms to perform block attention and grid attention, respectively. In the window partitioning mechanism, the input feature map is divided into non-overlapping windows, and each window undergoes block attention, focusing on local interactions within the window. The windows are then merged back to form the original feature map size through the window reverse process. Subsequently, the grid partitioning mechanism divides the feature map into a grid, where grid attention captures global interactions across the entire feature map. After each attention mechanism, a feed-forward network (FFN) is applied to enhance the feature representations.

The final stage of the MaxViT model includes a pooling layer (Pool) and a fully connected layer (FC) to produce the output. The pooling layer reduces the spatial dimensions to 1x1, and the fully connected layer generates the final predictions. The overall architecture of MaxViT ensures efficient and scalable processing of visual information, making it suitable for a wide range of vision tasks. The combination of local and global attention mechanisms allows the model to capture intricate details and broader context, leading to superior performance in various vision applications.

MaxViT excels in image classification, object detection, instance segmentation, image aesthetics assessment, and image generation, consistently outperforming recent state-of-the-art models. Ablation studies confirm the importance of each component in MaxViT, highlighting the critical role of grid attention for global interactions, MBConv

layers for generalization, and the optimal order of combining different layers. The sequential stacking of multi-axis attention modules proves more effective than parallel designs, and the proposed vertical layout scales better for larger models.

1.2. Model Validation Methods and Performance Metrics

Model validation methods are essential for assessing the performance and generalization capabilities of DL models. These techniques help estimate how well a model might perform on unseen data and aid in selecting the best model for real-world deployment. Two widely used validation techniques are holdout validation and K-fold cross-validation, each providing unique approaches to evaluate and validate machine learning models.

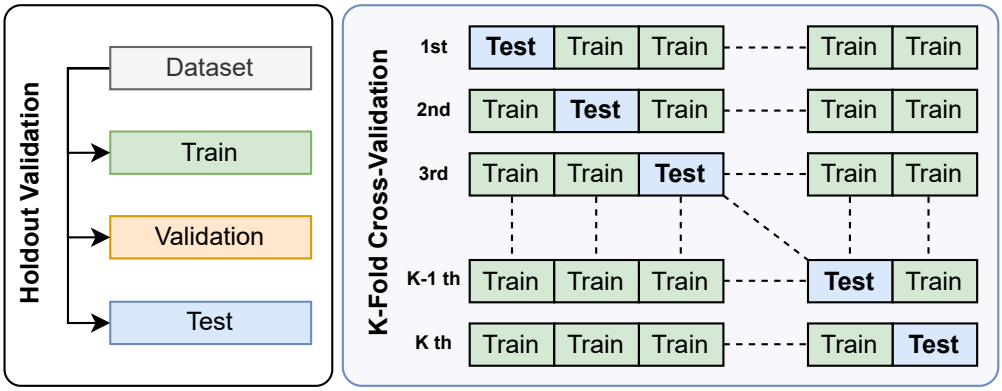

Figure 1.8. Comparison of Holdout Validation and K-Fold Cross-Validation methods. Holdout validation splits the dataset into training, validation, and test sets, while K-Fold Cross-Validation divides the dataset into K equal parts, iteratively using each part as the test set and the remaining parts for training.

Holdout validation splits the dataset into three subsets: a training set, a test set, and a validation set. This method ensures that a significant portion of the data is used for model training, while the test set is reserved for evaluating the model's performance on unseen instances. On the other hand, K-fold cross-validation divides the dataset into K equal parts or folds, using K-1 folds for training and the remaining fold for testing in each iteration. By iteratively using each fold as the test set and the others for training, this method aims for a more comprehensive assessment. It enhances the model's generalization ability by learning from diverse data subsets and averaging performance metrics across iterations. These two validation methods are illustrated in Figure 1.8.

Confusion matrices are crucial for understanding the performance of a classification model by displaying the relationships between true (actual) and predicted values. These matrices provide detailed information about true positives, false positives, true negatives, and false negatives, which are essential for evaluating the model's performance. Performance metrics derived from these confusion matrices help assess various aspects of a classification model's accuracy and reliability. Metrics such as "Accuracy" measure the ratio of correctly predicted instances to the total dataset, while "Sensitivity (Recall)" evaluates the rate of correctly identified true positives, indicating how many actual positives were accurately detected. "Precision" quantifies the proportion of true positive instances among those predicted as positive, focusing on minimizing false positives. The "F1 Score" offers the harmonic mean of precision and recall, balancing both false positives and false negatives. Additionally, Figure 1.9 visually represents the metrics used for evaluating a single event in the classification model, including formulas for metrics like Accuracy, Sensitivity, Precision, and F1 Score, which help assess the model's performance.

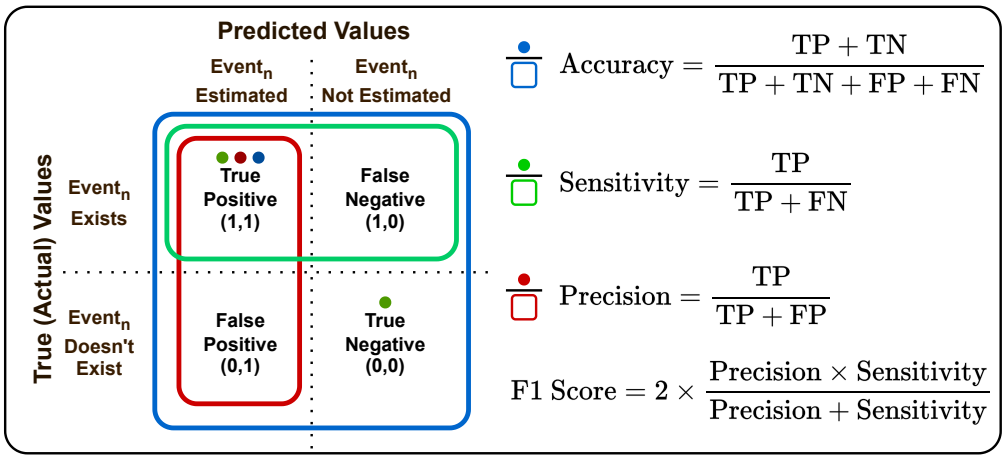

Figure 1.9. Confusion matrix and performance metrics for classification models. The matrix shows the relationships between true positives, false positives, true negatives, and false negatives. Performance metrics such as Accuracy, Sensitivity (Recall), Precision, and F1 Score are derived from these values to evaluate the model's performance.

CHAPTER 2

DEEP LEARNING-BASED DETECTION AND CLASSIFICATION OF CD36 USING AN IMMUNO-BIOSENSOR AT VARIOUS CONCENTRATIONS AND BELOW LIMIT OF DETECTION

2.1. Introduction

CD36 is a membrane protein that plays a significant role in various physiological and pathological processes, making it a critical target in biomedical research. It functions in lipid metabolism, immune response modulation, and angiogenesis (Pepino et al., 2014; Silverstein and Febbraio, 2009). In lipid metabolism, CD36 facilitates the uptake of fatty acids into adipose and muscle tissues, regulating energy homeostasis (Glatz, Nabben, and Luiken, 2022). In the immune system, CD36 aids in the recognition and phagocytosis of pathogens by immune cells, such as macrophages and dendritic cells, enhancing the body's defense mechanisms against infections (Baranova et al., 2008). Furthermore, CD36's role in angiogenesis implicates it in tumor growth and metastasis by contributing to the formation of new blood vessels that supply nutrients to growing tumors (Liao et al., 2022).

The pivotal role of CD36 in numerous biological processes emphasizes the urgent need for further investigation to fully understand its mechanisms and therapeutic potential. Consequently, advancements in CD36 detection, essential for diagnosing and monitoring various diseases, are of paramount importance. Immuno-biosensors emerge as a notable alternative to conventional methods like immunofluorescence and Western blot for CD36 analysis (Ladányi et al., 2018; Miao et al., 2001). While techniques such as flow cytometry and mass spectrometry have also been utilized, immuno-biosensors offer unique benefits (Dzik et al., 2009). They afford rapid, sensitive, and portable detection, rendering them particularly advantageous in scenarios requiring swift and on-site analysis (Bahadır and Sezgintürk, 2015). Operating through the identification and analysis of CD36 using specific antibodies, these sensors present a versatile approach for both biomedical research and diagnostic applications (Er and Odaci Demirkol, 2022; Zeybekler and Odaci, 2023).

Electrochemical immuno-biosensors operate on the principle of detecting binding events between CD36 and anti-CD36 through measurable electrical signals. These biosensors typically entail immobilizing anti-CD36 on an electrode surface. Upon binding of a CD36 antigen to the immobilized anti-CD36, it induces alterations in electrochemical properties at the electrode interface, such as impedance, current, or potential. Techniques such as cyclic voltammetry (CV), differential pulse voltammetry (DPV), and electrochemical impedance spectroscopy (EIS) are commonly utilized to discern the electrochemical responses elicited by antibody-antigen interactions (Kumar and Kalkal, 2021). However, the analysis of measurement outcomes varying with concentration necessitates expertise, and the absence of human analysis can lead to deficiencies, particularly in interpreting measurements taken at or below the Limit of Detection (LOD). At such low analyte concentration levels, traditional analysis methods may struggle to discriminate (Armbruster and Pry, 2008). Therefore, our focus lies in developing a model using deep learning (DL) techniques to detect the presence of specific analyte concentrations and accurately predict them. This model can offer high precision and accuracy in determining certain analyte concentrations and discriminating between concentrations below the LOD. Such capabilities are critical for effectively analyzing biosensor data and potentially responding more accurately to specific analyte concentrations.

DL, a powerful subfield of artificial intelligence, excels at recognizing complex patterns in datasets. The integration of DL analysis with electrochemical data confers substantial advantages for analyte classification. Zhou et al. introduced an intelligent back-propagation neural network to improve the selectivity of nonenzymatic electrochemical biosensors for glucose and lactate detection. Their study demonstrated that this approach could reliably predict concentrations in complex mixtures, showcasing significant potential for practical applications (Z. Zhou et al., 2022). Molinara et al. proposed a DL technique for the accurate detection and classification of organic pollutants in water using CV with low-cost disposable screen-printed electrodes. By modifying the electrodes with nanomaterials and transforming voltammetry data into RGB images via Gramian angular field transformations, the study achieved a classification accuracy of 100% for hydroquinone and benzoquinone using convolutional neural networks (CNNs). This approach significantly enhances the sensitivity and reliability of pollutant detection platforms (Molinara et al., 2022). Kayali et al. demonstrated the use of machine learning models to classify concentrations of ferrous ions (Fe^{2+}) in potassium ferrocyanide solutions, using square wave voltammetry (SQWV) and DPV. Their study showed that machine learning algorithms, including Random Forest and Backpropagation Neural Networks, achieved high accuracy in data classification, significantly improving the sensitivity and reliability of electrochemical detection methods (Kayali et al., 2023). Li et al. introduced an innova-

tive intelligent electrochemical sensing approach for the classification and quantification of volatile organic compounds (VOCs) in solution. By utilizing cyclic voltammetry with an ionic liquid (IL)-based electrolyte, the study analyzed various VOC analytes, achieving a 99.09% accuracy in classification using a specially designed 1-D convolutional neural network (1D-CNN). The system demonstrated impressive accuracy in both classification and quantification of methanol, with a 94.4% test accuracy and 98.18% accuracy in categorizing methanol volumes, respectively. Notably, a linear correlation between max current density and methanol volume was observed, with a LOD of 9.3 μL (Yaonian Li et al., 2024).

In our study, we harnessed DL algorithms, notably 1D Convolutional Neural Network (1D-CNN) and Long Short-Term Memory network (LSTM), to process and analyze datasets, thereby significantly enhancing the overall performance of immuno-biosensors. By employing these models, we were able to discern patterns within electrochemical signals that conventional analytical methods may overlook. This capability played a pivotal role in improving the classification accuracy of analytes by detecting subtle differences in the electrochemical responses arising from diverse antibody-antigen interactions. Moreover, our utilization of DL analysis enabled real-time data processing and decision-making, a critical aspect for point-of-care diagnostics. Additionally, this integration facilitated automated data interpretation, thus mitigating the potential for human error.

2.2. Methods

2.2.1. Preparation of PDA/Anti-CD36 Immuno-Biosensor Platform

Based on dopamine, Polydopamine (PDA) surface was created on the carbon electrode surface by electrochemical polymerization method. To obtain PDA, 1 mg/mL Dopamine solution was prepared in TBS (25 mM Tris, 140 mM HCl, 3 mM KCl) buffer. PDA was obtained from this prepared Dopamine solution using the CV technique for 5 cycles in the potential range between -0.5V and 0.5V at a scanning speed of 50 mV/s (J.-l. Wang et al., 2014).

2.2.2. Design of PDA/Anti-CD36 Immuno-Biosensor Recognition Surface

After obtaining the PDA surface, a biofunctional surface for the determination of CD36 was obtained by using Anti-CD36. 0.5% EDC (pH: 6.0; in 50 mM phosphate buffer) was first dropped onto the PDA surface and incubated for 15 minutes (Chiou and Wu, 2004). Then, the electrode surfaces were washed with pure water. After the surfaces were dried, 4.5 μL of 10 μM Anti-CD36 (pH: 7.4; 50 mM PBS) was dropped. It was left to incubate at +4 °C for 24 hours.

2.2.3. Electrochemical Measurement Techniques

2.2.3.1. Cyclic Voltammetry (CV)

Cyclic Voltammetry is a method used to explore the kinetics and mechanisms of electrochemical reactions. This technique involves varying the potential of an electrode at a controlled rate and recording the resulting current (Kissinger and Heineman, 2005). CV is applied in various fields such as studying redox reactions, identifying electrochemical mechanisms, characterizing electrode surfaces, and developing sensors (Ciobanu et al., 2023; González and Sequí-Castellano, 2021). During the potential sweep, the current generated by redox reactions at the electrode surface is measured, and the resulting current-potential plot (voltammogram) is analyzed to understand reaction mechanisms and kinetic parameters. The benefits of CV include fast and sensitive measurements, the ability to operate over a wide potential range, and the provision of both analytical and mechanistic insights. However, its drawbacks include the need for complex data interpretation and potential interference from other redox-active species in the solution.

2.2.3.2. Differential Pulse Voltammetry (DPV)

Differential Pulse Voltammetry is a technique used in electrochemical analysis to achieve high sensitivity and low detection limits. This method involves incrementally increasing the potential and applying short pulses at each step. DPV is employed in appli-

cations such as detecting trace amounts of analytes, biosensors, environmental analyses, and pharmaceutical analyses (Mamun et al., 2023; Noureldin et al., 2023; Yolchuyev and Aydoğmuş, 2023). The electrode potential is increased in small steps, and a short pulse is applied at each step while measuring the current. The differential current (the difference in current between two steps) is recorded, and the resulting differential current-potential graph is analyzed to obtain information about analyte concentration and other parameters. The advantages of DPV include high sensitivity and low detection limits, low noise levels, and the ability to work with small sample volumes. However, its disadvantages include the requirement for more complex instrumentation and longer analysis times. Both techniques play significant roles in electrochemical analysis and are used in different application areas; CV is suitable for more general and broad analyses, while DPV is preferred for more specific and sensitive measurements.

2.2.4. Electrochemical Characterization of the PDA/Anti-CD36 Immuno-Biosensor System

Electrochemical characterization of whether each modification step of the developed PDA/Anti-CD36 immuno-biosensor system was successful was performed using CV, DPV and EIS techniques. For this purpose, measurements were taken after each modification (Blank SPCE, PDA, PDA/Anti-CD36 and PDA/Anti-CD36/CD36). CV, DPV, and EIS measurements were performed with PBS (50 mM PBS, pH 7.4) containing 5 mM hexacyanoferrate (III) ($K_3[Fe(CN)_6]$, HCF) and 0.1 M KCl. CV and DPV measurements were performed in the potential range of -0.4 to +0.8 V with a scan rate of 20 mV/s.

2.2.5. Analytical Characterization and Applications of the PDA/Anti-CD36 Immuno-Biosensor

2.2.5.1. Establishing the Linear Detection Range and Determining the LOD for CD36 Determination

Analytical applications of the PDA/Anti-CD36 immuno-biosensor were performed using the DPV technique. To create the calibration scheme of the PDA/Anti-CD36

immuno-biosensor system, 12 different CD36 solutions were prepared in the concentration range of 0.005 ng/mL to 50 ng/mL. PDA/Anti-CD36 was applied to the immuno-biosensor surfaces and incubated for 30 minutes at room temperature. After incubation, the surfaces were washed with distilled water. Once the surfaces were dried, electrochemical measurements were taken three times.

Figure 2.1. Schematic illustration of immuno-biosensor surface preparation.

To determine the LOD of the PDA/Anti-CD36 immuno-biosensor, 10 different electrodes were prepared. Electrochemical measurements were taken in the presence of 0.05 ng/mL CD36, which is the lowest concentration in the calibration chart. Using the obtained data, the LOD of the PDA/Anti-CD36 immuno-biosensor was calculated. The scheme for the preparation of the immuno-biosensor system for CD36 determination is shown in Figure 2.1.

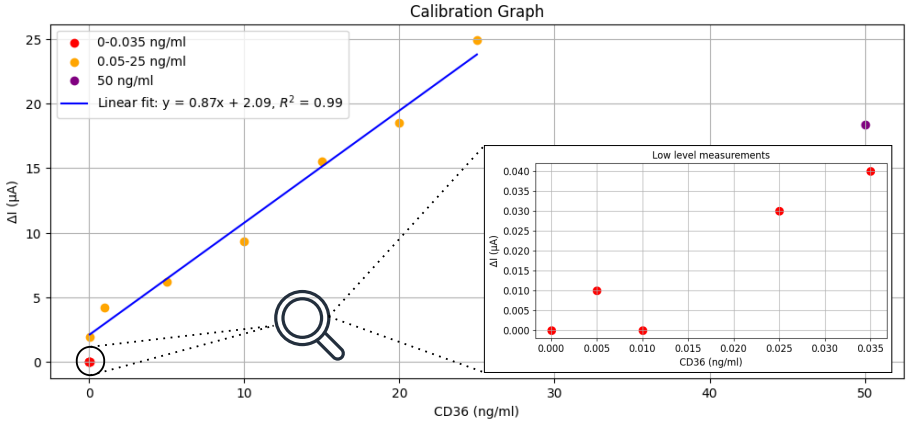

Figure 2.2. Calibration graph showing the effect of CD36 concentration on the response in the biosensor system. The linear detection range for CD36 determination with PDA/Anti-CD36 modified SPCE is highlighted, with a linear fit equation of $y = 0.87x + 2.09$ and $R^2 = 0.99$. Measurements were taken at a scan rate of 50 mV/s in 10 mL PBS (pH 7.4) with 5 mM $K_3[Fe(CN)_6]$ and an anionic redox solution containing 0.1 M KCl. The inset graph details low-level measurements.

2.2.6. Dataset

The dataset comprises electrochemical data obtained from CV and DPV measurements. These data were generated using electrochemical immuno-biosensors designed to detect the CD36 antigen. Measurements were conducted across various concentrations and cycles. Each analyte concentration level is represented by a separate class within the dataset. The corresponding data are stored in Excel files (.xlsx format) organized within class-specific folders. This dataset aims to capture the relationship between analyte concentration and the corresponding voltammetric and biosensor measurements, enabling the exploration of classification and predictive modeling tasks.

2.2.7. Addressing Low CD36 Concentration Detection Challenges

For CD36 concentrations below 0.05 ng/mL (specifically 0.005, 0.01, 0.025, and 0.035 ng/mL), electrochemical differentiation was not possible using traditional methods. Therefore, DL analysis was employed to distinguish between these lower concentrations. Multiple measurements were taken for each of the 12 different CD36 concentrations using both CV and DPV techniques, providing a comprehensive dataset for analysis. DL algorithms were applied to this dataset to enhance the detection sensitivity and accuracy at these lower concentrations.

2.2.8. Data Augmentation

The quantity of data is crucial for effectively training artificial intelligence models. A substantial amount of data ensures that the models can learn intricate patterns and relationships within the dataset. However, obtaining a large volume of data can be challenging economically and in time constraints. To address this challenge, we employed data augmentation techniques to generate additional data without extensive experimental efforts. In our experimental setup, which involves different cycles, we specifically focused on analyzing the initial five cycles of each measurement. This decision was motivated by the potential variability between cycles, particularly evident in CV and DPV measurements. By concentrating on the first five cycles, our aim was to capture the early behavior of the analytes during the voltametric measurements while minimizing potential variations

arising from factors such as electrode condition and analyte diffusion kinetics.

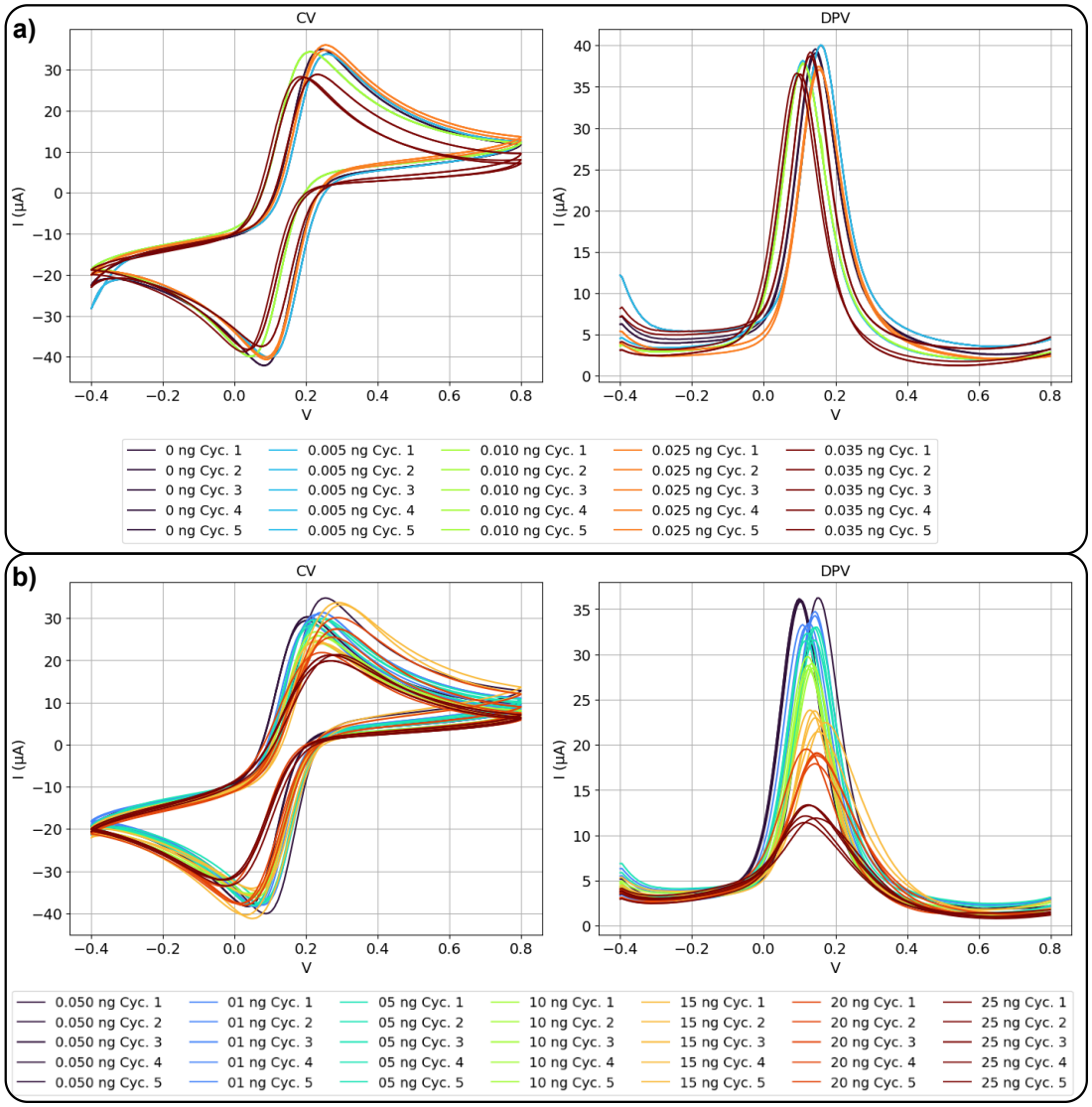

Figure 2.3. CV and DPV measurements for an analyte at various concentrations. The CV and DPV curves show the current (I) in microamperes (μA) versus the potential (V) in volts (V) across multiple cycles (Cyc.). The legend indicates different concentrations and cycles, with each color representing a specific concentration and cycle number. Panel (a) represents measurements below the LOD, while panel (b) represents measurements ranging from 0.050 ng (LOD) to 25 ng.

Data augmentation was performed using convex combination methods on measurements taken for each analyte over 5 cycles. For each class, the process involved creating linear combinations of signals between pairs of cycles. Specifically, 50 new signals were

generated for each pair of cycles using the following formula:

$$\text{new\_signal} = (1 - \alpha) \times \text{signal1} + \alpha \times \text{signal2} \tag{2.1}$$

where α is a weighting factor ranging from 0 to 1. Starting from the first cycle, new signals were generated as linear combinations between each pair of cycles, and all newly generated signals, along with the original cycle signals, were combined to create the final dataset. 500 new signals were generated for each class. These new signals were added to the existing dataset, resulting in a more extensive and diverse dataset for model training. This approach aims to improve the model's generalization capability and performance. The convex combination method is effective for data augmentation, especially in cases with limited datasets. This method generates new data points between existing examples, enabling the model to learn different patterns and enhancing model performance. This augmentation process ensured that our dataset met the requirements for DL by providing sufficient data for the models to learn from. Through the augmentation of our dataset, we empowered our DL models to discern more patterns and achieve superior results. This augmented dataset facilitated the production of more accurate and reliable outcomes, thereby enhancing the overall performance of our study.

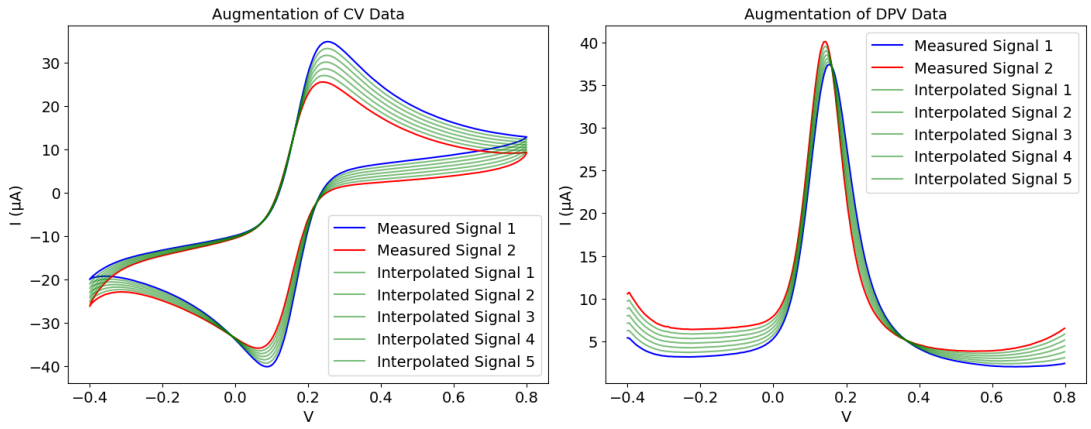

Figure 2.4. Five signals interpolated using convex combination between Cycle 1 as signal 1 and Cycle 5 as signal 2 for 0.025 ng of analyte are depicted with green lines.

2.2.9. Deep Learning Algorithms

2.2.9.1. One-Dimensional Convolutional Neural Network (1D-CNN)

In this study, we utilized a 1D-CNN to analyze CV and DPV data. The 1D-CNN can effectively capture patterns in one-dimensional data, such as time series or sensor signals, making it well-suited for biosensor data analysis. The model used in this study consists of four convolutional layers with varying filter numbers and sizes, as detailed in the first column of Table 2.1. These layers are designed to extract features at multiple scales, allowing the model to capture both fine and coarse-grained data patterns. Each convolutional layer is followed by a max-pooling layer to reduce dimensionality and prevent overfitting. The output of the last pooling layer is flattened and fed into two fully connected layers with Rectified Linear Unit (ReLU) activation functions to learn high-level representations of the input data. The ReLU activation function introduces non-linearity, enabling the model to learn complex relationships between the input and output. The output layer employs the softmax activation function to generate classification probabilities for each class. We trained the model using the categorical crossentropy loss function and the Adam optimizer with a learning rate of 0.001.

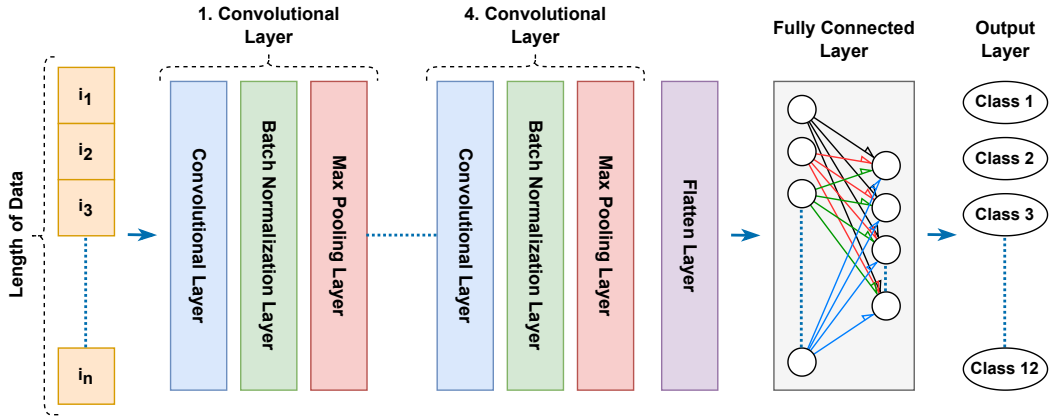

Figure 2.5. Representative 1D-CNN architecture.

Table 2.1. Comparison of the layer configurations for the 1D-CNN and 1D-CNN – LSTM models. Both models include multiple Conv1D layers with varying filter sizes and ReLU activations, followed by BatchNormalization and MaxPooling1D layers. The 1D-CNN model uses Flatten and several Dense layers for classification, while the 1D-CNN – LSTM model incorporates an LSTM layer before the Dense layers.

1D-CNN	**1D-CNN – LSTM**
Conv1D (Filters=256, Kernel Size=3, ReLU)	Conv1D (Filters=256, Kernel Size=3, ReLU)
BatchNormalization	BatchNormalization
MaxPooling1D (Pool Size=2)	MaxPooling1D (Pool Size=2)
Conv1D (Filters=128, Kernel Size=3, ReLU)	Conv1D (Filters=128, Kernel Size=3, ReLU)
BatchNormalization	BatchNormalization
MaxPooling1D (Pool Size=2)	MaxPooling1D (Pool Size=2)
Conv1D (Filters=64, Kernel Size=3, ReLU)	Conv1D (Filters=64, Kernel Size=3, ReLU)
BatchNormalization	BatchNormalization
MaxPooling1D (Pool Size=2)	MaxPooling1D (Pool Size=2)
Conv1D (Filters=32, Kernel Size=3, ReLU)	Conv1D (Filters=32, Kernel Size=3, ReLU)
BatchNormalization	BatchNormalization
MaxPooling1D (Pool Size=2)	MaxPooling1D (Pool Size=2)
Flatten	LSTM (Units=64, Return Sequences=True)
Dense (Units=128, ReLU)	Flatten
Dense (Units=64, ReLU)	Dense (Units=128, ReLU)
Dense (Units=32, ReLU)	Dense (Units=64, ReLU)
Dense (Units=12, Softmax)	Dense (Units=32, ReLU)
	Dense (Units=12, Softmax)

2.2.9.2. Hybrid Neural Network Model

In this study, The 1D-CNN – LSTM model starts by applying a 1D convolutional layer to the input data, which applies a set of filters to extract relevant features from the time-series data. This is followed by a batch normalization layer to normalize the output of the convolutional layer. The normalized feature maps are then fed into a max-pooling layer that reduces the dimensionality of the output. This sequence of convolutional, batch normalization, and max-pooling layers is repeated to further refine and extract features. Next, the output of the fourth max-pooling layer is fed into an LSTM layer, which processes the sequential data and uses memory cells to store information about past inputs. Finally, the output of the LSTM layer is passed through a flatten layer, which transforms the data into a one-dimensional vector. This flattened output is then fed into a fully connected

neural network consisting of several dense layers, which perform the classification task.

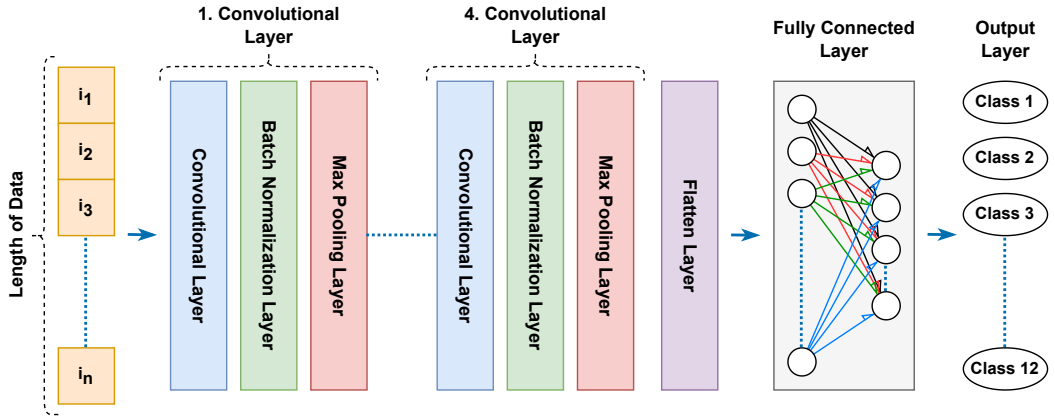

Figure 2.6. Architecture of the Hybrid 1D-CNN – LSTM model.

2.3. Results and Discussion

2.3.1. Data on Electrochemical Characterization of the PDA/Anti-CD36 Immuno-Biosensor System

Electrochemical characterizations of the PDA/Anti-CD36 immuno-biosensor surface were performed with CV, DPV and EIS measurement techniques.

CV is a popular electrochemical measurement technique widely used to provide information about electron transfer, reduction and oxidation reactions (Giffhorn, 2000). According to the CV voltammograms seen in Figure 2.7.a, I catodic values are 62.99 μA, 48.41 μA, 35.52 μA, 27.46 for blank SPCE, PDA, PDA/Anti-CD36 and PDA/Anti-CD36/CD36 surfaces, respectively. It was found to be μA. Electrochemical oxidation and reduction peaks of HCF were observed in the CV profiles of each modification step of the prepared immuno-biosensor surface. A significant decrease in the currents obtained after each modification step of PDA, PDA/Anti-CD36 and PDA/Anti-CD36/CD36 of blank SPCE was observed. This supports that all modifications were completed successfully.

DPV is a fast and sensitive technique widely used in biosensor systems. Therefore, the DPV technique is a method used to understand how the behavior of modified surfaces

changes after each modification step. Figure 2.7.b shows the voltammograms of the blank SPCE, PDA, PDA/Anti-CD36, PDA/AntiCD36/CD36 modified surfaces obtained using DPV. After each modification, the current values were found to be 47.59 μA, 21.43 μA, 16.71 μA and 9.15 μA, respectively. A significant decrease was observed in the DPV voltammograms obtained after each modification step of PDA, PDA/Anti-CD36 and PDA/Anti-CD36/CD36 of blank SPCE. The obtained DPV profiles support the CV voltammogram results.

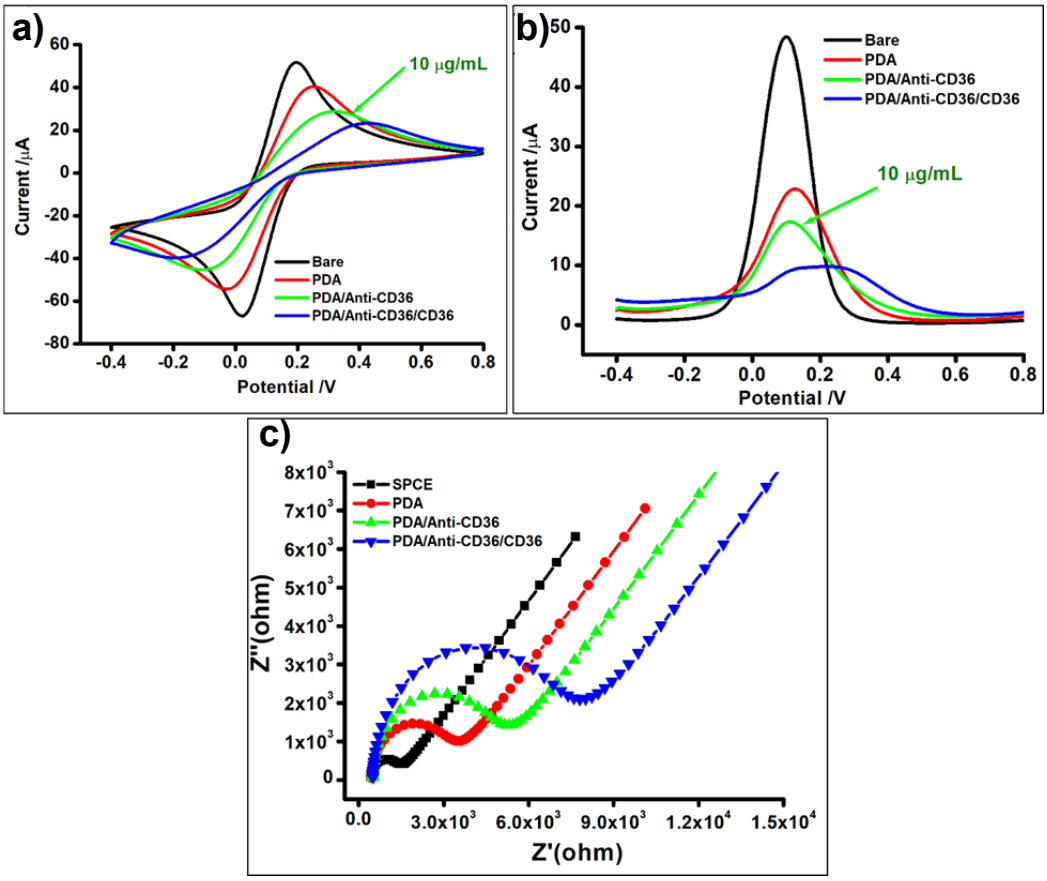

Figure 2.7. Electrochemical characterization results of blank SPCE, PDA, PDA/Anti-CD36 and PDA/Anti-CD36/CD36 surfaces a) CV voltammograms b) DPV voltammograms c) EIS diagrams (All measurements, scan rate: 50 mVs^{-1}. 10 mL PBS (pH 7.4) in the presence of an anionic redox solution containing 5 mM $K_3[Fe(CN)_6]$ and 0.1 M KCl).

Finally, the EIS technique was used for the prepared surfaces to determine how their resistance changed after each modification. In the EIS technique, the diameter of

the semicircle is proportional to the increase in load transfer resistance. Increasing the diameter of the semicircle indicates that an insulating layer is formed on the surfaces and electron transfer decreases. Figure 2.7.c shows the EIS diagrams of the empty SPCE, PDA, PDA/Anti-CD36 and PDA/Anti-CD36/CD36 surfaces. After each modification, it is seen that the resistance on the surface increases and the semicircle becomes larger. This supports the results obtained from CV and DPV techniques.

2.3.2. Data on Analytical Characterization and Applications of the PDA/Anti-CD36 Immuno-Biosensor

In order to determine the linear determination range of the PDA/Anti-CD36 immuno-biosensor, DPV measurements were performed in 12 different CD36 concentration ranges of 0-50 ng/mL. As seen in Figure 2.2, the linear determination range of the developed PDA/Anti-CD36 immuno-biosensor was determined to be 0.05-25 ng/mL. The calibration equation was determined as $y = 0.87x + 2.09$ and $R^2 = 0.9p$. it is thought that after 25 ng/mL CD36 concentration, a decrease in current intensity is observed due to CD36 creating steric hindrance on the PDA/Anti-CD36 biofunctional surface.

In order to determine the LOD of the developed PDA/Anti-CD36 immuno-biosensor, 10 different electrochemical measurements were carried out with a CD36 concentration of 0.05 ng/mL, which is the lowest point of the calibration graph. As a result of all these calculations, the detection limit of the PDA/Anti-CD36 immuno-biosensor was found to be ng/mL.

2.3.3. Deep Learning Model Performance Analysis

Figure 2.8 illustrates the performance comparison between two DL models, a 1D-CNN and a hybrid 1D-CNN – LSTM, on two different datasets: CV data and DPV data. For the CV data (panel a), the 1D-CNN model demonstrates higher stability and better overall performance in both training and validation. The 1D-CNN training accuracy rises rapidly, reaching close to 1.0, and the validation accuracy fluctuates between approximately 0.6 and 0.9. In contrast, the 1D-CNN – LSTM model's validation accuracy shows significant instability, often dipping below 0.6. Correspondingly, the validation loss for the 1D-CNN is consistently lower and more stable, staying around 0.5 after an initial drop, while the hybrid model experiences significant fluctuations with peaks above 2.0. For the DPV data

(panel b), both models achieve similar training accuracies, peaking around 1.0. However, the hybrid model's validation accuracy is slightly higher but more variable, fluctuating between 0.7 and 0.9. The 1D-CNN validation loss stabilizes around 0.5, whereas the hybrid model shows a lower and more stable validation loss, often remaining below 1.0 but with occasional spikes. These results suggest that the 1D-CNN model performs robustly and consistently on CV data, while the hybrid 1D-CNN – LSTM model exhibits potential advantages for DPV data, despite greater fluctuation in accuracy and loss.

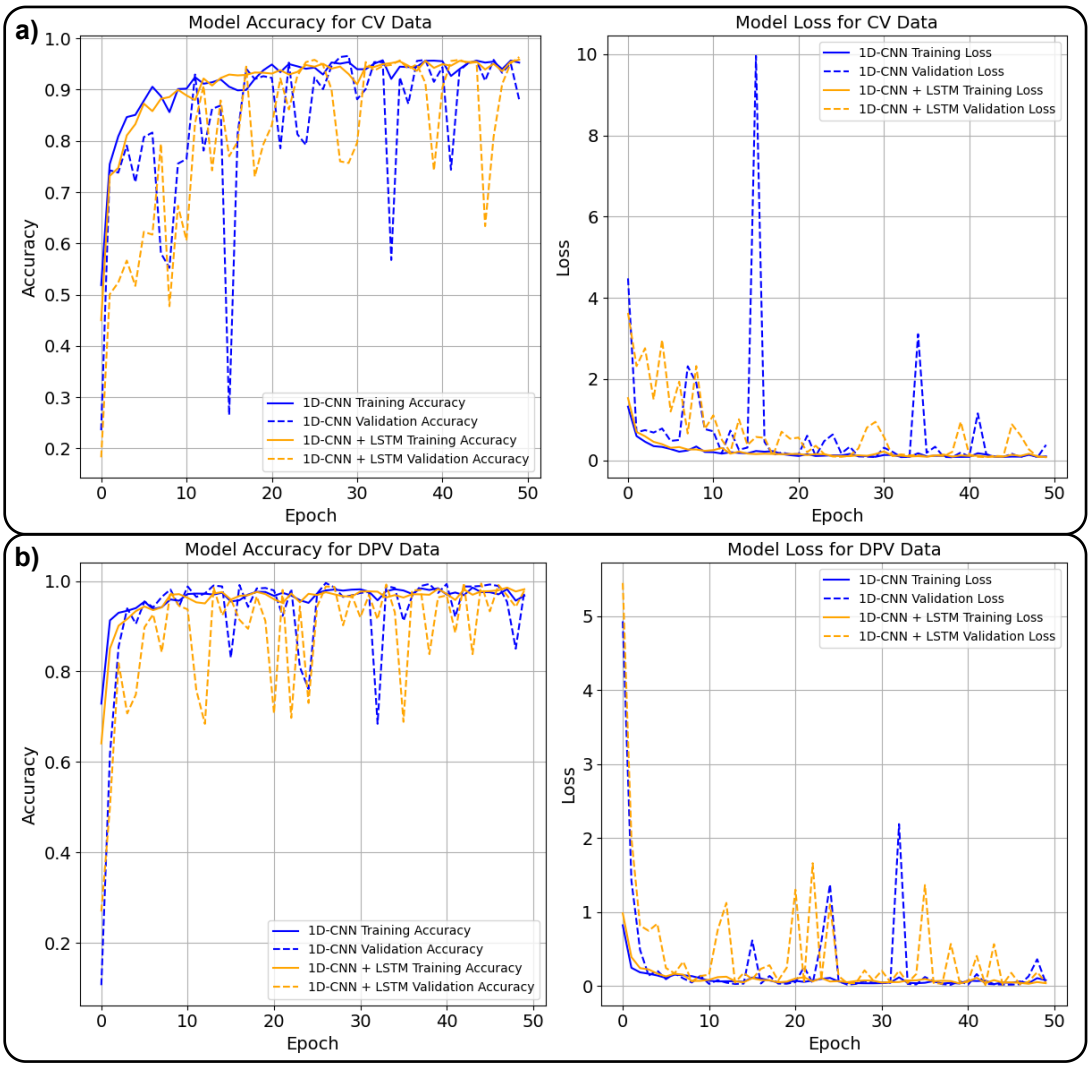

Figure 2.8. The graphs illustrate the training and validation accuracy (left) and loss (right) curves throughout the epochs. The training accuracy/loss is depicted in blue, while the validation accuracy/loss is shown in orange. These plots provide insights into the performance of the model during training, indicating how well it generalizes to unseen data.

The confusion matrices in Figure 2.9 provide a comprehensive comparison of the classification performance of the 1D-CNN and 1D-CNN – LSTM models across all analyte concentrations for both CV and DPV datasets. For the CV data, the 1D-CNN model achieves perfect classification at 0 ng, 0.010 ng, and all higher concentrations from 0.035 ng to 25 ng, but shows reduced performance at 0.005 ng with 50.00% correct classification, misclassifying 16.22% as 0 ng and 33.78% as 0.010 ng, and at 0.025 ng with 98.75% correct classification and 1.25% misclassified as 0.035 ng. The 1D-CNN – LSTM model, however, demonstrates improved performance at 0.005 ng, with 54.05% correct classification and reduced misclassifications of 13.51% as 0 ng and 32.43% as 0.010 ng, and at 0.025 ng with 96.67% correct classification and 3.33% misclassified as 0.035 ng, while maintaining perfect classification at 0 ng, 0.010 ng, and all higher concentrations.

For the DPV data, the 1D-CNN model achieves 100% correct classification at 0 ng, 0.010 ng, and all higher concentrations from 0.025 ng to 25 ng, but shows 86.49% correct classification at 0.005 ng with 13.51% misclassified as 0.010 ng. The 1D-CNN – LSTM model improves performance at 0.005 ng to 90.54% correct classification, with 9.46% misclassified as 0 ng, while also achieving perfect classification at 0 ng, 0.010 ng, and all higher concentrations.

These results indicate that the hybrid 1D-CNN – LSTM model generally offers improved performance over the 1D-CNN model, particularly at lower analyte concentrations. The LSTM component's ability to capture temporal dependencies enhances the model's discriminatory power, leading to more accurate classifications and reduced confusion between closely spaced concentrations. This demonstrates that incorporating LSTM into the CNN architecture significantly enhances the performance of electrochemical sensor data analysis, providing a robust framework for future research and development. The hybrid model's superior classification capabilities are consistent across all analyte concentrations, underscoring its potential for more reliable and precise electrochemical sensor data interpretation.

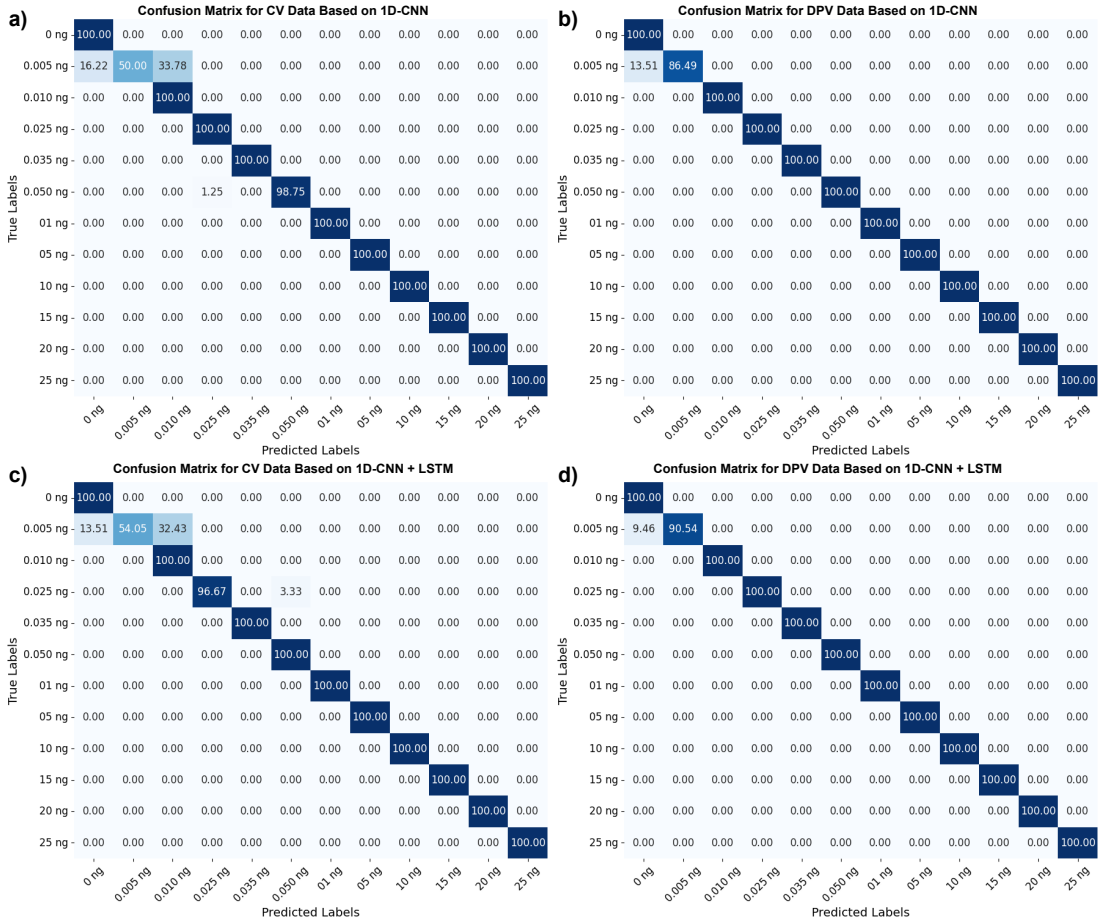

Figure 2.9. Confusion matrices showing the classification results for CV and DPV data, presented as percentages, achieved by 1D-CNN and 1D-CNN – LSTM models. (a) Confusion matrix for CV data using the 1D-CNN model, (b) Confusion matrix for DPV data using the 1D-CNN model, (c) Confusion matrix for CV data using the 1D-CNN – LSTM model, (d) Confusion matrix for DPV data using the 1D-CNN – LSTM model. The matrices illustrate the classification performance for various analyte concentrations, with true positive rates along the diagonal and misclassification rates off-diagonal.

Table 2.2 evaluates the performance of two models, 1D-CNN and a combined 1D-CNN – LSTM, across two methods: CV and DPV. For the CV method, the 1D-CNN model achieves an accuracy of 95.82%, a precision of 97.01%, a recall of 95.73%, and an F1-score of 95.52%. The addition of LSTM to the 1D-CNN slightly improves these metrics, resulting in an accuracy of 95.93%, the same precision of 97.01%, an increased recall of 95.89%, and an F1-score of 95.73%. In contrast, the DPV method demonstrates a marked improvement in performance. The 1D-CNN model under this method achieves an accuracy of 98.90%, a precision of 99.18%, a recall of 98.87%, and an F1-score of 98.96%. The combined 1D-CNN – LSTM model further enhances these metrics,

achieving an accuracy of 99.23%, a precision of 99.41%, a recall of 99.21%, and an F1-score of 99.28%. Overall, the DPV method significantly outperforms the CV method in all metrics. Additionally, incorporating LSTM with 1D-CNN consistently yields better performance in both methods, indicating the effectiveness of combining these techniques for improved model accuracy, precision, recall, and F1-score.

Table 2.2. Performance comparison of 1D-CNN and 1D-CNN – LSTM models using CV and DPV methods. The metrics evaluated are Accuracy, Precision, Recall, and F1-Score.

Method	**Model**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
CV	1D-CNN	95.82%	97.01%	95.73%	95.52%
	1DCNN + LSTM	95.93%	97.01%	95.89%	95.73%
DPV	1D-CNN	98.90%	99.18%	98.87%	98.96%
	1DCNN + LSTM	99.23%	99.41%	99.21%	99.28%

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a machine learning algorithm used for dimensionality reduction and visualization, beneficial for high-dimensional data. It maps multi-dimensional data to two or three dimensions for visualization, preserving the structure and relationships within the original data. In Figure 2.10, the t-SNE algorithm has been applied to CV and DPV data to visualize how well the 1D-CNN and 1D-CNN – LSTM models differentiate between different analyte concentrations. Panels (a) and (b) display the t-SNE results for CV and DPV data, respectively, using the 1D-CNN model. Panels (c) and (d) show the t-SNE results for the same types of data but using the 1D-CNN – LSTM model. The clusters of points in each panel represent different concentrations of the analyte. The distinct separation between clusters indicates that both models effectively capture the variations in the data corresponding to different analyte concentrations. The 1D-CNN model shows a clear clustering pattern in both CV and DPV data, suggesting it successfully distinguishes between varying concentrations. The addition of LSTM to the 1D-CNN model, as seen in panels (c) and (d), appears to enhance this differentiation further, resulting in more distinct and well-separated clusters. This improved separation highlights the superior performance of the combined 1D-CNN – LSTM model in capturing the underlying patterns of the CV and DPV data, thereby improving the model's ability to classify different analyte concentrations accurately.

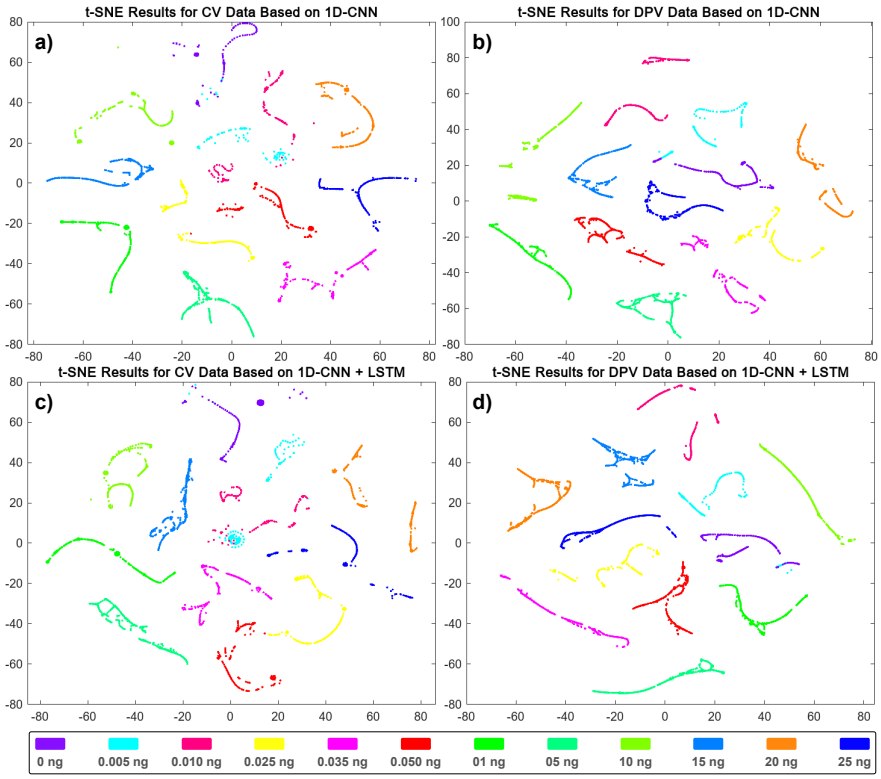

Figure 2.10. t-SNE results for CV and DPV data based on 1D-CNN and 1D-CNN – LSTM models. Panels (a) and (b) show the t-SNE results for CV and DPV data, respectively, using the 1D-CNN model. Panels (c) and (d) depict the t-SNE results for CV and DPV data, respectively, using the 1D-CNN – LSTM model. The different colors represent various concentrations of the analyte, as indicated in the legend.

The performance comparison between 1D-CNN and 1D-CNN – LSTM models on CV and DPV datasets reveals that the 1D-CNN model is more stable and performs better on CV data, with rapid training accuracy and stable validation accuracy and loss. In contrast, the hybrid model shows higher but more variable validation accuracy on DPV data. Confusion matrices indicate that the hybrid model generally offers improved performance, especially at lower analyte concentrations, due to its ability to capture temporal dependencies. Performance metrics show that the hybrid model consistently outperforms the 1D-CNN in both methods, achieving higher accuracy, precision, recall, and F1-scores. t-SNE visualizations further highlight the superior performance of the hybrid model in distinguishing between different analyte concentrations, demonstrating its effectiveness in capturing data patterns and improving classification accuracy.

2.4. Conclusion

In this study, we developed a DL-based approach for the detection and classification of CD36 using an immuno-biosensor at various concentrations, including those below the LOD. Our research highlights the significant role of CD36 in various physiological and pathological processes, emphasizing the need for advanced detection methods. Traditional techniques, while effective, often fall short in sensitivity and rapid analysis, particularly at low analyte concentrations.

We successfully fabricated a PDA/Anti-CD36 immuno-biosensor platform and characterized it using electrochemical techniques such as CV, DPV, and EIS. The results confirmed the successful modification of the sensor surface at each step, demonstrating its capability to detect CD36 with high sensitivity. The linear detection range of the developed PDA/Anti-CD36 immuno-biosensor was determined to be 0.05-25 ng/mL, with a calibration equation of $y = 0.87x + 2.09$ and an R^2 value of 0.99. The LOD was found to be 0.05 ng/mL.

To address the challenges of detecting low CD36 concentrations, we integrated DL models, specifically 1D-CNN and a hybrid 1D-CNN – LSTM network. These models were trained on augmented datasets derived from CV and DPV measurements. The data augmentation process, which involved generating 500 additional signals per class through convex combinations, significantly enhanced the dataset's size and diversity, improving the models' training and performance.

Our findings indicate that the hybrid 1D-CNN – LSTM model outperforms the standalone 1D-CNN model, particularly in classifying lower analyte concentrations. For the CV method, the 1D-CNN model achieved an accuracy of 95.82%, precision of 97.01%, recall of 95.73%, and F1-score of 95.52%. The hybrid model slightly improved these metrics, achieving an accuracy of 95.93%, precision of 97.01%, recall of 95.89%, and F1-score of 95.73%. For the DPV method, the 1D-CNN model achieved an accuracy of 98.90%, precision of 99.18%, recall of 98.87%, and F1-score of 98.96%. The hybrid model further enhanced these metrics, achieving an accuracy of 99.23%, precision of 99.41%, recall of 99.21%, and F1-score of 99.28%. The t-SNE visualizations further confirmed the models' effectiveness in distinguishing between different analyte concentrations, with the hybrid model showing more distinct and well-separated clusters.

In conclusion, the integration of DL techniques with electrochemical immuno-biosensors presents a powerful approach for the sensitive and accurate detection of CD36. This methodology not only enhances the performance of biosensors but also holds significant potential for point-of-care diagnostics and real-time data processing. Future work

could explore the application of this approach to other biomarkers and further refine the models to improve their robustness and generalizability.

CHAPTER 3

HAND GESTURE CLASSIFICATION WITH SURFACE ELECTROMYOGRAPHY (SEMG) SIGNALS USING VISION TRANSFORMERS

3.1. Introduction

Electromyography (EMG) is an essential technique for measuring and recording the electrical activity produced by skeletal muscles (Ibrahim et al., 2016; Mills, 2005). This method is pivotal in providing critical insights into muscle function and nerve-muscle interactions, thus playing a crucial role in diagnosing neuromuscular disorders such as muscular dystrophy, amyotrophic lateral sclerosis (ALS), and peripheral neuropathies (Bashford, Mills, and Shaw, 2019; Bora et al., 2021; Ullah and Iqbal, 2020). EMG's diagnostic capabilities enable clinicians to develop targeted treatment plans and monitor the progression of these conditions effectively. Beyond clinical diagnostics, EMG is widely employed in sports science to assess muscle performance, prevent injuries, and optimize training programs for athletes, underscoring its versatility in both clinical and non-clinical settings (S.-H. Liu et al., 2019; Mcmanus, Vito, and Lowery, 2020). Recent advancements in science and technology have further enhanced the application of EMG in intelligent systems, particularly within the field of Human-Computer Interaction (HCI) (Maity and Veer, 2023).

Hand gesture recognition, a critical component of HCI, significantly benefits from EMG data. By analyzing the electrical activity of muscles during various gestures, it is possible to develop intuitive control systems for a diverse range of applications (Qi et al., 2019; X. Zhou et al., 2020). This capability is especially valuable in the design of advanced medical devices, such as prosthetic limbs and robotic assistants, which require precise and natural user control (Godoy et al., 2022; Samuel et al., 2019). These devices rely on the accurate interpretation of muscle signals to perform desired actions, thereby enhancing the quality of life for users.

In recent years, deep learning (DL) methods have gained prominence in the field of hand gesture classification. DL networks, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated remarkable capability

in automatically extracting relevant features from input data, significantly enhancing classification performance. These networks can learn complex patterns from large datasets, making them highly effective for EMG signal analysis. To effectively utilize EMG data for gesture recognition, it is necessary to extract intrinsic information from the recorded signals. Conventional Time-Frequency Analysis (TFA) methods, such as Wavelet Transform (WT), and Short-Time Fourier Transform (STFT), are commonly employed to analyze EMG signals. These methods help transform raw EMG data into a format suitable for FE, which is a critical step in gesture discrimination. These methods' output can also be visualized as a 2D image, where axes represent time and frequency, enabling better understanding and analysis of signal characteristics. In 2020, Nahid et al. aimed to optimize control schemes for the DUFAB Hand prosthetic using sEMG data. They converted sEMG signals into images using the continuous wavelet transform (CWT), then employed transfer learning and CNN-LSTM architectures. Achieving high accuracies (99.72% and 99.83%), their approach outperformed traditional methods, particularly with transfer learning using AlexNet (Nahid, Rahman, and Ahad, 2020). In 2021, Elbeshbeshy et al. examined EMG signal analysis and classification for rehabilitation and motor control. They utilized time-frequency representations and DL models, achieving accuracies of 92.71% to 96.88% with pre-trained CNNs. Their end-to-end system showed superior performance over traditional classifiers, demonstrating the efficacy of DL in EMG analysis for prosthetic control and robotic applications (Elbeshbeshy, Rushdi, and El-Metwally, 2021). In 2022, Ozdemir et al. explored hand gesture classification using surface electromyography (sEMG) signals and transfer learning (TL) based on convolutional neural networks (CNNs). They collected 4-channel sEMG data from 30 subjects performing 7 hand gestures, then applied time-frequency (TF) analysis methods (STFT, CWT, HHT) to obtain TF images. Pre-trained CNN architectures were used to classify gestures. Results showed that the HHT method, combined with TL using the ResNet-50 model, achieved the best average accuracies (93.75% for SKCV and 94.41% for LOOCV) (Ozdemir et al., 2022b). In 2023, Buelvas et al. proposed a method for classifying hand gestures using EMG signals from amputee patients. They employed CWT to generate scalogram images for training a CNN, achieving high accuracy (94.49% in one-to-one methodology and 85.70% in all-to-one methodology) with low computational complexity (Buelvas, Montaña, and Serrezuela, 2023). In 2024, Mohapatra et al. introduced time–frequency domain deep neural network (TFDDNN), for recognizing hand gestures using multichannel electromyogram (MEMG) sensor data. By segmenting MEMG recordings into frames and applying CWT to obtain time–frequency representations (TFRs), they achieved robust recognition performance. The proposed TFDDNN model, integrated with a deep representation learning network (DRLN), outperformed existing methods, attaining accuracy rates of 92.73% and 80.33%

for multiclass-based recognition across different databases. This innovative approach was further implemented in a web application, enabling real-time recognition of hand gestures in Internet of Things (IoT) applications (Mohapatra, Aggarwal, and Tripathy, 2024).

In this study, we utilized the open-source dataset published by Ozdemir et al. in 2022 (Ozdemir et al., 2022a). Our primary objective is to develop advanced methods for hand gesture classification by combining Time-Frequency Analysis (TFA) of surface EMG (sEMG) signals with Vision transformers. We examined the effects of different window sizes and various transformer models on classification performance. Additionally, we performed a comparative analysis of traditional Short-Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT) methods. Our research makes significant contributions to the field by providing a high-resolution sEMG signal dataset from 40 subjects and determining the optimal number of channels for precise gesture recognition. The dataset provided by Ozdemir and colleagues is essential for training and evaluating the proposed models, ensuring their robustness and reliability.

In summary, this paper explores the integration of sEMG with advanced TFA and Vision transformers for gesture analysis, highlighting improvements in accuracy and efficiency. By leveraging the strengths of DL and signal processing techniques, we aim to enhance the performance of sEMG-based gesture recognition systems. Our findings demonstrate the feasibility and potential of using sEMG and Vision transformers for developing intuitive and responsive control systems, paving the way for future advancements in intelligent human-machine interfaces.

3.2. Methodology

3.2.1. Dataset

In our research, we made use of the open-source dataset presented by Ozdemir et al. in 2022 (Ozdemir et al., 2022a). The dataset provided constitutes a valuable resource for the Biomedical Signal Processing Society, comprising 4-channel time-series sEMG signals from 40 participants who executed 10 distinct hand gestures. This comprehensive dataset proves particularly beneficial for tasks like hand-gesture detection and classification, holding considerable promise for the development of AI models, particularly in DL research aimed at predicting hand gestures. The data is structured into repetitive cycles and encompasses both raw and filtered formats, thus facilitating its integration into DL models and serving as a benchmark for evaluating other models.

Each participant's data was meticulously captured using a 4-channel MP36 BIOPAC device with Ag/AgCl surface electrodes positioned on specific muscle sites, ensuring the acquisition of high-quality recordings. The signals were sampled at 2 kHz and underwent filtering to eliminate noise, enabling researchers to focus on muscle-generated sEMG signals. Furthermore, demographic data and responses from health-related surveys were gathered to guarantee the accuracy and relevance of the recordings. The dataset is accompanied by detailed MATLAB code for the automated segmentation of gesture moments, thereby enhancing its usability for further analysis. Overall, this dataset provides a sturdy foundation for research in hand gesture recognition, model training, and validation in machine learning and DL applications, as well as in the development of HCI and biorobotics applications.

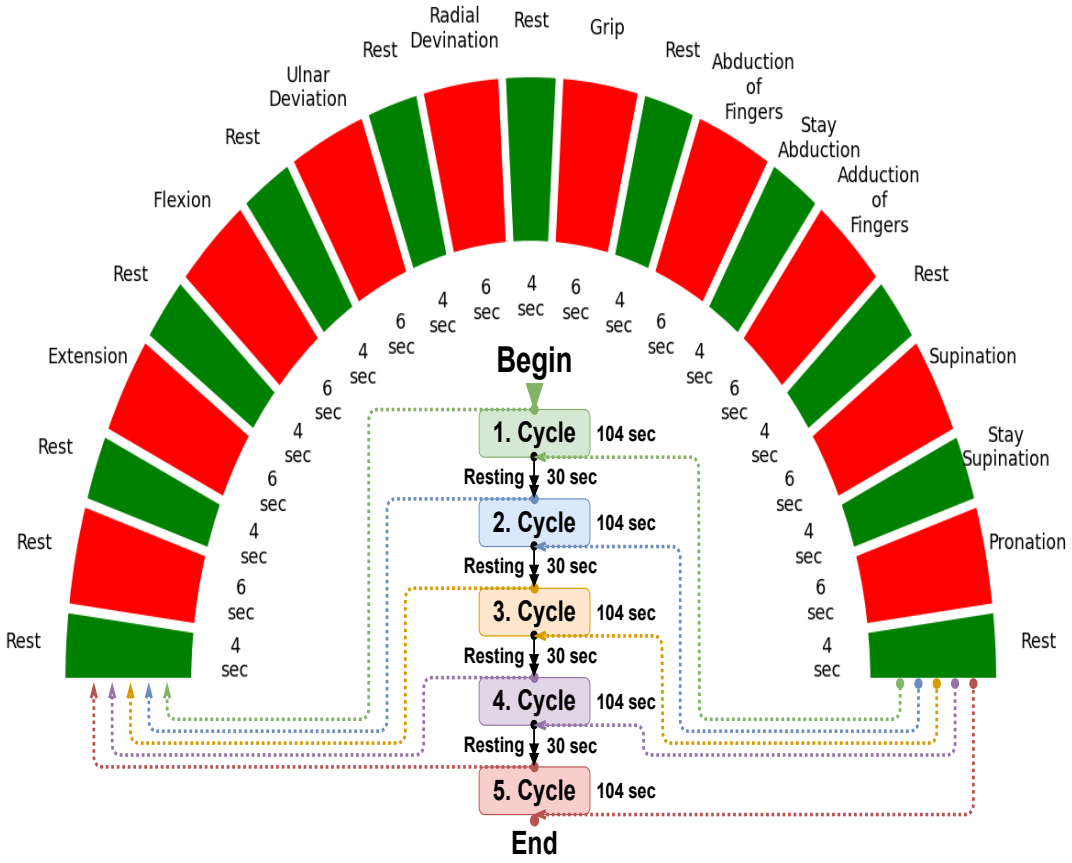

Figure 3.1. Figure depicting a series of hand gestures and rest intervals, organized into five cycles of exercises.

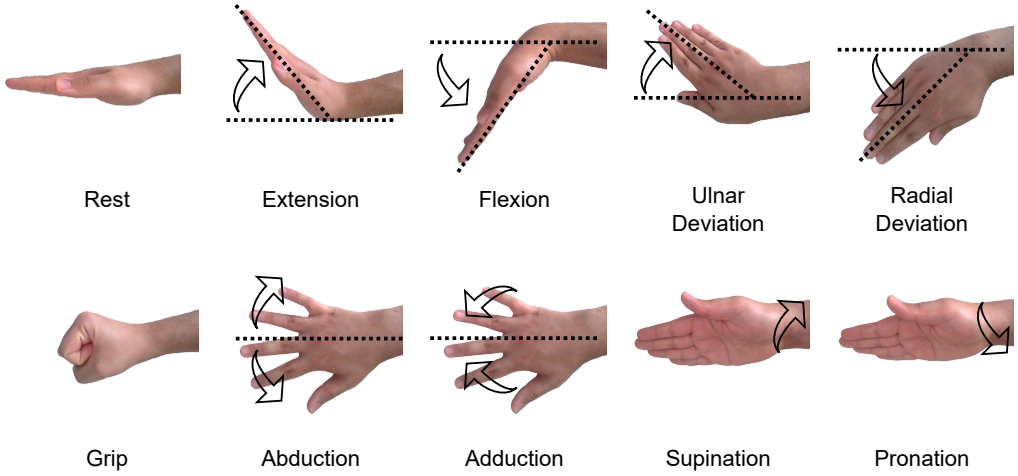

Figure 3.2. Ten different hand gestures demonstrated in sequence.

The dataset includes 10 different hand gestures, which are demonstrated in sequence in Figure 3.2. These gestures are Rest, Extension, Flexion, Ulnar Deviation, Radial Deviation, Grip, Abduction, Adduction, Supination, and Pronation. Each gesture represents a distinct position or movement of the hand.

3.2.2. Preprocessing and Windowing for sEMG Data

The dataset's raw sEMG signals undergo crucial preprocessing steps to extract the muscle activity signal and eliminate noise. A sixth-order Butterworth band-pass filter ranging from 5 to 500 Hz is applied to remove artifacts, while a second-order notch filter at 50 Hz minimizes power-line interference. Following this, segmentation of the signals is performed to isolate sEMG recordings corresponding to seven hand gestures per cycle, as depicted in the experimental design. Segments are obtained by identifying 4-second steady-state periods during maximal muscle contraction, effectively bypassing transitional delays. Subsequently, signal windowing is applied using rectangular sliding windows of lengths 250, 500, and 1000 ms, with corresponding shifts of 50, 100, and 200 ms, respectively, and overlapping windows.

3.2.3. Time-Frequency Analysis for Image-Based sEMG Analysis

Time-Frequency Analysis (TFA) methods play a crucial role in transforming 1D sEMG signals into 2D images, facilitating image-based analysis. By operating in the joint time-frequency (TF) domain, TFA approaches offer a comprehensive representation of signal characteristics, capturing signal energy distribution across time and frequency. This conversion process enhances the effectiveness of signal analysis, particularly for classification studies, where TF images serve as valuable inputs for DL models. Recent advancements in the field have witnessed the widespread adoption of TF images in conjunction with DL models, leading to notable outcomes in various studies (Yao Li et al., 2024; Ozdemir, Cura, and Akan, 2021). Among the traditional TFA methods, STFT and CWT stand out for their effectiveness in generating TF images. These images, derived through STFT and CWT, are subsequently employed to fine-tune pre-trained various ViT models for the classification of hand gestures. In the following sections, we provide a concise overview of STFT and CWT, elucidating their roles in the transformation process.

3.2.3.1. Short-Time Fourier Transform (STFT) for sEMG Image Generation

Short-Time Fourier Transform (STFT) is a powerful tool for analyzing non-stationary signals, such as sEMG signals, by providing a time-frequency representation. Mathematically, the STFT is defined as:

$$X(t,\omega) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j\omega\tau}\, d\tau \tag{3.1}$$

where:

- $x(\tau)$ is the signal to be analyzed,
- $w(\tau - t)$ is the window function centered at time t,
- ω is the angular frequency,
- j is the imaginary unit ($\sqrt{-1}$),
- τ is the integration variable.

STFT provides a way to examine the spectral content of the signal as it changes over time by applying the Fourier transform to a sliding window of the signal. This method allows for the decomposition of a signal into its constituent frequencies at different points in time, offering insights into the time-varying characteristics of the signal. In practice, the STFT is computed by dividing the signal into overlapping segments, applying a window function, and performing the Fourier transform on each segment. This results in a two-dimensional time-frequency representation, useful for analyzing non-stationary signals like sEMG. By visualizing the STFT, one can observe the evolution of different frequency components, aiding in the interpretation of underlying physiological processes.

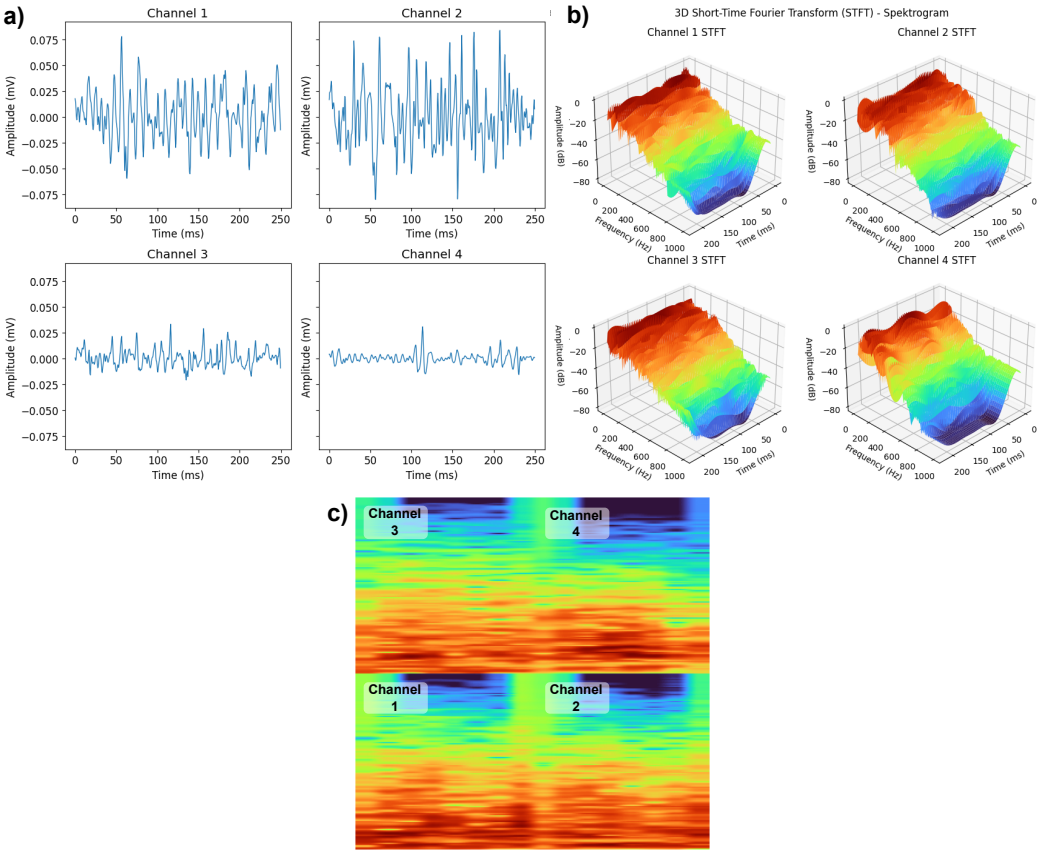

Figure 3.3. Analysis of sEMG signal recorded from four channels. a) Time-domain representation of sEMG signals recorded from four channels, showing the amplitude variations over time for each channel. b) 3D STFT spectrograms for each channel, illustrating the frequency content of the signals over time. c) Combined spectrogram view for all channels, presenting a comprehensive visualization of the frequency distribution and intensity over time. This data represents an sEMG signal captured over a 250 ms period during a resting state.

3.2.3.2. Continuous Wavelet Transform (CWT) for sEMG Image Generation

In this study, we utilized the Continuous Wavelet Transform (CWT) to generate sEMG images. CWT is a powerful tool for analyzing non-stationary signals like sEMG because it provides a time-frequency representation of the signal. Unlike the STFT, which uses a fixed window size, CWT can vary its window size to provide better frequency and time localization. This adaptability makes CWT particularly suitable for sEMG signal analysis, where the signal characteristics can vary rapidly over time.

To identify the most suitable wavelet parameters, eight different wavelets were tested using the `pywt.cwt` function. The wavelets tested included Bump (bump), Derivative of Gaussian (cgau1), Complex Morlet (cmor), Frequency B-Spline (fbsp), Gaussian (gaus1), Mexican Hat (mexh), Morlet (morl), and Shannon (shan). Each wavelet was applied to the sEMG data to create time-frequency representations, which were then fed into a convolutional neural network (CNN) for classification. The performance of each wavelet was evaluated based on accuracy, precision, recall, and F1-score. The results are summarized in Table 3.1.

Table 3.1. Performance metrics for various wavelets

Wavelet	**Accuracy**	**Precision**	**Recall**	**F1-score**
bump	0.77	0.76	0.77	0.76
cgau1	0.78	0.79	0.78	0.77
cmor	0.79	0.79	0.79	0.79
fbsp	0.77	0.78	0.77	0.77
gaus1	0.66	0.65	0.66	0.65
mexh	0.67	0.68	0.67	0.67
morl	0.70	0.72	0.7	0.7
shan	0.78	0.79	0.78	0.78

Among the wavelets tested, the Complex Morlet wavelet (cmor) achieved the highest performance across all metrics. This wavelet's superior performance can be attributed to its ability to provide a good balance between time and frequency localization, which is crucial for accurately capturing the features of sEMG signals. The Complex

Morlet wavelet used in this study is mathematically defined as:

$$\psi(t) = \frac{1}{\sqrt{\pi B}} \exp\left(-\frac{t^2}{B}\right) \exp\left(j2\pi Ct\right) \tag{3.2}$$

where B represents the bandwidth and C denotes the center frequency. The combination of these parameters allows the wavelet to effectively analyze the frequency content of the sEMG signals over time.

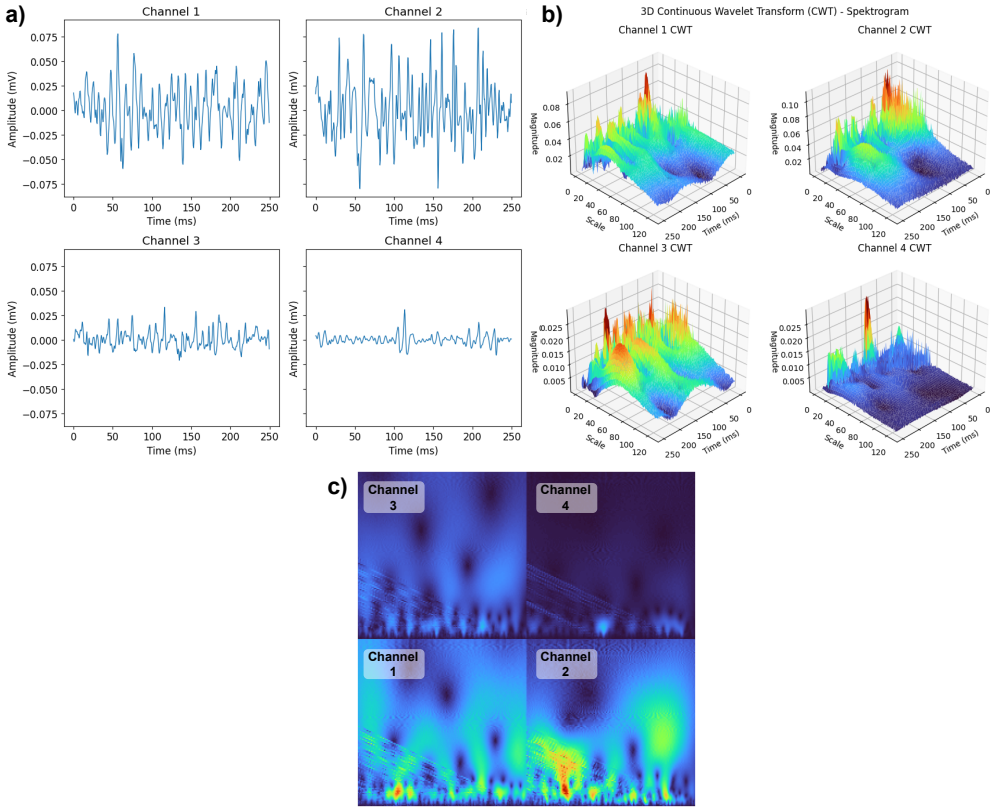

Figure 3.4. Examination of sEMG signals from four channels. (a) Displays the time-domain representation of sEMG signals from four channels, highlighting the amplitude changes over time for each channel. (b) Shows the 3D CWT spectrograms for each channel, depicting the frequency content of the signals over time. (c) Presents a combined CWT spectrogram view for all channels, offering a detailed visualization of the frequency distribution and intensity over time. This data captures an sEMG signal over a 250 ms period during a resting state.

The CNN architecture used for classification included three convolutional layers,

each followed by batch normalization and max-pooling layers to reduce the spatial dimensions. This was followed by fully connected layers with dropout regularization to mitigate overfitting, and a final SoftMax layer to produce the output probabilities. This detailed architecture ensured that the model could effectively learn and generalize from the CWT-transformed sEMG images.

In this study, it was found that the Complex Morlet wavelet yielded the highest accuracy in a simple Convolutional Neural Network (CNN) model for the CWT of this dataset. Using this parameter, CWT was applied to generate images across the entire dataset, which were then prepared for use in the respective transformers for CWT applications.

3.2.3.3. Application of ViT-Based Models to sEMG Image Data

We investigated the efficacy of various ViT models in classifying sEMG signals transformed into image representations using STFT and CWT with various window sizes and overlap ratios. Specifically, we explore the performance of three state-of-the-art transformer architectures: Base Vision Transformers (ViT), Shifted Window Transformer (Swin), and Multi-Axis Vision Transformer (MaxViT).

The sEMG image dataset is split into training, testing, and validation sets with a ratio of 70% for training, 15% for testing, and 15% for validation, ensuring robust evaluation of the models. Each sEMG image is associated with a hand gesture label, allowing the models to learn the mapping between image features and corresponding gestures. We applied both fine-tuning (FT) and feature extraction (FE) techniques to all three transformer models to assess their capabilities in different training scenarios.

By evaluating the performance of Vision Transformers, Swin Transformer, and MaxViT on the sEMG image dataset, we aim to identify the most effective architecture for hand gesture recognition tasks. Through comprehensive experimentation and analysis, we gain insights into the suitability of various ViT models for processing sEMG data.

3.3. Results and Discussion

In this section, we present the findings of our study on hand gesture recognition using the dataset of 4-channel surface sEMG signals provided by Ozdemir et al. (2022). The dataset consists of recordings from 40 participants performing 10 distinct hand gestures, providing a robust foundation for developing and evaluating various ViT models.

In our preprocessing stage, we segmented the sEMG signals using windowing techniques with window lengths of 250 ms, 500 ms, and 1000 ms. Each window was overlapped with the next data segment by shifts of 50 ms for the 250 ms window, 100 ms for the 500 ms window, and 200 ms for the 1000 ms window.

Subsequently, we investigated the effectiveness of transforming these segmented signals into time-frequency images using STFT and CWT. These time-frequency representations served as inputs to various ViT architectures, including base ViT, Swin, and MaxViT.

According to the results shown in Figure 3.5, different training methods significantly impact model performance. The images used in this study were generated using STFT. The FT method yields substantially superior performance compared to the FE method. Models subjected to FT, including MaxViT, Swin, and base ViT, achieved nearly 100% accuracy in both training and validation phases and exhibited low loss values. Notably, there was a rapid decrease in loss within the first 10 epochs, followed by stabilization. In contrast, when using the FE method, accuracy rates were markedly lower. The MaxViT model demonstrated lower accuracy compared to other models in the FE approach, while the ViT model achieved the highest accuracy. Training losses started higher and decreased more slowly compared to the FT method. In conclusion, the FT method enhances overall model performance, yielding higher accuracy rates and lower losses. This indicates that training the entire model is more effective than relying solely on pre-learned features. Within this context, the base ViT model exhibited the best performance in the FE approach, whereas the MaxViT model showed the lowest performance. These findings provide valuable insights into model selection and training strategies. Figures 3.6 and 3.7 illustrate the impact of different window lengths used in STFT on the images generated for model training. The performance of the three models varies with these STFT configurations. These figures provide a comparative analysis of how different window lengths in the STFT process influence the accuracy and loss metrics across the three models.

Figure 3.8 illustrates the accuracy and loss curves for CWT images with a 500 ms window, comparing two different scenarios: FT and FE. In the FT scenario, the accuracy curves for the MaxViT, Swin, and base ViT models show a similar upward trend. The MaxViT model achieves the highest validation accuracy, surpassing 97% after approximately the 10th epoch. The Swin and base ViT models also reach high accuracy levels, but slightly lower than MaxViT. The loss curves indicate that the training and validation losses for all three models decrease rapidly. The MaxViT model achieves the lowest validation loss, indicating the best performance, while the Swin and base ViT models also reach low validation loss levels, but slightly higher than MaxViT. In the FE scenario, the accuracy curves for the MaxViT, Swin, and base ViT models show an upward

trend, but reach lower accuracy levels compared to the FT scenario. The base ViT model achieves the highest validation accuracy, reaching around 0.70 after approximately the 60th epoch. The loss curves show that the training and validation losses for all three models decrease but remain higher compared to the FT scenario. The base ViT model achieves the lowest validation loss, indicating the best performance, while the Swin and MaxViT models also reach low loss levels, but slightly higher than MaxViT.

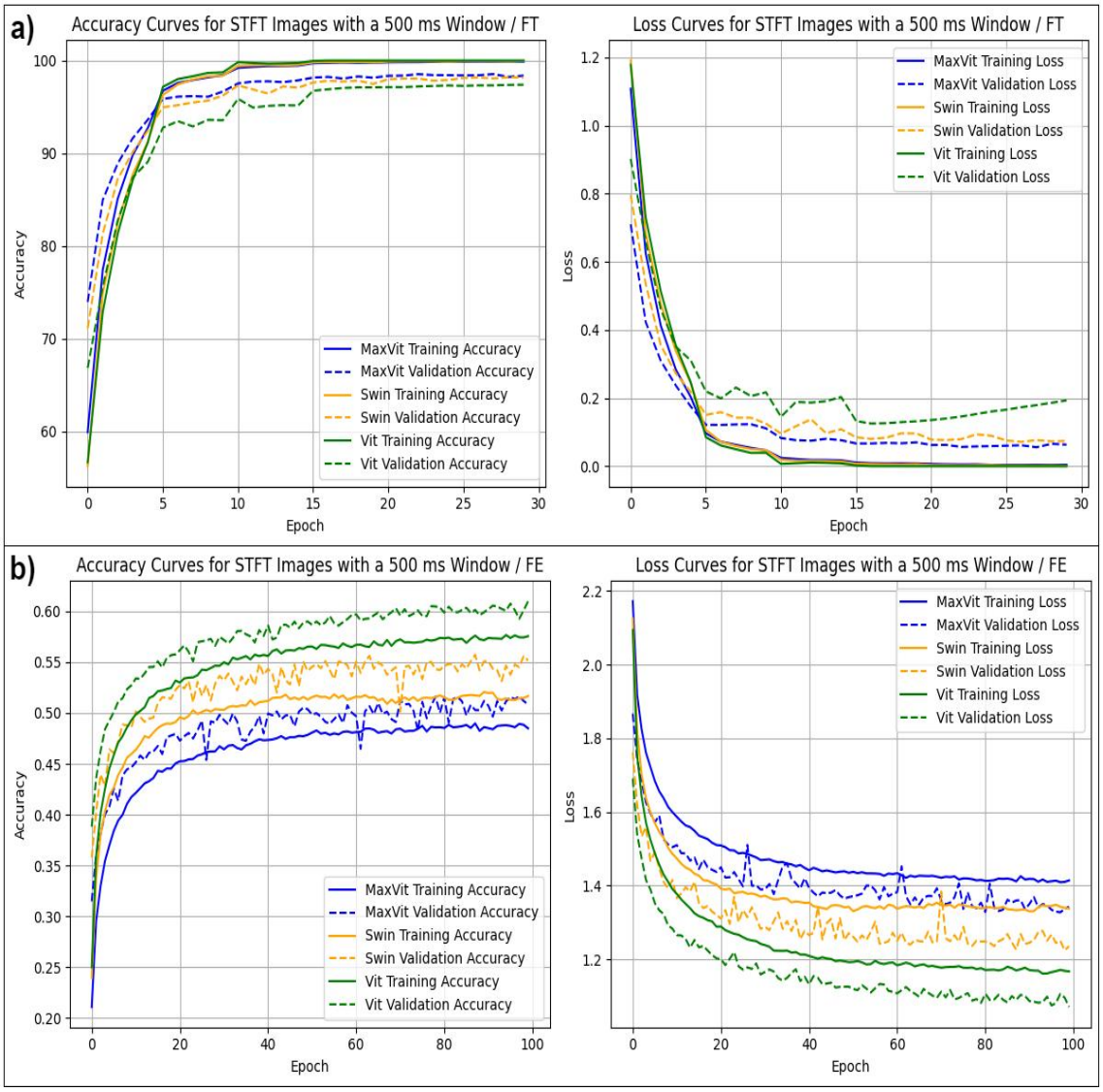

Figure 3.5. Comparison of training and validation performance for MaxViT, Swin, and base ViT models using STFT images created with a window size of 500 ms a) Shows the accuracy and loss graphs obtained with FT for all three models.b) Shows the graphs obtained when all three models are used as feature extractors.

Figures 3.9 and 3.10 depict the effect of varying window lengths used in CWT on

the images generated for model training. The performance of the three models changes with these CWT configurations. These figures offer a comparative analysis of how different window lengths in the CWT process impact the accuracy and loss metrics across the three models.

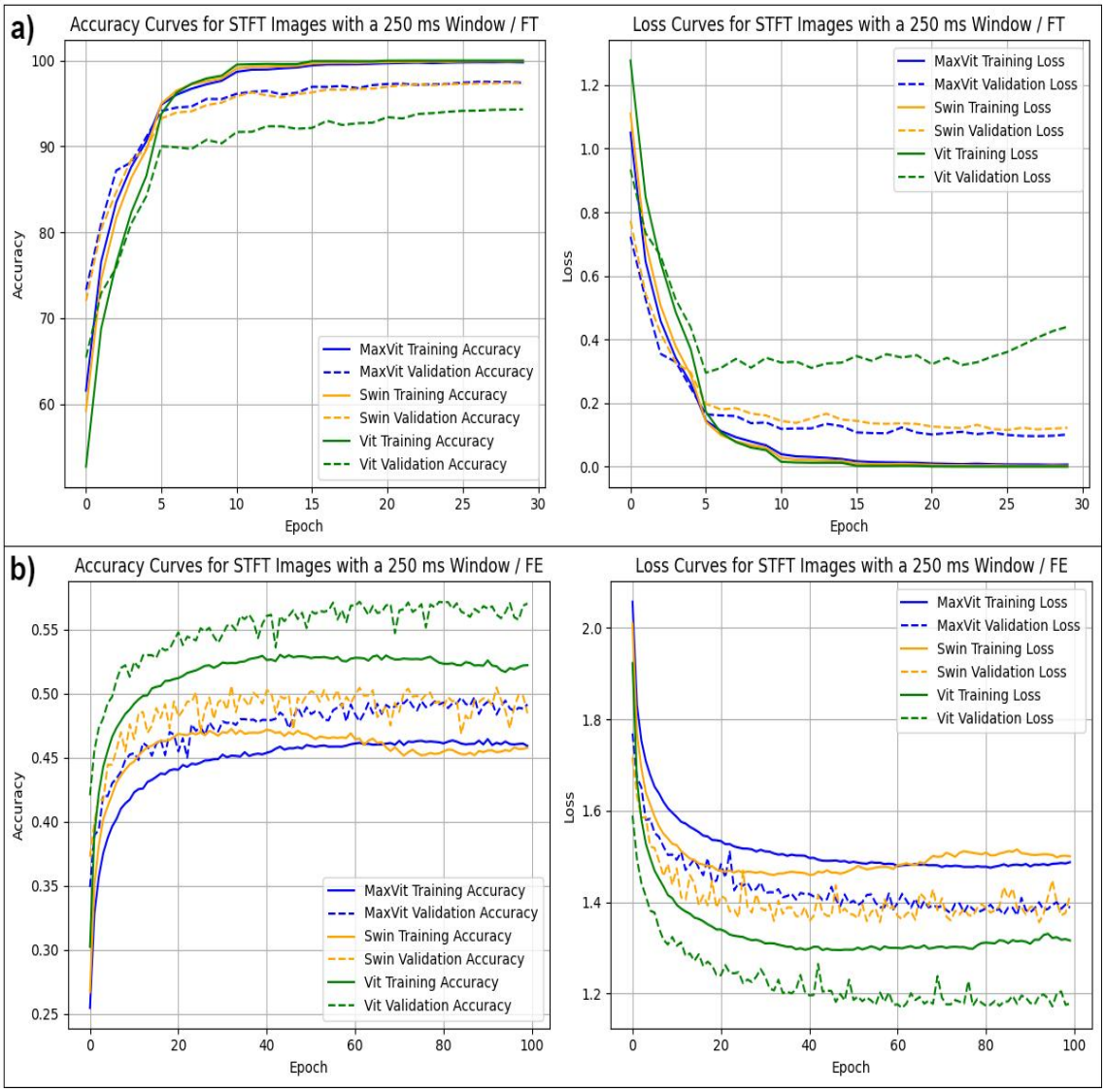

Figure 3.6. Comparison of training and validation performance for MaxViT, Swin, and base ViT models using STFT images created with a window size of 250 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

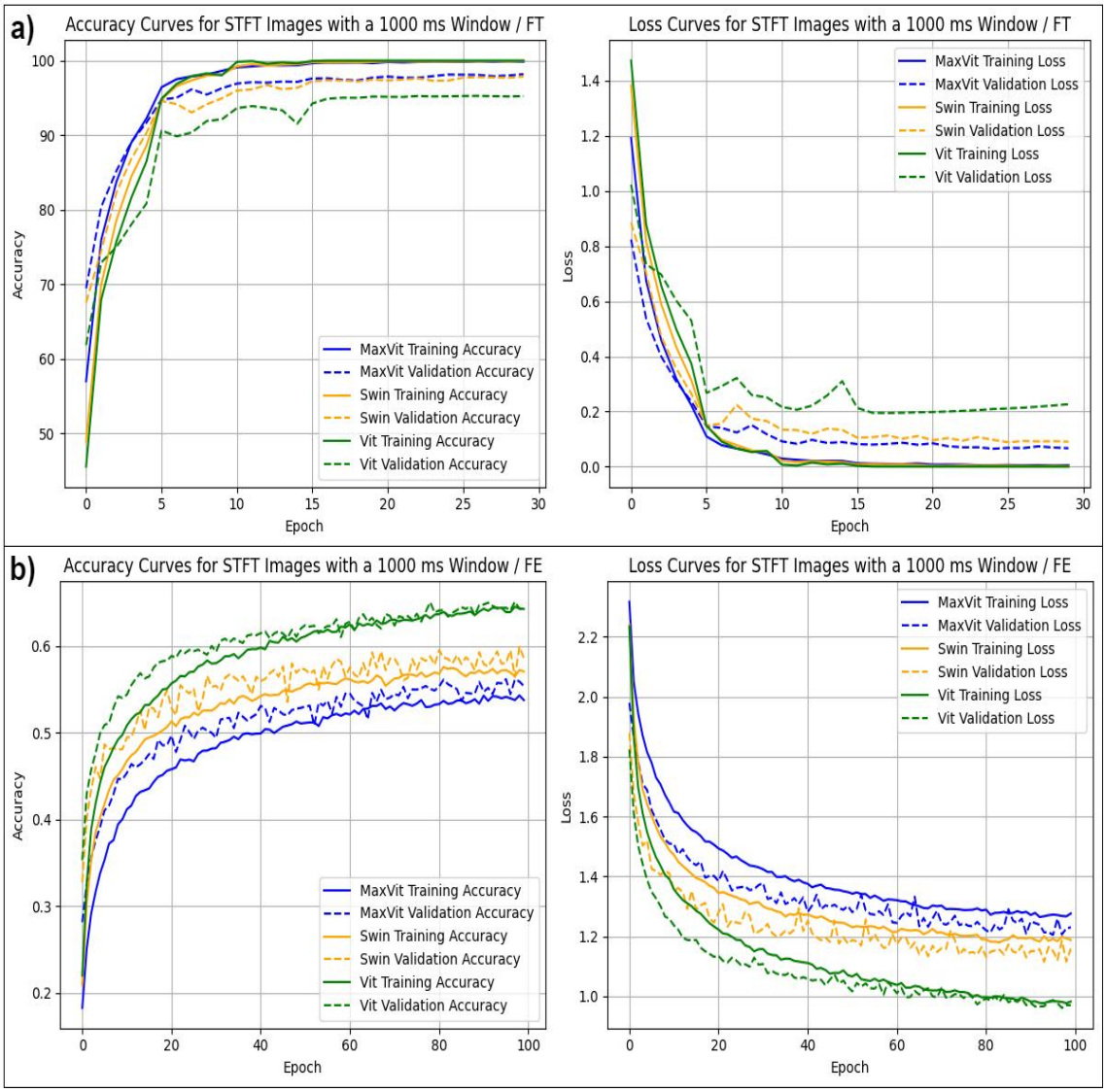

Figure 3.7. Comparison of training and validation performance for MaxViT, Swin, and base ViT models using STFT images created with a window size of 1000 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

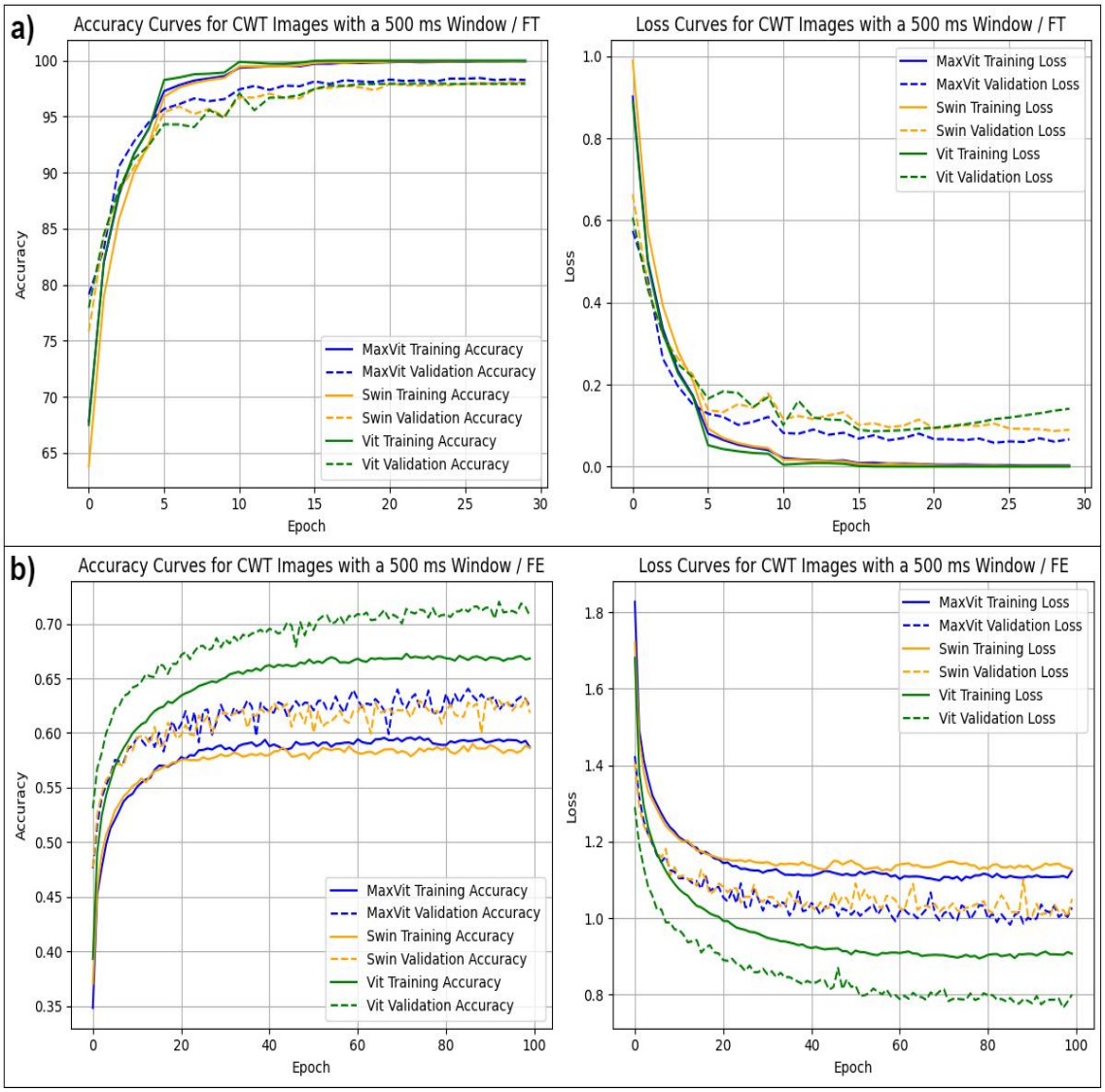

Figure 3.8. Comparison of training and validation performance for MaxViT, Swin, and base ViT models using CWT images created with a window size of 500 ms a) Shows the accuracy and loss graphs obtained with FT for all three models.b) Shows the graphs obtained when all three models are used as feature extractors.

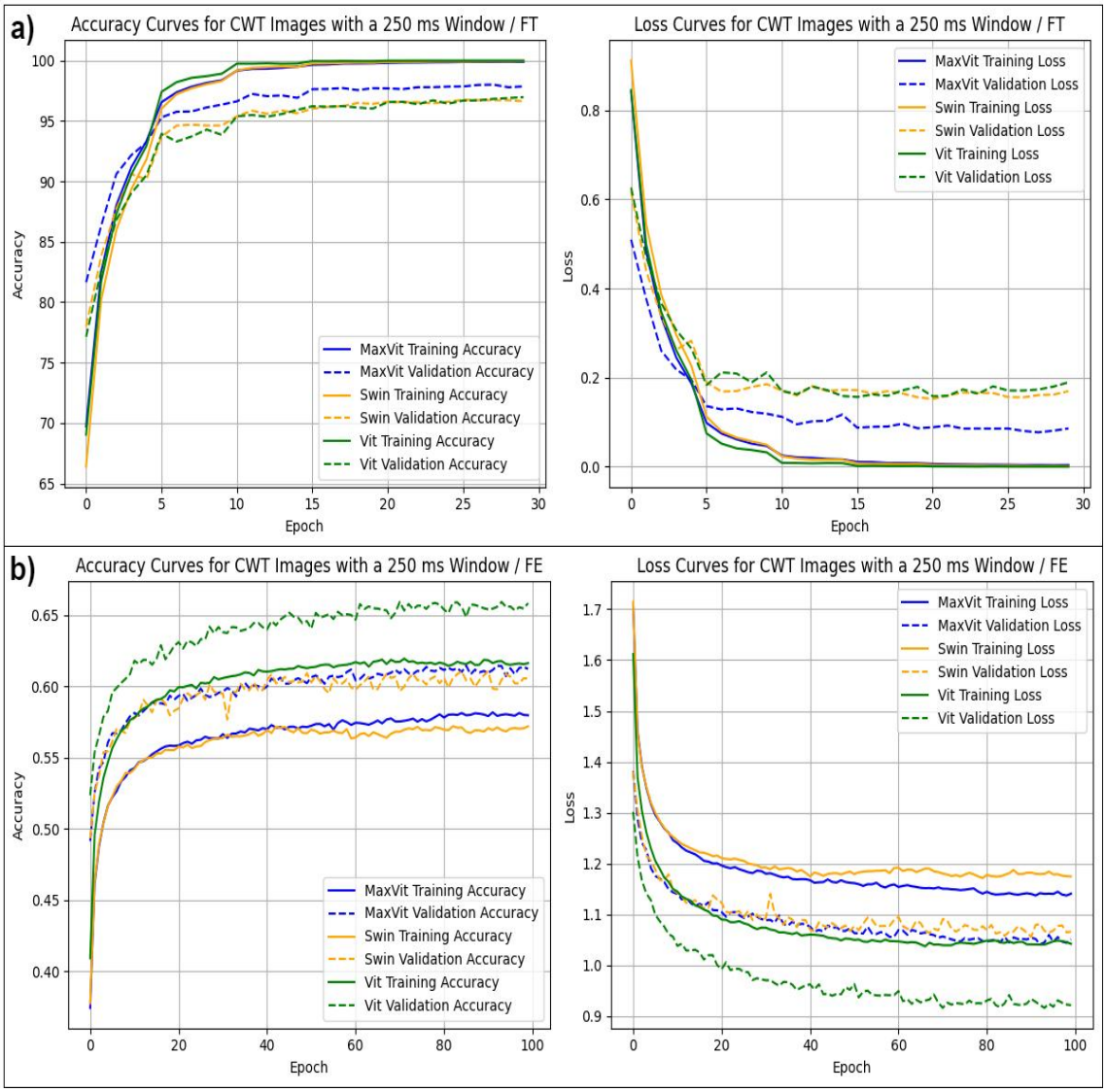

Figure 3.9. Comparison of training and validation performance for MaxViT, Swin, and base ViT models using CWT images created with a window size of 250 ms a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

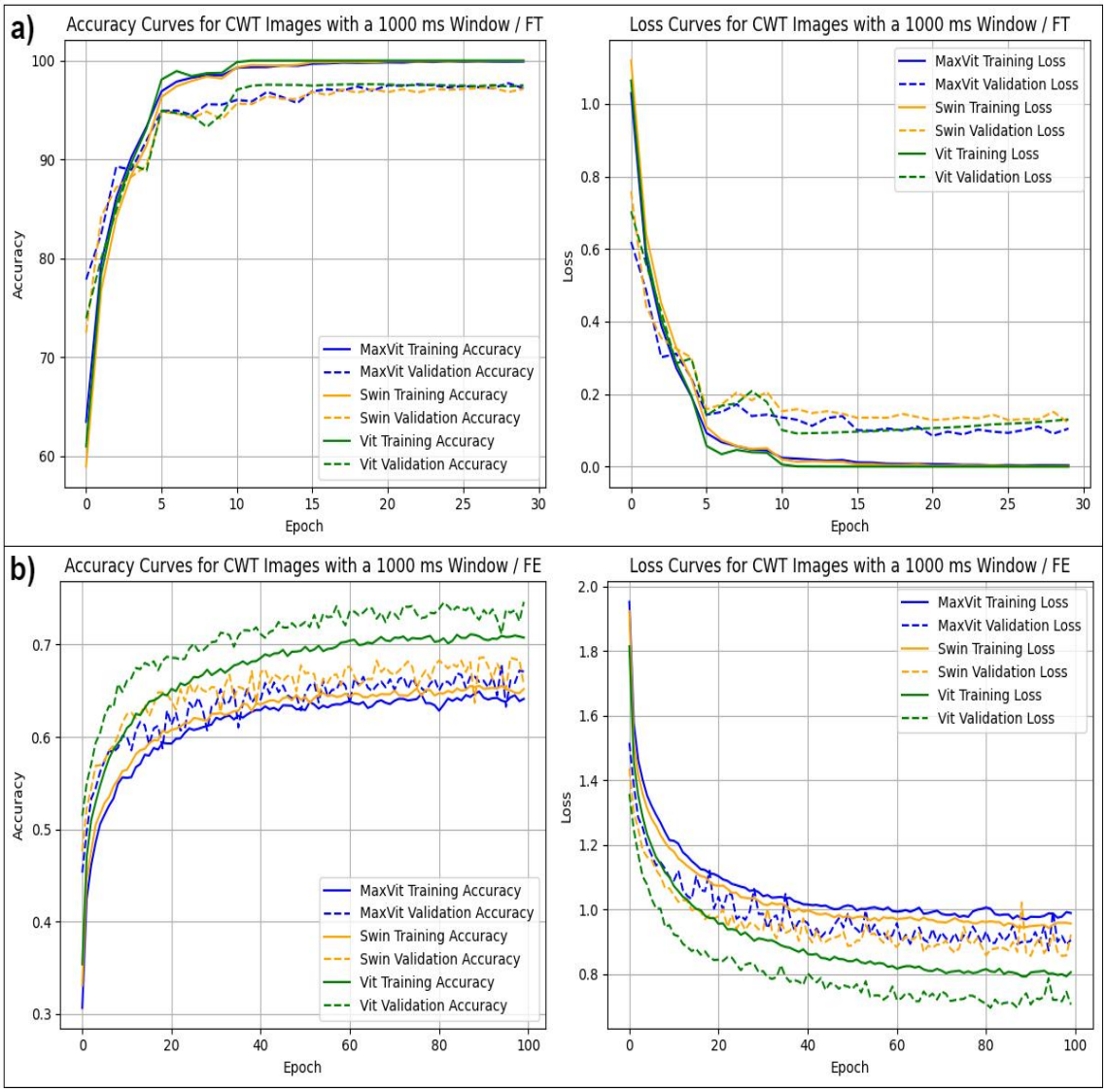

Figure 3.10. Comparison of training and validation performance for MaxViT, Swin, and base ViT models using CWT images created with a window size of 1000 ms a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

Figure 3.11 presents confusion matrices for STFT images using three different models (MavVit, Swin, and base ViT) with a 500 ms window. The top row (a, c, e) shows the models when fine-tuned, while the bottom row (b, d, f) shows the models used as feature extractors. In the fine-tuned configuration, all three models exhibit strong diagonal dominance, indicating high classification accuracy. MavVit (a) demonstrates the highest accuracy, achieving near-perfect classification for several classes, with minimal misclassifications. Swin (c) also performs well, showing high accuracy with few misclassifications. Base ViT (e), while still accurate, shows more misclassifications compared to MavVit and Swin.

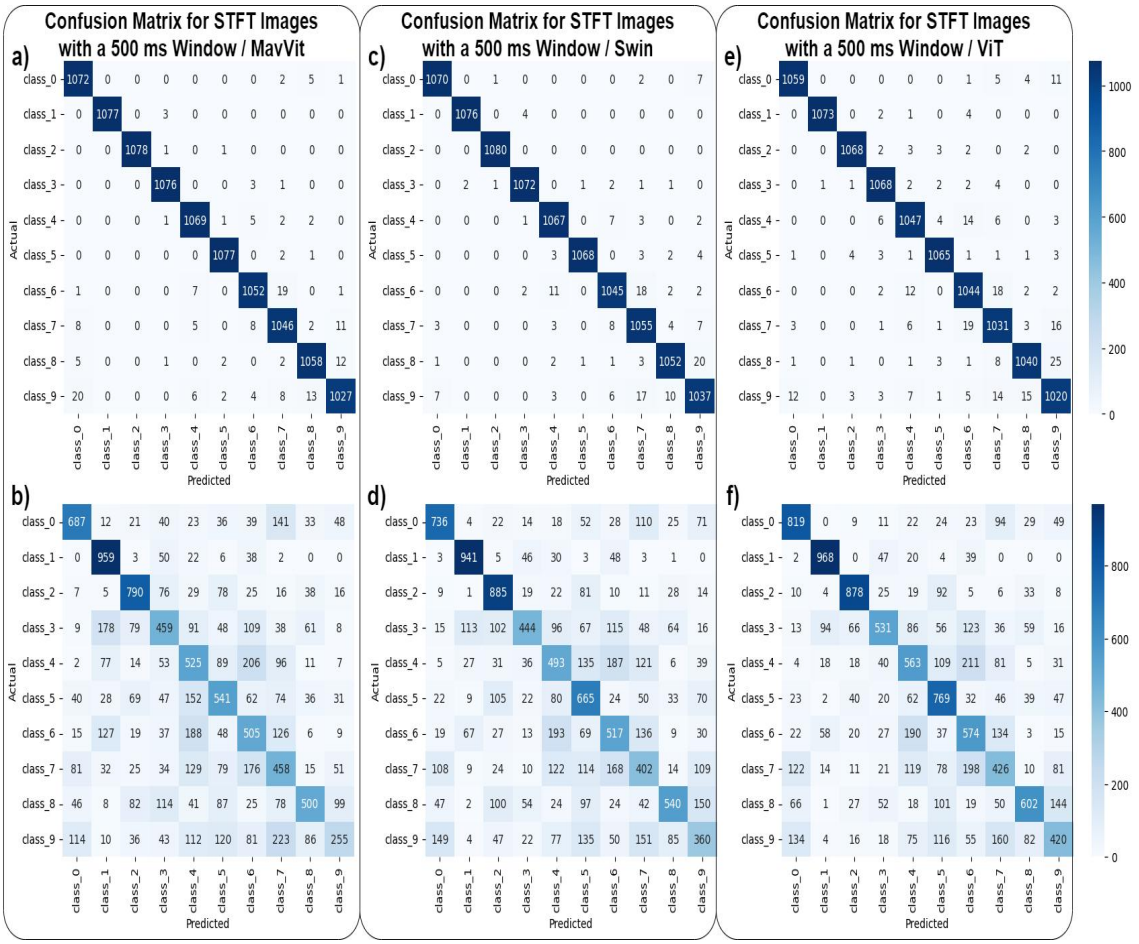

Figure 3.11. Comparison of classification performance via confusion matrices for MaxViT, Swin, and base ViT models using STFT images created with a window size of 500 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

When used as feature extractors, the models show a more distributed pattern of misclassifications, indicating lower accuracy. MavVit (b) and Swin (d) still perform relatively well but have higher off-diagonal values, reflecting more instances of misclassification across various classes. Base ViT (f) outperforms the other models in this configuration, showing relatively better accuracy and fewer misclassifications compared to MavVit and Swin. Base ViT (f) demonstrates substantial misclassifications in several classes but still maintains a better overall performance as a feature extractor. Overall, the results suggest that fine-tuning the models significantly enhances their classification accuracy for STFT images, with MavVit demonstrating the best performance in the fine-tuned configuration. In contrast, base ViT performs better as a feature extractor, highlighting the importance of model configuration for optimal performance.

Figures 3.12 and 3.13 present confusion matrices that illustrate the impact of

varying window lengths in the STFT on the generated images used for model testing. These figures provide a comparative analysis of how different window lengths in the STFT process affect the classification accuracy across the three models. The performance variations of the three models with these STFT configurations are clearly depicted, highlighting the influence of window length on model accuracy during testing.

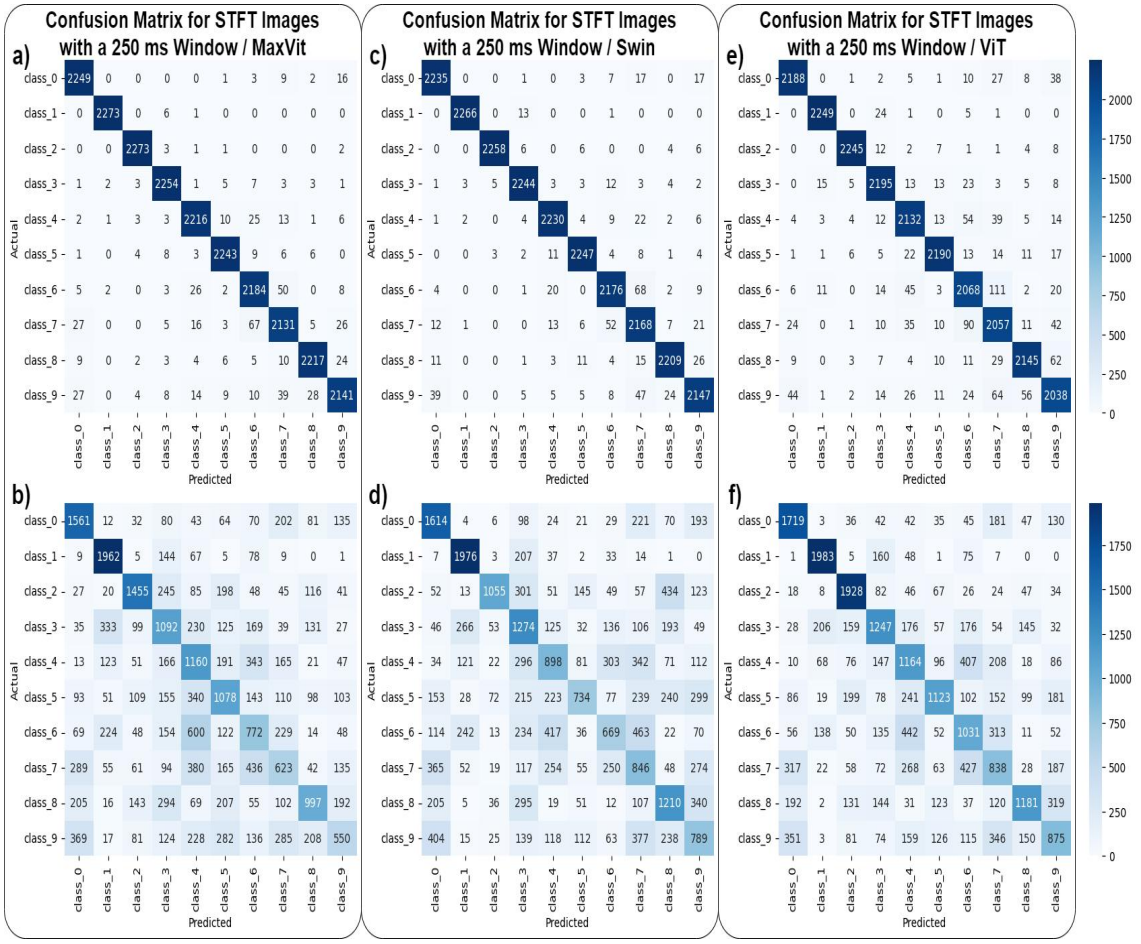

Figure 3.12. Comparison of classification performance via confusion matrices for MaxViT, Swin, and base ViT models using STFT images created with a window size of 250 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

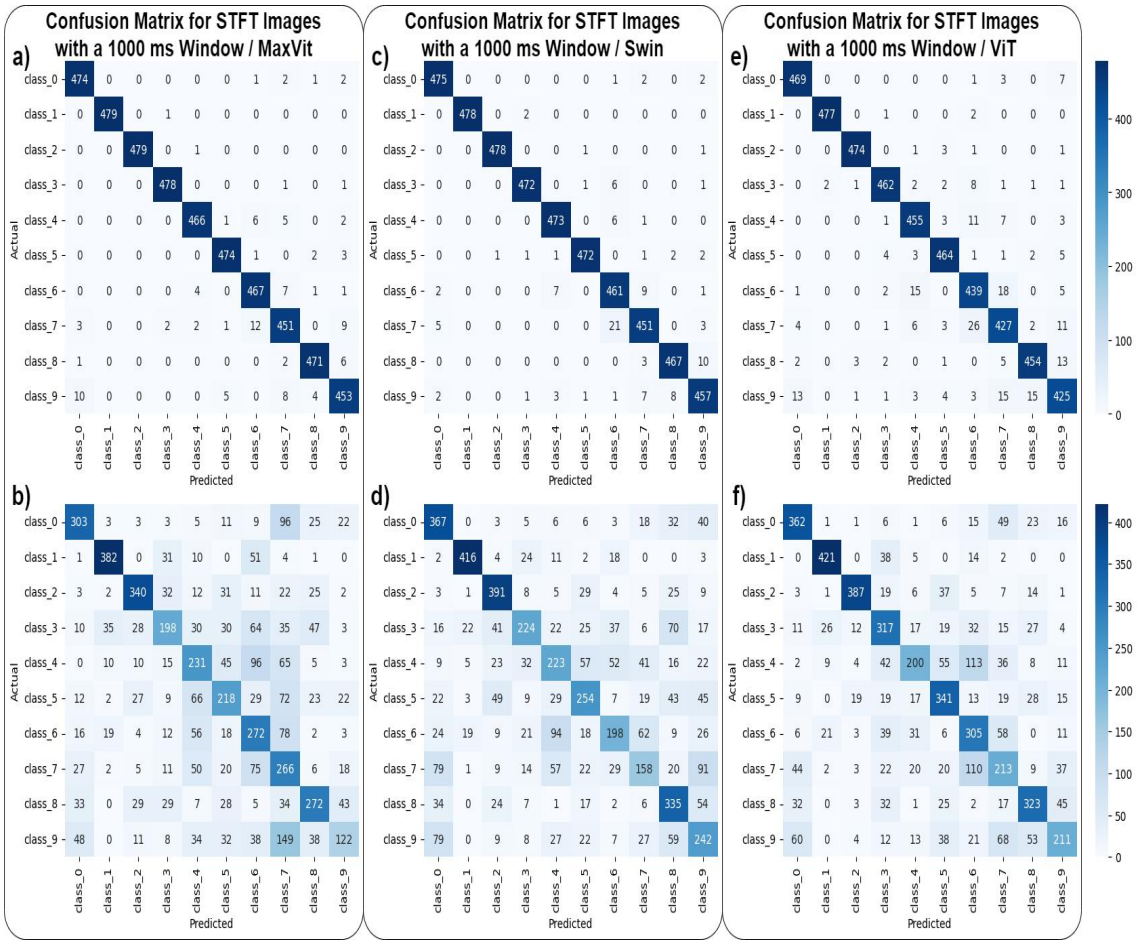

Figure 3.13. Comparison of classification performance via confusion matrices for MaxViT, Swin, and base ViT models using STFT images created with a window size of 1000 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

Figure 3.14 presents confusion matrices for CWT images using three different models with a 500 ms window. The top row (a, c, e) shows the models when fine-tuned, while the bottom row (b, d, f) shows the models used as feature extractors. In the fine-tuned configuration, MavVit demonstrates the highest classification accuracy, achieving near-perfect classification for several classes, with minimal misclassifications. Swin also performs well, showing high accuracy with few misclassifications. ViT, while still accurate, shows more misclassifications compared to MavVit and Swin.

When used as feature extractors, the models show a more distributed pattern of misclassifications, indicating lower accuracy. MavVit and Swin still perform relatively well but have higher off-diagonal values, reflecting more instances of misclassification across various classes. ViT outperforms the other models in this configuration, showing relatively better accuracy and fewer misclassifications compared to MavVit and Swin.

Overall, the results suggest that fine-tuning the models significantly enhances their classification accuracy for CWT images, with MavVit demonstrating the best performance in the fine-tuned configuration. In contrast, ViT performs better as a feature extractor, highlighting the importance of model configuration for optimal performance.

Figures 3.14 and 3.15 display confusion matrices that demonstrate the effect of varying window lengths in the CWT on the generated images used for model testing. These figures provide a comparative analysis of how different window lengths in the CWT process impact the classification accuracy across the three models. The performance variations of the three models with these CWT configurations are clearly illustrated, highlighting the influence of window length on model accuracy during testing.

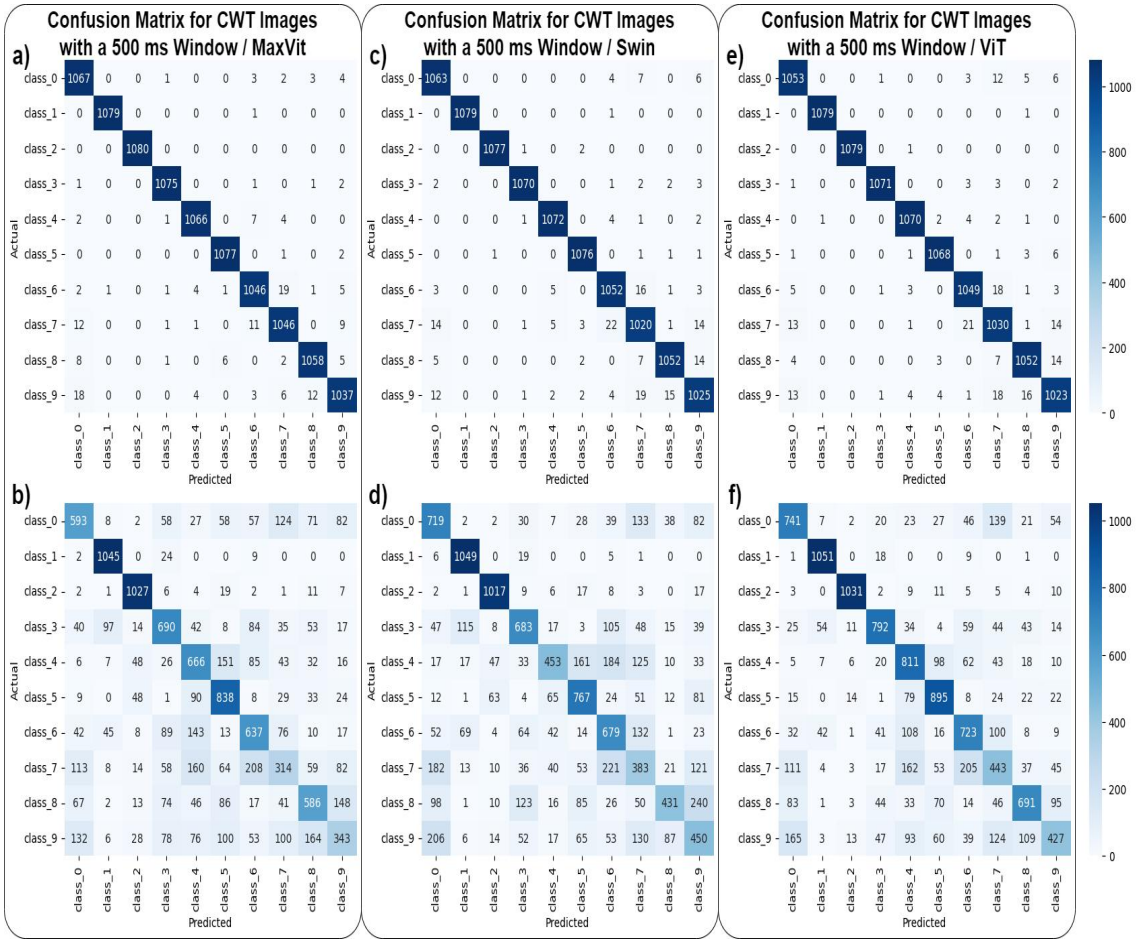

Figure 3.14. Comparison of classification performance via confusion matrices for MaxViT, Swin, and base ViT models using CWT images created with a window size of 500 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

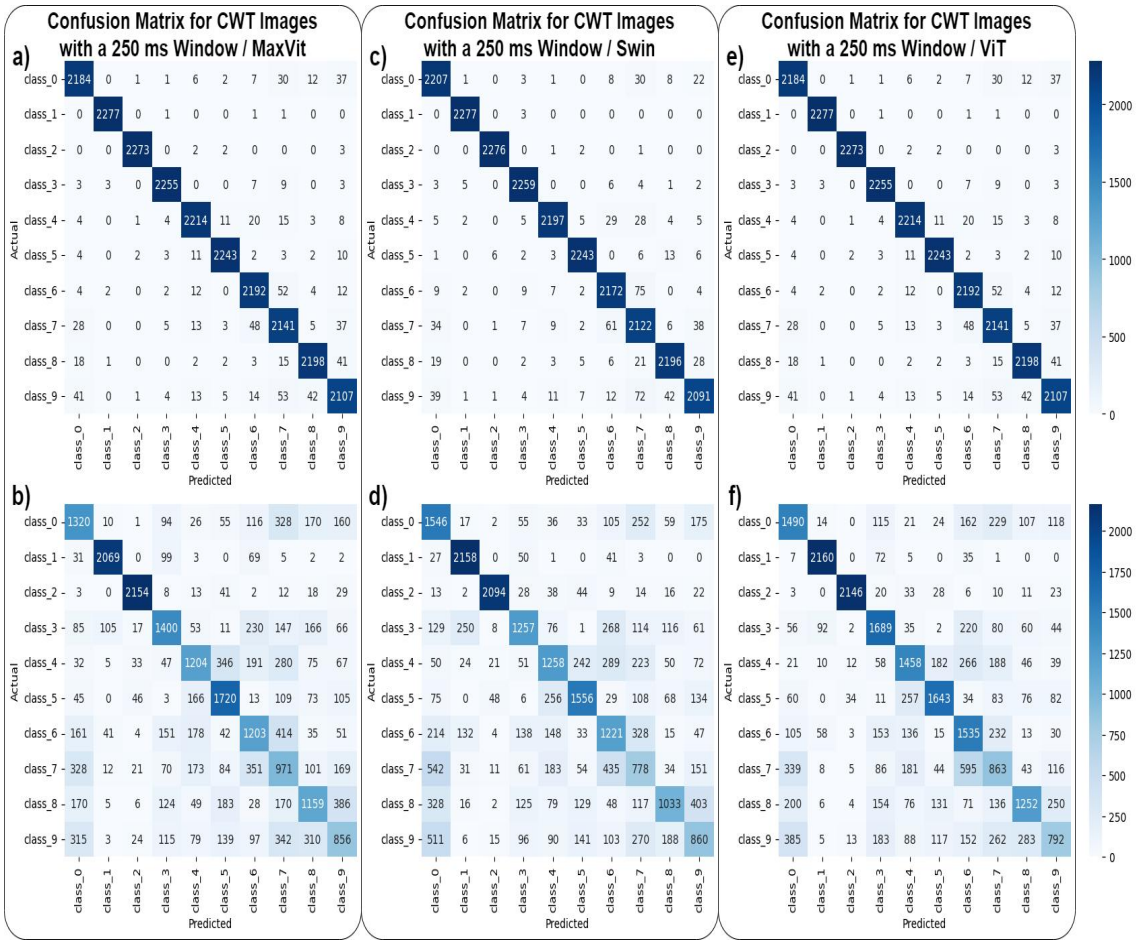

Figure 3.15. Comparison of classification performance via confusion matrices for MaxViT, Swin, and base ViT models using CWT images created with a window size of 250 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

In Table 3.2, the performance metrics of three models (MaxViT, Swin, and ViT) are evaluated using two methods (CWT and STFT) with varying window lengths (250 ms, 500 ms, and 1000 ms). The evaluation is conducted in two stages: Feature Extractor and FT, with metrics including Accuracy, Precision, Recall, and F1 Score.

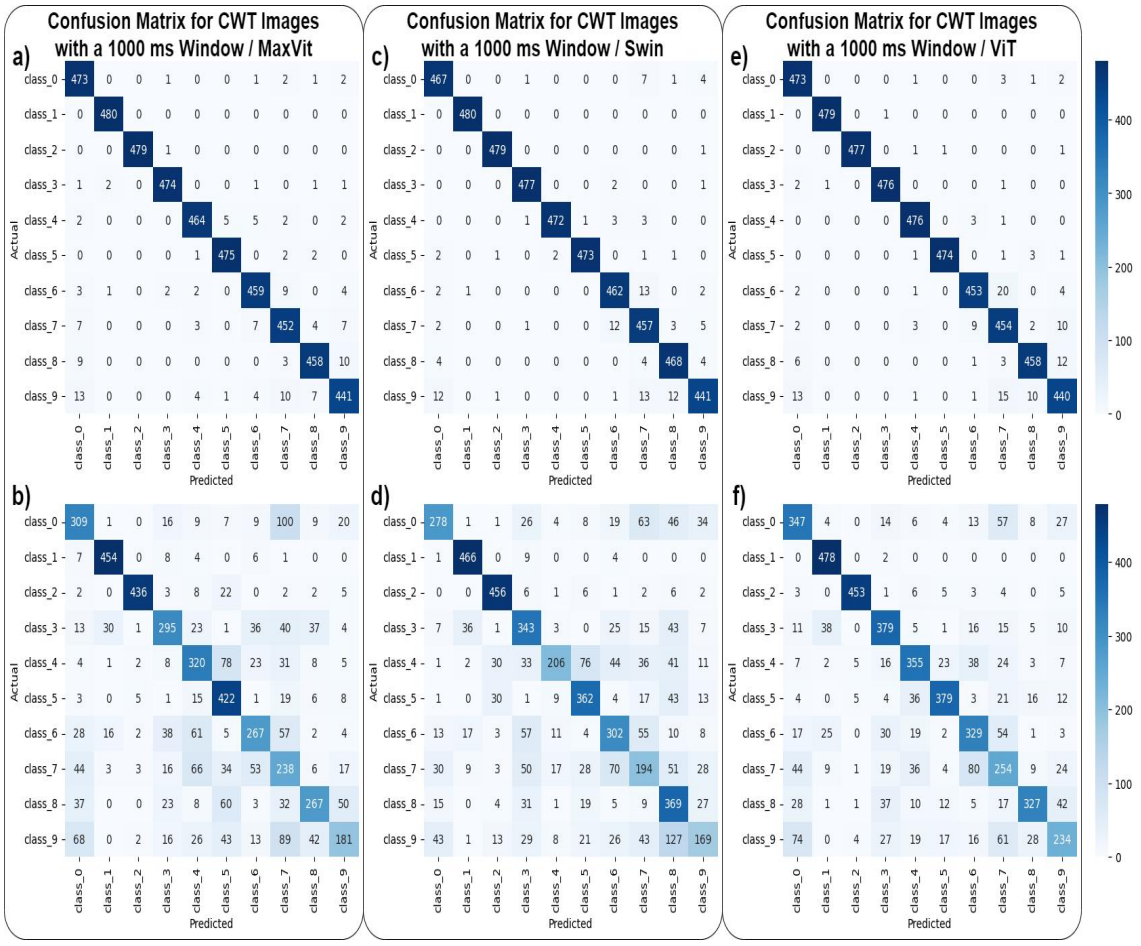

Figure 3.16. Comparison of classification performance via confusion matrices for MaxViT, Swin, and base ViT models using CWT images created with a window size of 1000 ms. a) Shows the performance metrics for FT, with MaxViT, Swin, and base ViT models. b) Displays the performance metrics for FE, with MaxViT, Swin, and base ViT models.

For the CWT method, the ViT model consistently outperformed the MaxViT and Swin models across all window lengths in the Feature Extractor stage. Specifically, the ViT model achieved the highest accuracy of 73.65% with a window length of 1000 ms, compared to MaxViT's 66.44% and Swin's 65.52%. Precision, recall, and F1 scores followed a similar trend, with ViT showing superior performance. In the FT stage, all models exhibited significant improvements in performance metrics. The MaxViT model achieved an accuracy of 98.40% with a window length of 500 ms, while Swin and ViT models also showed high accuracies of 98.40% and 97.90%, respectively, for the same window length. The precision, recall, and F1 scores were consistently high across all models, indicating effective FT.

Under the STFT method, the ViT model again demonstrated superior performance in the Feature Extractor stage, particularly with a window length of 1000 ms, achieving an

accuracy of 64.17%. The MaxViT and Swin models showed lower accuracies of 54.25% and 58.50%, respectively, for the same window length. Precision, recall, and F1 scores were also higher for the ViT model. In the FT stage, the performance metrics for all models improved significantly. The MaxViT model achieved the highest accuracy of 97.80% with a window length of 500 ms, while the Swin and ViT models showed accuracies of 97.60% and 94.70%, respectively, for the same window length. Precision, recall, and F1 scores were similarly high, indicating that FT effectively enhanced the models' performance.

Overall, the ViT model demonstrated superior performance in the Feature Extractor stage across both methods and all window lengths. The FT stage resulted in substantial improvements in performance metrics for all models, with accuracies, precision, recall, and F1 scores all being very high (mostly above 94%). Increasing the window length generally improved the performance metrics in the FE stage, but the improvement was less pronounced in the FT stage. These results suggest that while the initial FE benefits from larger window lengths, FT is crucial for achieving high performance across all models and methods.

Table 3.2. Performance metrics (Accuracy, Precision, Recall, and F1 Score) of MaxViT, Swin, and Vit models using CWT and STF methods with varying window lengths (250, 500, and 1000) in both FE and FT stages.

Method	**Model**	**Window Length**	**Feature Extraction**				**Fine Tuning**			
			Accuracy	*Precision*	*Recall*	*F1 Score*	*Accuracy*	*Precision*	*Recall*	*F1 Score*
CWT	MaxViT	250	61.65%	61.97%	61.65%	61.64%	96.90%	96.90%	96.90%	96.90%
		500	62.40%	60.86%	62.40%	61.19%	98.40%	98.40%	98.40%	98.40%
		1000	66.44%	67.43%	66.44%	66.18%	97.00%	97.00%	97.00%	97.00%
	Swin	250	60.36%	61.02%	60.36%	60.12%	96.70%	96.70%	96.70%	96.70%
		500	61.40%	61.83%	61.40%	60.65%	98.00%	98.00%	98.00%	98.00%
		1000	65.52%	66.21%	65.52%	64.58%	97.40%	97.40%	97.40%	97.40%
	Vit	250	65.91%	66.02%	65.91%	65.52%	96.90%	96.90%	96.90%	96.90%
		500	70.42%	70.16%	70.42%	69.83%	97.90%	97.90%	97.90%	97.90%
		1000	73.65%	73.80%	73.65%	73.44%	97.10%	97.10%	97.10%	97.10%
STF	MaxViT	250	49.34%	49.15%	49.34%	48.55%	97.30%	97.30%	97.30%	97.30%
		500	52.58%	52.96%	52.58%	51.92%	98.40%	98.40%	98.40%	98.40%
		1000	54.25%	56.54%	54.25%	54.38%	97.80%	97.80%	97.70%	97.70%
	Swin	250	48.53%	50.16%	48.53%	48.05%	97.30%	97.30%	97.30%	97.30%
		500	55.40%	55.72%	55.40%	54.89%	98.40%	98.40%	98.40%	98.40%
		1000	58.50%	58.40%	58.50%	57.81%	97.60%	97.60%	97.60%	97.60%
	Vit	250	57.41%	57.38%	57.41%	56.97%	94.30%	94.40%	94.30%	94.30%
		500	60.65%	60.98%	60.65%	60.33%	97.40%	97.40%	97.40%	97.40%
		1000	64.17%	64.86%	64.17%	63.95%	94.70%	94.70%	94.70%	94.70%

3.4. Conclusion

In this study, we explored the integration of sEMG with advanced TFA and various ViT models for hand gesture classification. Utilizing the open-source dataset, we aimed to enhance the performance of sEMG-based gesture recognition systems by combining TFA methods such as STFT and CWT with various ViT models, including base Vision Transformers (ViT), Swin Transformer, and MaxViT.

Our findings demonstrate that FT various ViT models significantly improves classification accuracy compared to FE methods. Among the models tested, MaxViT exhibited the highest performance when fine-tuned, achieving near-perfect accuracy in both training and validation phases. Specifically, MaxViT achieved an accuracy of 98.40% with a 500 ms window length for both STFT and CWT methods. The base ViT model, while performing well in the FE approach, showed lower accuracy compared to fine-tuned models, with the highest accuracy of 73.65% for CWT and 64.17% for STFT in the feature extraction stage.

The comparative analysis of STFT and CWT methods revealed that both techniques are effective in transforming sEMG signals into time-frequency images suitable for various ViT models. Overall, this study underscores the potential of combining sEMG with advanced TFA and various ViT models to develop intuitive and responsive control systems for many applications, including prosthetic devices and HCI systems. By leveraging the strengths of ViT and signal processing techniques, we believe there is significant potential for future advancements in intelligent human-machine interfaces, which could ultimately enhance the quality of life for users.

CHAPTER 4

PHASE-OTDR-BASED ELECTRICAL CURRENT SENSING USING DEEP LEARNING

4.1. Introduction

Fiber optic current sensors offer several advantages, including lightweight, small size, and good insulating ability, over traditional current sensing technologies such as electromagnetic sensors or Hall effect sensors, making them a popular choice for a wide range of applications, including generation and distribution of electricity, nuclear power plants, on-aircraft lightning-detection, and railway stray current detection (Aerssens et al., 2012; Kurosawa, 2014; Nguyen, Ely, and Szatkowski, 2015; S. Xu et al., 2014). Ampere's law enables contactless current sensing by measuring the magnetic field created around the current of interest. Optical fiber, as a magneto-optic material, responds to the magnetic field according to the Faraday effect. That response can be observed in the polarization properties of a lightwave traveling inside the fiber, which allows measuring the magnetic field. Fiber can almost uniformly be exposed to the current-induced magnetic field by being coiled around a current-carrying wire, so the current of interest can be quantified from magnetic field response.

Ideally, circularly polarized light passing through such a coil gets a Faraday phase shift that is directly proportional to the current of interest. This phase information can be acquired by interferometric detection methods such as Sagnac, Michelson, or reciprocal reflection (Frosio and Dändliker, 1994; Kersey and Jackson, 1986; Nicati and Robert, 1988) interferometer. However, the demodulated phase may contain a dominating nonlinear term coming from the intrinsic linear birefringence of the fiber. Spun fiber, a special type of fiber manufactured by spinning the preform while drawing, provides a linear birefringence suppression, so it is widely preferred for Faraday sensing applications. Optical fiber sensitivity to the Faraday effect is relatively poor due to the magneto-optic properties of silica (Cruz, Andres, and Hernandez, 1996). High-current sensing applications take advantage of it since the maximum phase shift is limited to π or 2π depending on the sensing configuration. However, sensing low-level currents becomes more challenging due to several noise sources, such as temperature drift and vibrations. Proper packaging and isolation of the sensing coil can reduce the external noise effect, but the photodetector

noise remains a problem to overcome.

A recent study on interferometric detection-based current sensing involves a popular optical fiber monitoring tool, Phase-Sensitive Optical Time Domain Reflectometer (Phase-OTDR), which allows observing the phase change of Rayleigh backscattered lightwaves as a function of position in the fiber (Sirin, Aldogan, and Wuilpart, 2022). It combines the Phase-OTDR with a Fiber Bragg Gratings (FBG)-inscribed spun fiber coil. Simulations in the mentioned study show that the photodetector thermal noise poses a restriction on the accurately measurable minimum current. The data corrupted by the photodetector noise in Phase-OTDR is obtained for each positional sample point on the sensing coil, constituting an array of phase values involving complex temporal (positional) patterns. Extraction of the especially low-level current value from those patterns requires deep learning-based analysis and data-driven approaches due to their ability to extract insights and make more accurate predictions from complex datasets (LeCun, Bengio, and Hinton, 2015; Z. Li et al., 2022).

Deep learning (DL) has gained popularity because of its capacity to automatically learn features from large datasets. DL techniques may overcome the limitations of traditional methods, which depend on manually crafted features, set a benchmark for future methods, and improve further as more data becomes available (Y. Yang et al., 2021). By utilizing different types of DL models, the analysis of complex temporal patterns becomes feasible, effectively differentiating between low and similar signal levels, leading to a remarkable improvement in detection system performance (Cho et al., 2020; Goodfellow, Bengio, and Courville, 2016; Sikdar, Liu, and Kundu, 2022). The success of these methods across various domains underscores the key benefits of DL-based data-driven approaches, such as enhanced accuracy, automated feature engineering, scalability, flexibility, and interpretable results (Alibabaei et al., 2022; Belhadi et al., 2021). These advantages make DL-based analysis and data-driven approaches valuable tools for addressing a wide range of problems in various fields, including medical imaging and autonomous vehicles to finance and beyond (Alzubaidi et al., 2021; He, Wang, and Zhou, 2022; Lin et al., 2022; Odusami, Maskeliūnas, and Damaševičius, 2022). Utilizing these approaches is crucial for maximizing the insights from complex datasets and developing efficient solutions to real-world problems, making them indispensable for professionals working with large, complex datasets. One of the primary benefits of DL-based analysis is its ability to learn directly from data, eliminating the need for explicit feature extraction or manual preprocessing. Additionally, this approach can reduce the need for labor-intensive and error-prone data preprocessing steps often required in traditional methods to extract relevant features (Farsi et al., 2021; Karimi et al., 2020). Moreover, DL-based analysis can surpass traditional methods in speed, particularly when handling large amounts of

data. Once trained, these models can rapidly process new data and generate results without extensive manual analysis or preprocessing (Ditria et al., 2020). Overall, DL-based analysis offers a powerful and flexible approach to extracting insights from complex data sets, and its ability to learn directly from data, reduce the need for manual preprocessing, and surpass traditional methods in speed makes it an indispensable tool for professionals working in various domains.

In the field of fiber optic sensors, DL has shown promise in improving the accuracy and reliability of sensor systems, particularly in cases with small sample sizes or limited computational resources. Shi et al. proposed a multi-event classification method using DL and support vector machine (SVM) for Phase-OTDR distributed optical fiber sensing systems (Y. Shi et al., 2020). They used a deep convolutional neural network (CNN) to extract features from the raw vibration signals and then applied an SVM to classify the events. The proposed method outperformed other traditional classification methods and achieved high accuracy. In 2021, Shi et al. also proposed an easy-access method for event recognition of Phase-OTDR sensing systems based on transfer learning (Y. Shi et al., 2021). They used a pre-trained deep neural network to extract features from the vibration signals. The proposed method achieved high accuracy with limited training data. Rao et al. reviewed recent advances in Phase-OTDR, including its principles, applications, and challenges (Rao et al., 2021). The review highlighted the potential of Phase-OTDR in various fields, such as energy, aerospace, and transportation, and provided insights into the future development of Phase-OTDR. Kandamali et al. reviewed machine learning (ML) methods for event identification and classification in Phase-OTDR systems (Kandamali et al., 2022), where the current state-of-the-art methods were analyzed in terms of their advantages and limitations. The review provided guidance for researchers to choose appropriate ML methods for their applications. Finally, Ming Wang et al. proposed a Phase-OTDR pattern recognition method based on CNN–LSTM (M. Wang et al., 2023). They used CNN to extract spatial features and an LSTM to capture the temporal patterns of the vibration signals. The proposed method achieved high accuracy in event recognition and could be applied to various Phase-OTDR systems.

In this study, we propose a novel DL-based approach to enhance the performance of a Phase-OTDR-based current sensing system by accurately classifying phase data corrupted by photodetector noise into intervals of electrical current values. Our approach employs three DL models: 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bi-LSTM. These models were trained and tested by the data acquired from the Phase-OTDR-based current sensing simulation environment. To the best of our knowledge, this is the first DL-based optical current sensing study. Our proposed approach, employing DL models, reduces the maximum error margin under 20 A, addressing the challenge of accurately measuring

low-level currents. Accurate low-level current measurements are crucial for applications such as on-aircraft lightning detection, railway stray current detection, and nuclear power plants. While fiber optic current sensors offer a contactless current sensing method, the sensitivity of optical fiber to the Faraday effect is relatively low, making the measurements of low-level currents challenging due to various noise sources. The proposed DL-based approach overcomes this challenge by effectively analyzing and interpreting the complex temporal patterns in the data, resulting in high levels of current distinction and significantly decreasing detection limits. The success of this approach highlights several key advantages, including increased accuracy and the ability to interpret noisy data, making it a promising solution for future optical current sensing applications.

4.2. Methodology

4.2.1. Phase-OTDR-based Current Sensing Simulations and Data Preparation

Phase-OTDR components and the current sensing system are illustrated in Figures 4.1 and 4.2 respectively.

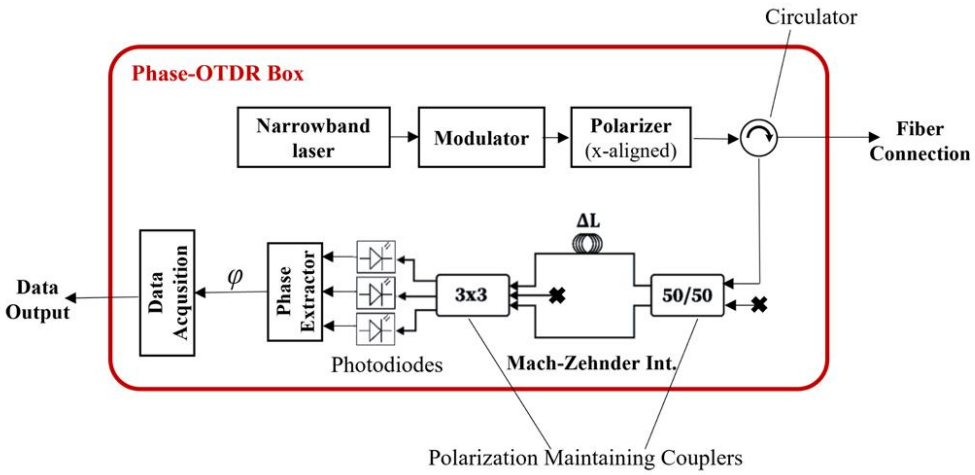

Figure 4.1. Phase-OTDR Interrogator Unit.

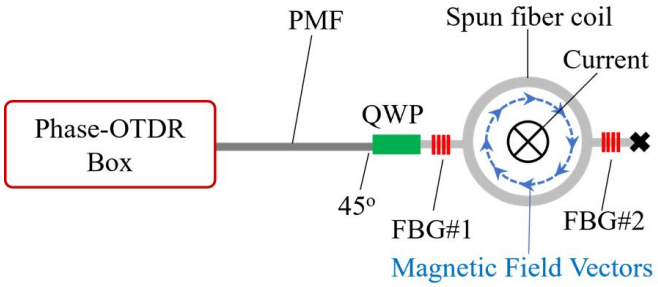

Figure 4.2. Phase-OTDR-based current sensing system. PMF: Polarization Maintaining Fiber, QWP: Quarter-Wave Plate, FBG: Fiber Bragg Grating.

The working principle of a Phase-OTDR-based current sensor involves using coherent light generated by a narrowband laser. This light is time-gated by a modulator, and the resulting optical pulse is linearly polarized by a polarizer in parallel with the x-axis of polarization-maintaining components in the system. The pulse is then transmitted through polarization-maintaining fiber (PMF) to a quarter-wave plate (QWP), which converts the polarization from linear to circular. This circular polarization is essential for the linear response of the pulse phase to the Faraday effect.

Before entering the current sensing coil, a small part of the light is reflected by the first of two 20 dB-reflective fiber Bragg gratings (FBGs). After the pulse is exposed to the Faraday effect along the coil, some reflects back from the second FBG. As the reflected lightwaves pass through the QWP, they are linearly polarized in the y-axis. The PMF carries them to a polarization-maintaining coupler, which directs them to a Mach-Zehnder interferometer. One arm of the interferometer is imbalanced by double the distance between the FBGs, allowing the reflected lightwaves to be interfered in a 3x3 coupler. The interfered light power is detected by three photodiodes to prevent signal fading.

The detector outputs can be expressed by equations 4.1, 4.2 and 4.3 (Masoudi and Newson, 2017a, 2017b):

$$P_1(t,x) = A(x) + B(x)\cos\left(\varphi(t,x)\right) + N_1(t,x) \tag{4.1}$$

$$P_2(t,x) = A(x) + B(x)\cos\left(\varphi(t,x) - \frac{2\pi}{3}\right) + N_2(t,x) \tag{4.2}$$

$$P_3(t,x) = A(x) + B(x)\cos\left(\varphi(t,x) + \frac{2\pi}{3}\right) + N_3(t,x) \tag{4.3}$$

where $\varphi(t,x)$ is the Faraday phase shift, A and B are constants, and $N_{1,2,3}(t,x)$ are additive white Gaussian noise (AWGN) terms coming as the detector noise. The standard deviation of AWGN is the minimum detectable power given as integrated noise in some datasheets. DC component of the detected powers can be removed by substituting their average from themselves as in Equation 4.4.

$$\tilde{P}_{1,2,3}(t,x) = P_{1,2,3}(t,x) - \frac{1}{3}\sum_{n=1}^{3} P_n(t,x) \tag{4.4}$$

Finally, Faraday phase shift which is proportional to the current to be measured at time t can be extracted as in Equation 4.5:

$$\varphi(t,x) = \arctan \frac{\tilde{P}_2(t,x) - \tilde{P}_3(t,x)}{\tilde{P}_1(t,x)} \propto I(t) \tag{4.5}$$

Phase-OTDR simulation tool that is used to generate test and training data is based on an extended version of the one-dimensional impulse response model (Electrical Engineers, Engineering, and Technology, 1985) so that the FBGs and polarization properties are also included besides the Rayleigh backscattering (Sirin, 2022). The simulated sensing coil is circularly winded eight times with a radius of 0.1 meters around the current carrying wire. Detector-integrated noise is assumed to be 17 nW. The pulse length is 2 meters. Parameters of the spun fiber, PMF, FBGs, and detectors are determined according to the components available in the market. The simulation tool is developed in MATLAB 2020a. The dataset is generated using parallel computation on a computer with a 4-core Intel i7. Simulating one measurement for 100 different current values takes 18 minutes on average.

When any of the detector outputs in equations 4.1-4.3 is plotted, three peaks are observed due to FBGs where the middle peak corresponds to the interference of FBG reflections as shown in 4.3.a Current sensing part of the trace is the middle peak, also called FBG interference zone. Considering the same pulse shot, phase values obtained from each spatial point in this region carry information about the same current value. Thus, they must be equal as shown in 4.3.b under ideal conditions, such as no noise, proper polarization of the pulse, and perfect birefringence conditions. When the detector noise is included, they differ significantly from each other, as shown in 4.3.c, which makes reading the current value difficult.

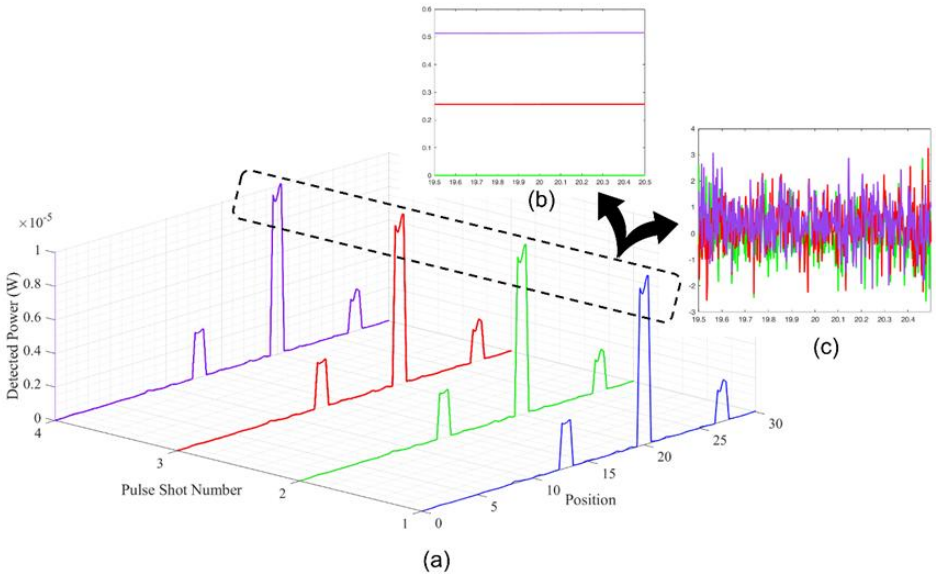

Figure 4.3. a) Power trace P_1 as a function of position x and pulse shot number (corresponds to time t). Middle peaks are FBG interference zone where the current values can be read from. b) Extracted phase when noise is off. c) Extracted phase when noise is on.

There are 401 spatial points in the FBG interference zone, which is the length of the phase data in each pulse shot. Current values range from 0 A to 99 A with a step size of 1 A in each pulse shot, so the overall phase data of one measurement is a 100 x 401 matrix. A total of 173 measurements were allocated for the DL models, distributed as follows: 52% for training (90 measurements), 19% for validation (33 measurements), and 29% for testing (50 measurements). Moreover, K-fold cross-validation was performed on the dataset. Our objective was to develop a model that could accurately classify current values based on the collected data. To achieve this, we analyzed the results of different scenarios, as shown in Table 4.1, to determine the most effective classification methods and identify improvement areas. In addition, scenarios were illustrated Figure 4.4. We studied different buffer sizes during classification and changed the current levels within the same class to see how it affected classification accuracy. By analyzing the results of each scenario, we identified the most effective methods to classify the data accurately and found the proper margin between current values for reduced error.

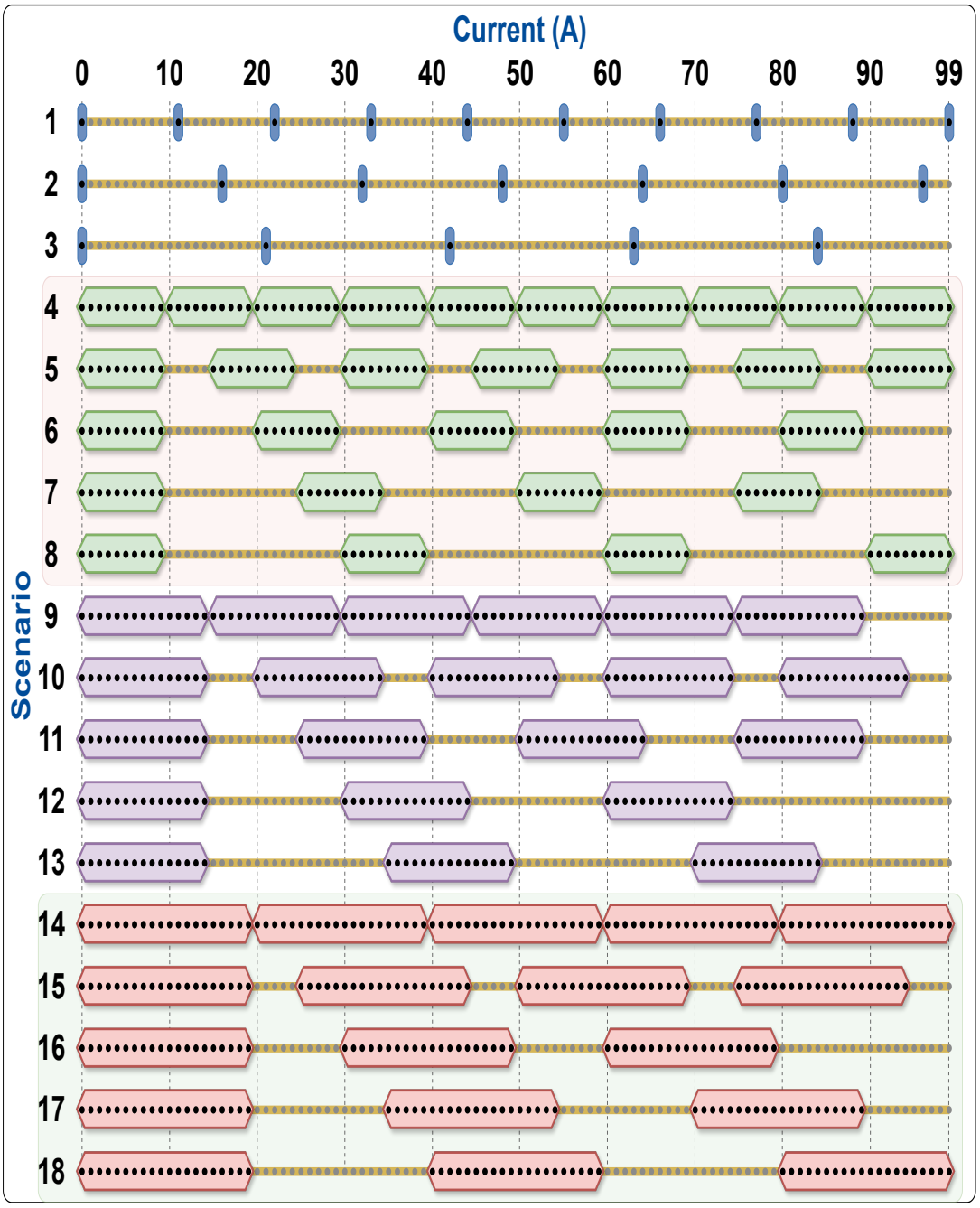

Figure 4.4. Scenario-specific data selection from the dataset. The yellow line represents the entire dataset, while the black dots enclosed indicate the selected values preferred for the specific scenario. Correspondingly, the gray dots dispersed along the yellow line signify the data points omitted from consideration within the scope of the given scenario.

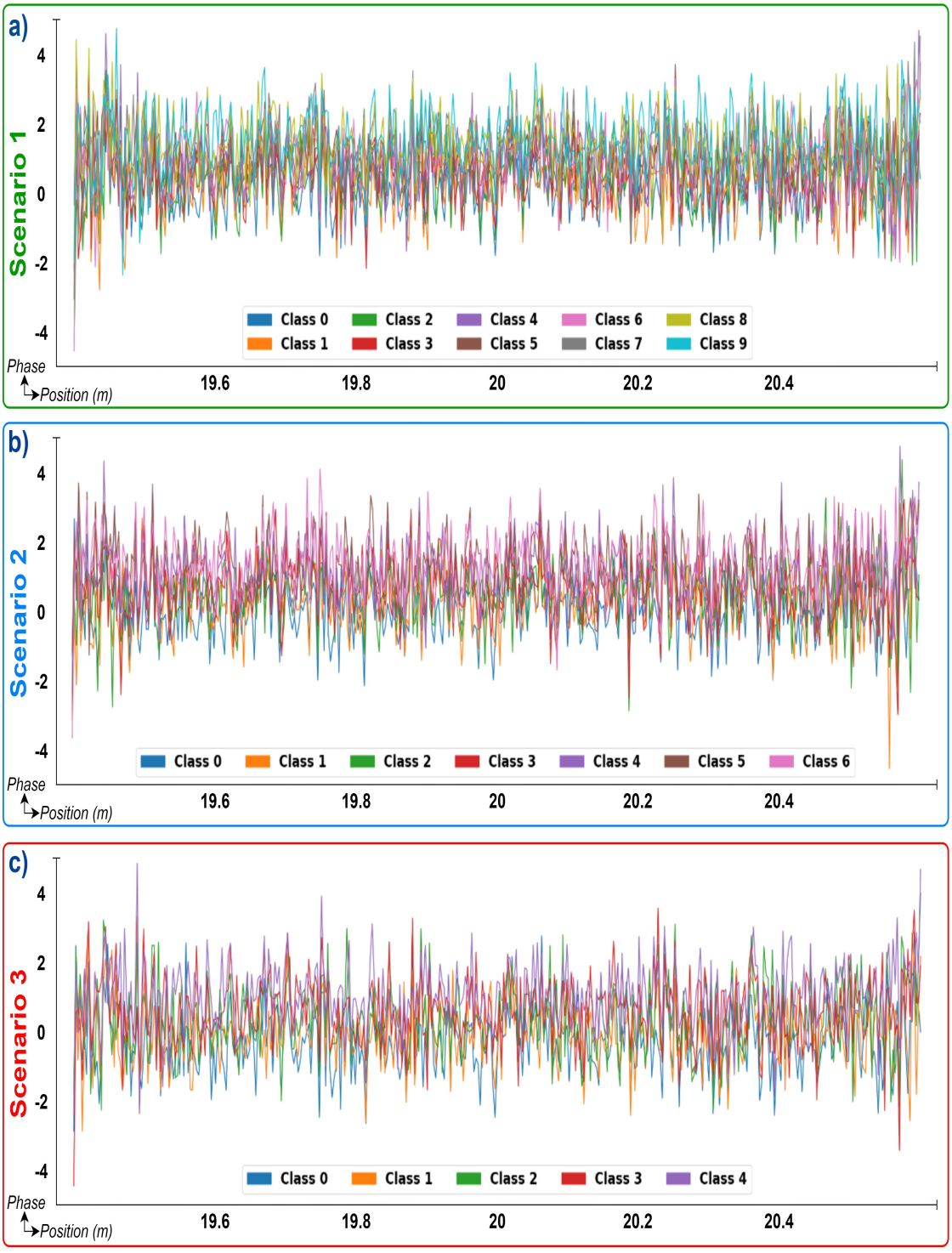

Figure 4.5. Figure illustrates the first three scenarios from Table 4.1, where specially designed current values are represented as distinct classes: a) Ten current values in the first scenario, b) Seven current values in the second scenario, and c) Five current values in the third scenario.

Table 4.1. This table categorizes current values observed during the experiment into distinct scenarios. Each row represents a scenario with class intervals, and columns denote class ranges from Class 0 to Class 9.

Scenario	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Total Number of Classes
1	0 A	11 A	22 A	33 A	44 A	55 A	66 A	77 A	88 A	99 A	10
2	0 A	16 A	32 A	48 A	64 A	80 A	96 A				7
3	0 A	21 A	42 A	63 A	84 A						5
4	0-9 A	10-19 A	20-29 A	30-39 A	40-49 A	50-59 A	60-69 A	70-79 A	80-89 A	90-99 A	10
5	0-9 A	15-24 A	30-39 A	45-54 A	60-69 A	75-84 A	90-99 A				7
6	0-9 A	20-29 A	40-49 A	60-69 A	80-89 A						5
7	0-9 A	25-34 A	50-59 A	75-84 A							4
8	0-9 A	30-39 A	60-69 A	90-99 A							4
9	0-14 A	15-29 A	30-44 A	45-59 A	60-74 A	75-89 A					6
10	0-14 A	20-34 A	40-54 A	60-74 A	80-94 A						5
11	0-14 A	25-39 A	50-64 A	75-89 A							4
12	0-14 A	30-44 A	60-74 A								3
13	0-14 A	35-49 A	70-84 A								3
14	0-19 A	20-39 A	40-59 A	60-79 A	80-99 A						5
15	0-19 A	25-44 A	50-69 A	75-94 A							4
16	0-19 A	30-49 A	60-79 A								3
17	0-19 A	35-54 A	70-89 A								3
18	0-19 A	40-59 A	80-99 A								3

In Table 4.1 provides a detailed overview of the various scenarios created to categorize the current values observed during the experiment.

In this study, we analyzed the effects of different scenarios on classification accuracy. In each scenario, different buffer sizes were used during the classification process, and current levels within the same class were varied. As a result of these analyses, we identified the most effective methods for accurately classifying the data and found the appropriate margin between current values to reduce error. Additionally, by examining the results of each scenario, we identified areas for improvement.

4.2.2. Deep Learning Algorithms

4.2.2.1. One-Dimensional Convolutional Neural Network (1D-CNN)

In this study, we utilized a 1D-CNN to analyze Phase-OTDR data. The 1D-CNN can capture patterns in one-dimensional data, such as time series or sensor signals,

making it an ideal choice for analyzing Phase-OTDR data. The model used in this study comprises three layers of convolutional filters with different filter numbers and filter sizes, as shown in the first column of Table 4.2. These layers aim to extract features at various scales, allowing the model to capture fine and coarse-grained data patterns. To reduce the dimensionality of the output and prevent overfitting, each convolutional layer is followed by a max-pooling layer.

Table 4.2. Details of 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bi-LSTM model architectures and their respective total number of trainable parameters.

1D-CNN	**1D-CNN – LSTM**	**1D-CNN – Bi-LSTM**
Conv1D (Filters=32, ReLU)	Conv1D (Filters=32, ReLU)	Conv1D (Filters=32, ReLU)
MaxPooling1D	MaxPooling1D	MaxPooling1D
Conv1D (Filters=64, ReLU)	Conv1D (Filters=64, ReLU)	Conv1D (Filters=64, ReLU)
MaxPooling1D	MaxPooling1D	MaxPooling1D
Conv1D (Filters=128, ReLU)	Conv1D (Filters=128, ReLU)	Conv1D (Filters=128, ReLU)
MaxPooling1D	MaxPooling1D	MaxPooling1D
Flatten	LSTM (Filters=128, ReLU)	Bi-LSTM (Filters=128, ReLU)
Dense (Units=64, ReLU)	Dropout (rate=0.5)	Dropout (rate=0.5)
Dense (Units=32, ReLU)	LSTM (Filters=64, ReLU)	Bi-LSTM (Filters=64, ReLU)
Dense (n Classes, Softmax)	Dropout (rate=0.5)	Dropout (rate=0.5)
	Dense (Units=64, ReLU)	Dense (Units=64, ReLU)
	Dense (Units=32, ReLU)	Dense (Units=32, ReLU)
	Dense (n Classes, Softmax)	Dense (n Classes, Softmax)
The total number of trainable parameters with 10 units at the last dense layer:		
426730	218602	469127

Furthermore, the output of the last pooling layer is flattened and fed into two fully connected layers with Rectified Linear Unit (ReLU) activation functions to learn high-level representations of the input data. The ReLU activation function introduces non-linearity into the model, enabling it to learn complex relationships between the input and output. The output layer of the model uses the softmax activation function to generate the classification probabilities for each class.

4.2.2.2. Hybrid Neural Network Models

This study also used two hybrid models: ID-CNN - LSTM and ID-CNN - Bi-LSTM. Both models combine ID-CNN layers with LSTM or Bi-LSTM layers to create

hybrid structures.

The first model, ID-CNN - LSTM, begins with three Conv1D layers. These layers have 32, 64, and 128 filters respectively, and each uses the ReLU activation function. Each Conv1D layer is followed by a MaxPooling1D layer. These layers extract features from the input data. Next, there is an LSTM layer with 128 filters, followed by another LSTM layer with 64 filters. After each LSTM layer, Dropout layers are added to prevent overfitting, with a dropout rate of 0.5. Finally, there are two Dense layers; the first Dense layer has 64 units, and the second Dense layer has 32 units, both using the ReLU activation function. The model ends with a Dense layer with n classes and a Softmax activation function.

The second model, ID-CNN - Bi-LSTM, also begins with three Conv1D layers. These layers have 32, 64, and 128 filters respectively, and each uses the ReLU activation function. Each Conv1D layer is followed by a MaxPooling1D layer. These layers extract features from the input data. Next, there is a Bi-LSTM layer with 128 filters, followed by another Bi-LSTM layer with 64 filters. After each Bi-LSTM layer, Dropout layers are added to prevent overfitting, with a dropout rate of 0.5. Finally, there are two Dense layers; the first Dense layer has 64 units, and the second Dense layer has 32 units, both using the ReLU activation function. The model ends with a Dense layer with n classes and a Softmax activation function.

These hybrid models use Conv1D and MaxPooling1D layers for feature extraction, and LSTM or Bi-LSTM layers for processing time series data. Dropout layers are used to prevent overfitting, and Dense layers are used for classification. This structure enhances both the accuracy and generalization capability of the models.

4.2.2.3. Validation Techniques

In this study, two validation methods were employed: Holdout Validation and K-Fold Cross-Validation. These validation methods ensure that the models are thoroughly evaluated and their performance is accurately measured, enhancing the reliability of the results.

In Holdout Validation, the dataset is divided into three subsets: training, validation, and testing. The model is trained on the training set, validated on the validation set, and tested on the test set. The model with the minimum validation loss is saved. The validation accuracy results are presented in Figure 4.7 and Table 4.3. The saved model's performance is evaluated using confusion matrices and other performance metrics, as shown in Figure 4.8 and Table 4.4.

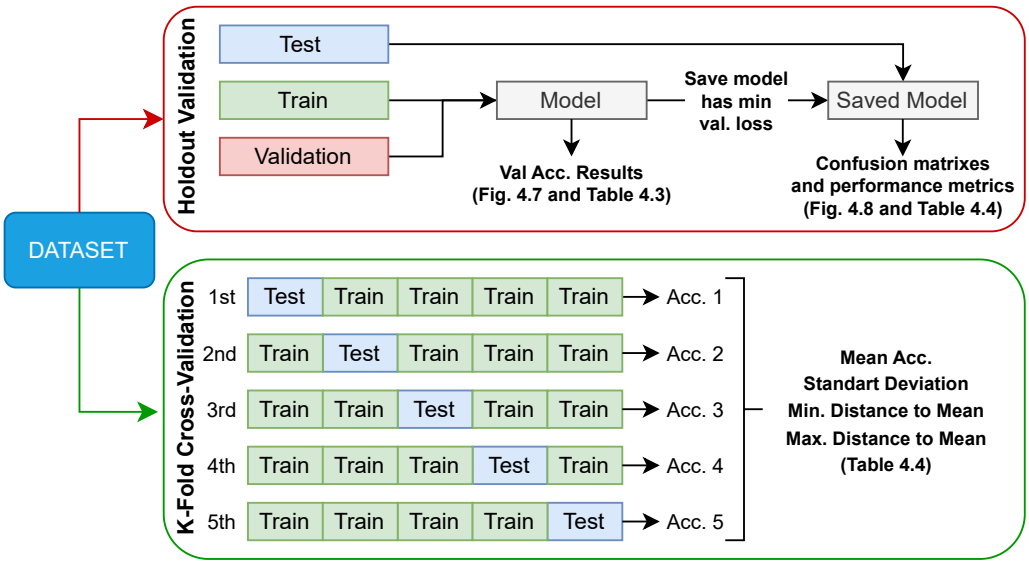

Figure 4.6. A flow diagram illustrating the methods employed during dataset evaluation. The dataset has been assessed using two distinct approaches. The first method partitioned the dataset into test, train, and validation sets and fed into DL models. As a second approach, K-fold cross-validation was applied to evaluate the dataset.

In K-Fold Cross-Validation with k=5, the dataset is divided into 5 subsets (folds). The model is trained and validated 5 times, each time using a different fold as the validation set and the remaining 4 folds as the training set. This process is repeated 5 times, and the validation accuracy for each fold is recorded (Acc. 1, Acc. 2, Acc. 3, Acc. 4, and Acc. 5). The mean accuracy and standard deviation are calculated and presented in Table 4.4.

4.3. Results and Discussion

Fiber optic current sensors offer several advantages, such as inherent electrical isolation, multiplexing capability, compact size, and lightweight. By using Phase-OTDR to interrogate the optical fiber for current sensing purposes, it is possible to monitor several current wires simultaneously using a single fiber line. However, the problem of photodetector thermal noise can significantly affect the accuracy in low-current regimes.

Many traditional methods for analyzing Phase-OTDR data focus on noise reduction through signal processing techniques (Bai, Lin, and Zhong, 2021; Wu et al., 2015). However, these processes are time-consuming and lack standardized pre-processing methods. Additionally, such techniques might result in the loss of crucial features and a struggle

to identify intricate patterns within the data. To address these challenges, DL-based approaches present a promising solution for classifying Phase-OTDR-based signals. These approaches offer increased accuracy, scalability, and flexibility, along with automated feature engineering. Recent studies have demonstrated the effectiveness of DL methods in event classification and recognition for Phase-OTDR-based sensing systems.

Table 4.3. Validation accuracies of DL models for classifying different current levels in the data in various scenarios.

Scenario	**Distinct Current Values In A Class**	**Skipped Data**	**Total Number of Classes**	**Max Validation Acc in 50 Epochs**		
				1D-CNN	***1D-CNN – LSTM***	***1D-CNN – Bi-LSTM***
1	-	10	10	0.882	0.909	0.885
2	-	15	7	0.961	0.970	0.978
3	-	20	5	0.994	0.994	0.994
4	10	-	10	0.727	0.752	0.747
5	10	5	7	0.904	0.915	0.919
6	10	10	5	0.977	0.983	0.983
7	10	15	4	0.993	0.999	0.998
8	10	20	4	0.999	0.999	1.000
9	15	-	6	0.841	0.852	0.852
10	15	5	5	0.943	0.950	0.952
11	15	10	4	0.983	0.987	0.987
12	15	15	3	0.999	0.999	0.998
13	15	20	3	1.000	1.000	1.000
14	20	-	5	0.879	0.896	0.889
15	20	5	4	0.954	0.968	0.967
16	20	10	3	0.994	0.994	0.994
17	20	15	3	0.999	1.000	1.000
18	20	20	3	0.999	1.000	1.000

This study aimed to investigate the effectiveness of DL models in classifying data from different current levels into distinct classes. The dataset used in the study included current values ranging from 0 A to 99 A, and various scenarios were generated by grouping the data differently. For each scenario, Table 4.3 presented the validation accuracies achieved by three DL models using the evaluation method of partitioning the dataset into training, testing, and validation subsets. The results showed that the proposed DL-based approaches significantly enhance the performance of Phase-OTDR-based current sensing and improve classification accuracy, especially in the presence of photodetector thermal noise.

The validation accuracy curve, delineated in Figure 4.7, illustrates the dynamic shifts in model accuracy as the number of epochs increases. For this demonstration, the first three scenarios were selected. Within these scenarios, it was observed that the accuracy of the models increased rapidly in the first few epochs and then started to plateau. After 10 epochs, the 1D-CNN model achieved an accuracy of around 99% (Figure 4.7.a), while both the 1D-CNN – LSTM and 1D-CNN – Bi-LSTM models achieved, but only after 18 epochs (Figure 4.7.b and 4.7.c). A closer inspection of Table 4.3 revealed that although the 1D-CNN model reached high validation accuracy more quickly (as depicted in Figure 4.7), the other two models exhibited superior performance in various scenarios.

In scenarios 4 to 8, the model's response to a varying number of distinct current values skipped between classes was investigated. The models exhibit a lower accuracy rate in Scenario 4, which encompasses 10 current values per class. However, as seen in Scenario 5, the models achieve higher accuracies with ten current values per class and five skipped data points. This implies that introducing skipped data might aid the models in better differentiating the different sparse patterns and increase the model's accuracy. This positive trend persisted in Scenario 6, where ten different current values coexist with ten skipped data points, and a notable improvement in the models' performance was observed. This suggests that the increased sparsity introduced by diverse current values and skipped data can lead to better model performance with higher accuracy. The models reach high accuracy levels in Scenarios 7 and 8, where ten different current values within a class are combined with fifteen and twenty skipped data points, respectively. These results show that it underlines its robustness in handling sparsity in complex data.

In scenarios 9 through 13, each class is created by a sequence of fifteen consecutive current values. Additionally, scenarios 14 to 18 are created such that each class encompasses a sequence of twenty consecutive current values. These scenarios were formulated to analyze the effect of modifying the number of distinct current values within a class on the accuracy of that class. An increase in the number of consecutive current values within a class positively impacted the accuracy rate. Incorporating these scenarios helps in developing a more comprehensive understanding of how alterations in the distribution of current values within a class can influence the accuracy of results.

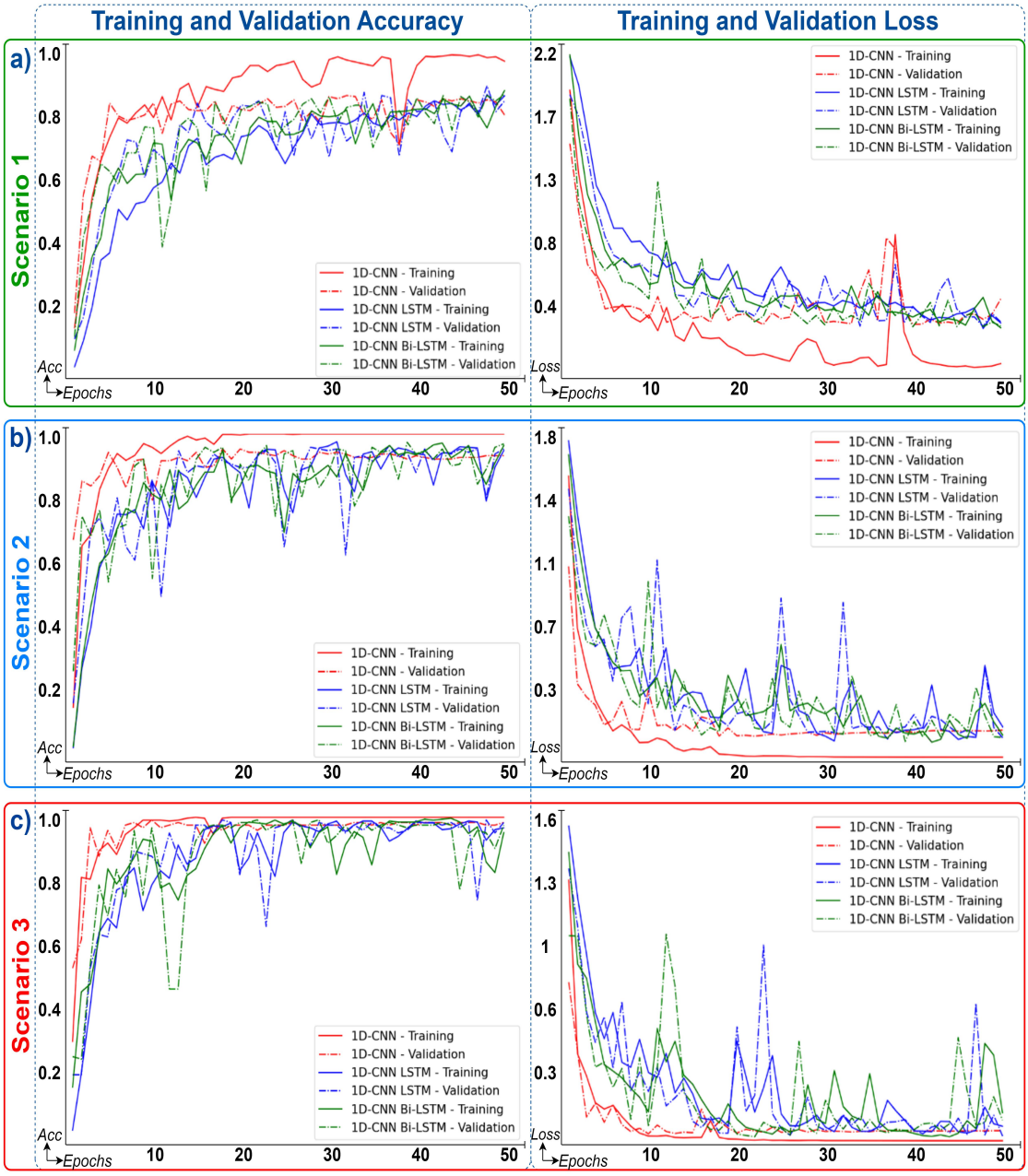

Figure 4.7. Graphical representation of validation accuracy and loss across epochs for three distinct scenarios (a, b, and c). Each graph illustrates the performance trajectory of three different models: 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bi-LSTM, facilitating a detailed comparison of their respective performances over the progression of epochs.

Evaluating the classification performance of ML models is crucial in various fields, and the confusion matrix is a widely used tool for such assessment. The confusion matrix represents the true and predicted labels for each class, providing insight into the model's performance and facilitating the computation of metrics such as accuracy,

precision, and recall. We save the weights of a model that has the lowest validation loss value out of 50 epochs during training and then feed the test data into the saved model to create the confusion matrix, which displays the number of correctly and incorrectly classified samples for each class, providing valuable insights into the model's classification performance. We present the confusion matrices for the first three scenarios selected from Table 4.1, with Figure 4.8 displaying the true and predicted labels for each class, providing a comprehensive view of the classification performance of the models.

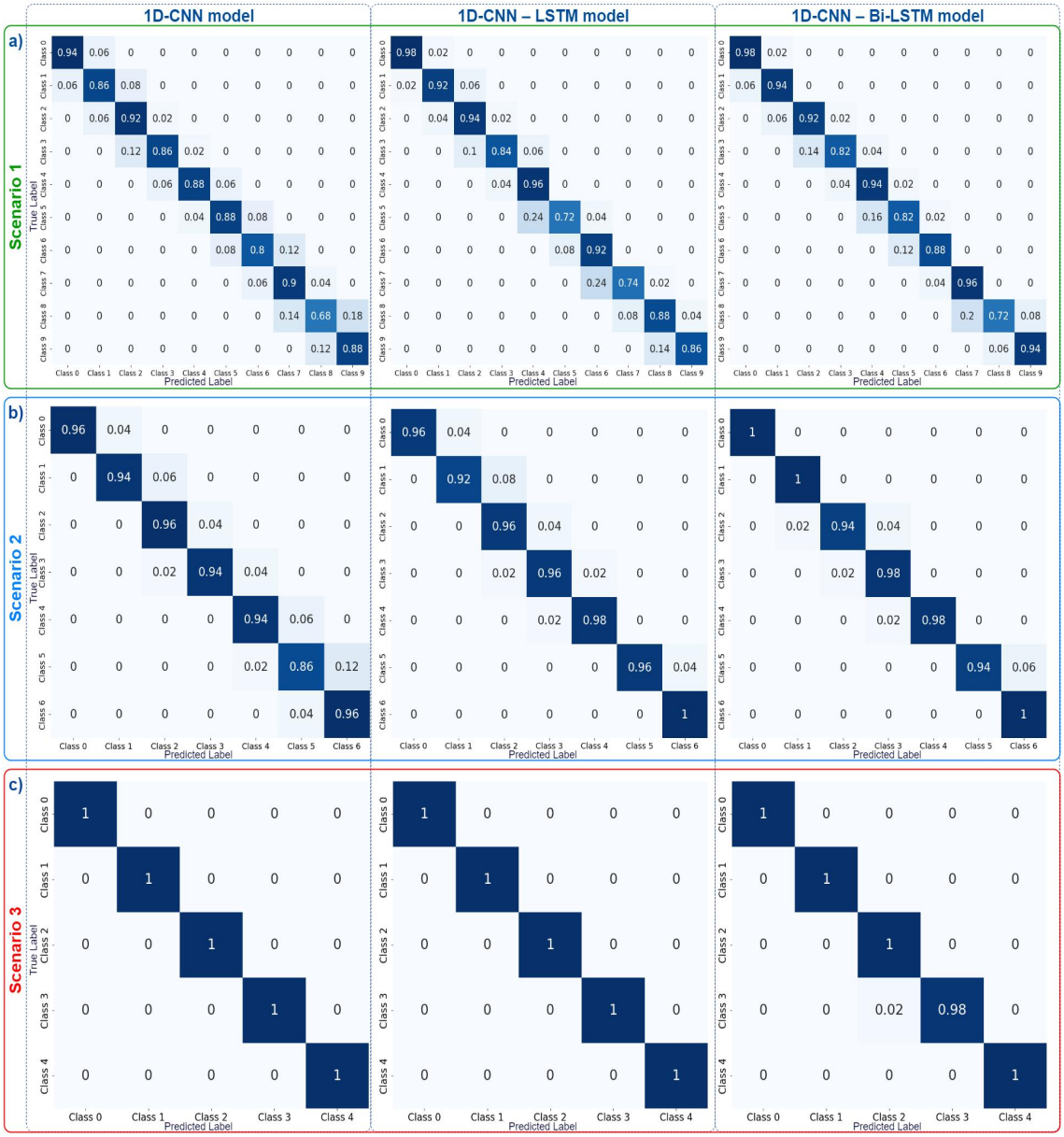

Figure 4.8. Classification performance illustrated by confusion matrices. Each matrix represents one of the first three scenarios, evaluated with different DL models.

Table 4.4. Comparison of Performance Metrics for Three Deep Learning Models Across 18 Scenarios.

Scenario	Model	Holdout Validation				K-Fold Cross-Validation	
		Accuracy	*Sensitivity*	*Precision*	*F1 Score*	*Mean Acc*	*SD*
1	1D-CNN	0.860	0.935	0.940	0.859	0.876	0.018
	1D-CNN – LSTM	0.876	0.979	0.980	0.876	0.841	0.049
	1D-CNN – Bi-LSTM	0.892	0.940	0.942	0.891	0.828	0.097
2	1D-CNN	0.937	1.000	1.000	0.937	0.962	0.015
	1D-CNN – LSTM	0.963	1.000	1.000	0.963	0.929	0.038
	1D-CNN – Bi-LSTM	0.977	1.000	1.000	0.977	0.951	0.032
3	1D-CNN	1.000	1.000	1.000	1.000	0.990	0.010
	1D-CNN – LSTM	1.000	1.000	1.000	1.000	0.959	0.067
	1D-CNN – Bi-LSTM	0.996	1.000	1.000	0.996	0.978	0.018
4	1D-CNN	0.732	0.848	0.862	0.728	0.816	0.006
	1D-CNN – LSTM	0.751	0.877	0.885	0.749	0.802	0.012
	1D-CNN – Bi-LSTM	0.751	0.874	0.886	0.750	0.807	0.011
5	1D-CNN	0.897	0.948	0.951	0.897	0.948	0.005
	1D-CNN – LSTM	0.917	0.935	0.939	0.917	0.927	0.006
	1D-CNN – Bi-LSTM	0.922	0.964	0.966	0.922	0.926	0.019
6	1D-CNN	0.979	0.992	0.992	0.979	0.988	0.003
	1D-CNN – LSTM	0.984	0.986	0.986	0.984	0.983	0.001
	1D-CNN – Bi-LSTM	0.985	0.994	0.994	0.985	0.982	0.005
7	1D-CNN	0.994	1.000	1.000	0.993	0.997	0.001
	1D-CNN – LSTM	0.994	0.996	0.996	0.994	0.997	0.002
	1D-CNN – Bi-LSTM	0.994	0.996	0.996	0.993	0.989	0.012
8	1D-CNN	0.999	1.000	1.000	0.999	0.999	0.001
	1D-CNN – LSTM	0.999	0.998	0.998	0.999	0.998	0.001
	1D-CNN – Bi-LSTM	1.000	1.000	1.000	1.000	0.999	0.001
9	1D-CNN	0.833	0.904	0.912	0.834	0.885	0.005
	1D-CNN – LSTM	0.842	0.949	0.949	0.842	0.855	0.042
	1D-CNN – Bi-LSTM	0.846	0.911	0.918	0.846	0.883	0.005
10	1D-CNN	0.935	0.954	0.957	0.935	0.967	0.004
	1D-CNN – LSTM	0.945	0.962	0.964	0.944	0.947	0.016
	1D-CNN – Bi-LSTM	0.952	0.970	0.971	0.952	0.959	0.004
11	1D-CNN	0.983	0.993	0.993	0.983	0.968	0.004
	1D-CNN – LSTM	0.987	0.997	0.997	0.987	0.959	0.005
	1D-CNN – Bi-LSTM	0.988	0.999	0.999	0.988	0.937	0.027
12	1D-CNN	0.998	0.999	0.999	0.998	0.992	0.002
	1D-CNN – LSTM	0.998	1.000	1.000	0.998	0.987	0.004
	1D-CNN – Bi-LSTM	0.998	1.000	1.000	0.998	0.978	0.010
13	1D-CNN	1.000	1.000	1.000	1.000	1.000	0.000
	1D-CNN – LSTM	0.999	1.000	1.000	0.999	0.999	0.001
	1D-CNN – Bi-LSTM	1.000	1.000	1.000	1.000	0.999	0.001
14	1D-CNN	0.872	0.901	0.918	0.871	0.913	0.007
	1D-CNN – LSTM	0.888	0.948	0.949	0.888	0.909	0.006
	1D-CNN – Bi-LSTM	0.881	0.918	0.923	0.881	0.907	0.005
15	1D-CNN	0.954	0.977	0.978	0.954	0.978	0.002
	1D-CNN – LSTM	0.966	0.986	0.986	0.966	0.968	0.005
	1D-CNN – Bi-LSTM	0.967	0.984	0.984	0.967	0.969	0.005
16	1D-CNN	0.992	0.995	0.995	0.992	0.996	0.002
	1D-CNN – LSTM	0.990	0.990	0.990	0.990	0.991	0.005
	1D-CNN – Bi-LSTM	0.992	0.991	0.991	0.992	0.994	0.002
17	1D-CNN	0.995	0.994	0.994	0.995	0.998	0.002
	1D-CNN – LSTM	0.997	0.997	0.997	0.997	0.997	0.001
	1D-CNN – Bi-LSTM	0.996	0.994	0.994	0.996	0.998	0.001
18	1D-CNN	1.000	1.000	1.000	1.000	0.999	0.001
	1D-CNN – LSTM	0.999	1.000	1.000	0.999	0.997	0.004
	1D-CNN – Bi-LSTM	1.000	1.000	1.000	1.000	0.999	0.001

As observed in Figure 4.8, the use of expanded buffer regions significantly facilitates the distinction of different classes. These extended buffer regions play a crucial role in enhancing classification accuracy. Particularly in the first two scenarios depicted in Figure 4.8, it can be observed that incorrect predictions are often close to the diagonal direction, highlighting the effectiveness of the DL approach. By employing buffer regions in Scenario 1 for 10 A, Scenario 2 for 15 A, and Scenario 3 for 20 A levels, an improvement in classification accuracy is observed as the buffer regions expand. This phenomenon demonstrates the model's capability to accurately discern even close current values and the contribution of expanded buffer regions to this success. Additionally, considering the achievement of 100% classification accuracy at the 20 A level, it is evident that discernibility is maintained even with smaller buffer regions.

Table 4.4 presents the performance metrics of three DL models for different scenarios. Each scenario is labeled in the first column, and the corresponding model is specified in the second column. The performance metrics for holdout validation, encompassing accuracy, sensitivity, precision, and F1 score, are reported in columns 3 to 6. These metrics are calculated by recording the epoch with the lowest validation loss over 50 epochs and evaluating the test dataset using the saved model associated with that recorded epoch. Moreover, for dataset analysis, the second method utilized was K-fold cross-validation, and its results were reported. The average accuracy and standard deviation across the five folds were respectively reported in columns 7 and 8. The K-fold cross-validation results emphasize the stability and consistency exhibited by the model's performance. Low standard deviation values point towards the model's reliability across various folds, while high average values signify strong overall performance.

In essence, our study's comprehensive evaluation of classification performance through performance metrics and K-fold cross-validation highlights the profound significance of our findings within the realm of optical current sensing. Not only do these outcomes enrich the existing knowledge, but they also offer invaluable insights for researchers and practitioners engaged in similar classification tasks. The application of DL-based methodologies to enhance the assessment of fiber optic sensor systems underscores the pivotal role that our research plays in advancing this domain. Moreover, the exploration of Scenarios 1 to 18 serves as compelling evidence of the inherent adaptability, robustness, and precision demonstrated by our DL models across a diverse array of class configurations. Their consistent performance under varying conditions, encompassing different current values, instances of skipped data, and varying class distributions, further underscores their potential for real-world applications. Our findings provide valuable insights for researchers and practitioners working on similar classification tasks and can contribute to the development of more accurate models in optical current sensing applica-

tions. The potential impact of this study is significant, as it can enhance the performance of Phase-OTDR-based current sensing in the presence of photodetector thermal noise and advance classification tasks in similar fields. The success of this approach highlights several key advantages, including increased accuracy and the ability to interpret noisy data, making it a promising solution for future optical current sensing applications.

4.4. Conclusion

The focus of this study is to introduce DL methods to improve the efficiency of Phase-OTDR-based current sensing systems. To the best of our knowledge, this study is the first DL-based optical current sensing implementation, offering unique insights and advancements in the field. One of the challenges in the accurate measurement of low-current regimes in Phase-OTDR-based current sensing is the presence of photodetector thermal noise. The proposed method presents numerous advantages compared to traditional approaches for classifying Phase-OTDR-based current measurements. It eliminates the need for preprocessing by utilizing raw data, offering speed and simplicity in the process. In this study, a dataset consisting of simulation-based measurements covering a range of current values was created. Three DL algorithms, namely 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bi-LSTM, were applied to scenarios with different current values as corresponding classes. The results highlighted the effectiveness of the 1D-CNN and hybrid models in accurately classifying current values based on Phase-OTDR-based current measurement data. The study also analyzed different scenarios, considering variations in buffer sizes and current levels within the same class, to identify the most effective methods for accurate classification and reducing error. Our analysis of the Phase-OTDR-based current sensing dataset highlighted the noteworthy performance of our method in accurately distinguishing between current levels. It demonstrated its effectiveness in complex classification tasks, achieving a 100% accuracy rate when there was a 20 A buffer zone between the current levels. Even as the buffer zones were gradually reduced, our method consistently maintained high accuracy, reaching 97% and 89% for current levels of 15 A and 10 A, respectively. These findings provide valuable insights for researchers and practitioners in related fields. It is emphasized that DL-based approaches play a critical role in enhancing fiber optic-based electrical current detection performance, especially in challenging scenarios like low current levels and noisy data. The proposed method presents several advantages by enhancing the performance of Phase-OTDR-based current sensing systems and creating new possibilities for improving contactless electrical current measurements. The classification approach used in this work has the potential to

be extended to various time-series data that require high-resolution class determination while maintaining accuracy, even for data that may not be easily distinguishable. The study specifically focuses on differentiating between low and similar current levels and finding the limit of detection to a specified value. In conclusion, this research presents an opportunity to enhance the performance of Phase-OTDR-based current sensing systems and explore new avenues for improving contactless electrical current measurements.

CHAPTER 5

DEEP LEARNING-BASED PHASE OTDR EVENT DETECTION USING IMAGE-BASED DATA TRANSFORMATION

5.1. Introduction

Event detection and classification involve identifying, analyzing, and categorizing patterns within event streams, providing insights into real-world occurrences (L. Liu et al., 2021; Porumb et al., 2020). This process is crucial in various fields, from emergency response to monitoring social media trends and threats (Zishan Ahmad et al., 2022; Kim and Kwon, 2022; Kim et al., 2020; Mredula et al., 2022). It helps understand, respond to, and utilize data effectively in a rapidly changing world.

Using diverse sensor systems enhanced by deep learning, event detection and classification address specific challenges in their domains. Despite the complexities in identifying and classifying events within complex datasets, deep learning technologies integrated with sensor systems demonstrate adaptability and efficacy. For example, deep learning models in acoustic sensors can precisely analyze sound, identifying noise sources in urban areas, which is vital for addressing noise pollution (Bonet-Solà, Vidaña-Vila, and Alsina-Pagès, 2023; Vidaña-Vila et al., 2021). In image processing and video analytics, deep neural networks detect object movements and monitor traffic flow, useful in surveillance and pedestrian observation (Becattini, Palai, and Bimbo, 2022; Ionescu et al., 2019; Su et al., 2022). Seismic sensors classify earthquakes based on magnitude and location (Lomax, Michelini, and Jozinović, 2019; Saad, Hafez, and Soliman, 2021), while wearable devices in healthcare detect medical events like heart arrhythmias or epileptic seizures (Beniczky et al., 2021; Lee et al., 2022; Tang et al., 2021). Deep learning models like CNNs and RNNs autonomously acquire features from diverse sensing modalities, enhancing event detection and classification accuracy.

Phase-sensitive Optical Time-Domain Reflectometer (Phase-OTDR) monitors acoustic events along optical fibers by sending light pulses and measuring the Rayleigh backscattered light. Unlike traditional sensors, Phase-OTDR provides distributed sensing, monitoring the entire length of the fiber and detecting perturbations at different points. This

technology is useful for real-time monitoring of large areas or structures, such as pipelines, railways, and security systems (Mahmoud, 2023; Wu et al., 2019; Xie et al., 2023).

Combining deep learning and distributed acoustic sensing has advanced event detection and classification along fiber-optic lines. Zhao et al. combined Markov Transition Fields and deep learning for event classification, achieving 96.16% accuracy (X. Zhao et al., 2022). Yang et al. applied these methods to high-speed railway security, achieving 99.5% precision and 555 FPS inference speed (Yang, Zhao, and Chen, 2022). In 2023, Yang et al. introduced a real-time event classification method with 94.23% accuracy (N. Yang et al., 2023). Barantsov et al. achieved over 98% accuracy with CNN architectures (Barantsov et al., 2023). Cao et al. created a Phase-OTDR event dataset, achieving over 82% recognition accuracy with SVM and CNN models (Cao et al., 2023). Chen et al. used this dataset with a Dendrite Net-based approach, achieving 98.6% accuracy (X. Chen et al., 2023). Ni et al. and Sun et al. further improved recognition accuracy using data augmentation and feature selection methods (Ni, Hu, and Yu, 2023; Sun and Fang, 2023).

Our study introduces an image-based analysis approach for Phase-OTDR data, converting 1D time series data into visually informative representations using techniques like Gramian Angular Difference Field (GADF), Gramian Angular Summation Field (GASF), and Recurrence Plots (RP). These images are organized into a multi-channel representation and processed for compatibility with deep learning models, enhancing event discrimination and data analysis. This approach leverages transfer learning, adapting pre-trained models for optical fiber sensor data, leading to more effective and adaptable analytical outcomes.

5.2. Methodology

This study introduces a novel approach for analyzing 1D Phase-OTDR data, representing a significant advancement in optical fiber sensing. By employing image-based data transformation techniques such as GADF, GASF, and RP, raw data is converted into informative images, facilitating a comprehensive understanding of complex fiber optic sensing data. The study details a Phase-OTDR setup methodology that captures 15,612 samples across various disturbance events, demonstrating the effectiveness of CNNs and transfer learning for accurate event discrimination. Furthermore, it addresses data size reduction, significantly compressing the dataset from 2.03 GB to 180 MB through innovative image-based conversion methods, thus offering substantial improvements in the analysis of complex fiber optic sensing data.

5.2.1. Measurement Method

In the Phase-OTDR setup, light generated by the laser is amplified using an Erbium-Doped Fiber Amplifier (EDFA) and then filtered to remove amplified spontaneous emission (ASE) noise. An Acousto Optic Modulator (AOM), driven by a function generator (FG), converts the continuous light into light pulses with a duration of 400 ns, corresponding to a spatial resolution of 40 m, and a repetition frequency of 12.5 or 8 kHz depending on the length of the leading fiber. The circulator directs the pulse into the sensing fiber, which consists of three sections. The first section is the leading fiber, 5 or 10 km in length, which is a bare fiber without any covering tube or jacket on the Acrylate coating and is kept isolated from vibration in a soundproof box. The second and third sections, each 50 m long, are armored. The second section is subjected to various vibrational events. Rayleigh backscattered light is directed by the circulator to the photodetector (PD), where its intensity is detected and recorded on the data acquisition card (DAQ). The collected data from the trailing 120 m fiber is then processed on a personal computer (PC).

5.2.2. Dataset

The dataset, publicly shared through the work of Cao et al., was collected by sending successive 10,000 pulses and detecting the intensity of light backscattered from the last 120 meters of the sensing fiber (Cao et al., 2023). Each measurement, taken with a single pulse, contains 12 intensity values of the Rayleigh backscattered signal at 12 equidistant spatial points along the 120-meter fiber section. Consequently, one sample is represented as an intensity matrix with dimensions of 12x10,000. During the measurement of each sample, one of six distinct disturbance events was applied to the second fiber section: background, digging, knocking, watering, shaking, and walking. Figure 5.1 illustrates one sample measurement data for each event.

The dataset, demonstrating significant scale, consists of a total of 15,612 samples, providing a comprehensive resource for training deep learning models. It has been divided into training and test sets in an 8:2 ratio and shared accordingly. Event measurements are conveniently stored in tag files corresponding to each event type. To offer insights into the dataset, a summary of the samples for each event is presented in Table 5.1. All event samples are available in .mat format, and an inherent class imbalance is evident from the varying sample counts across different event categories.

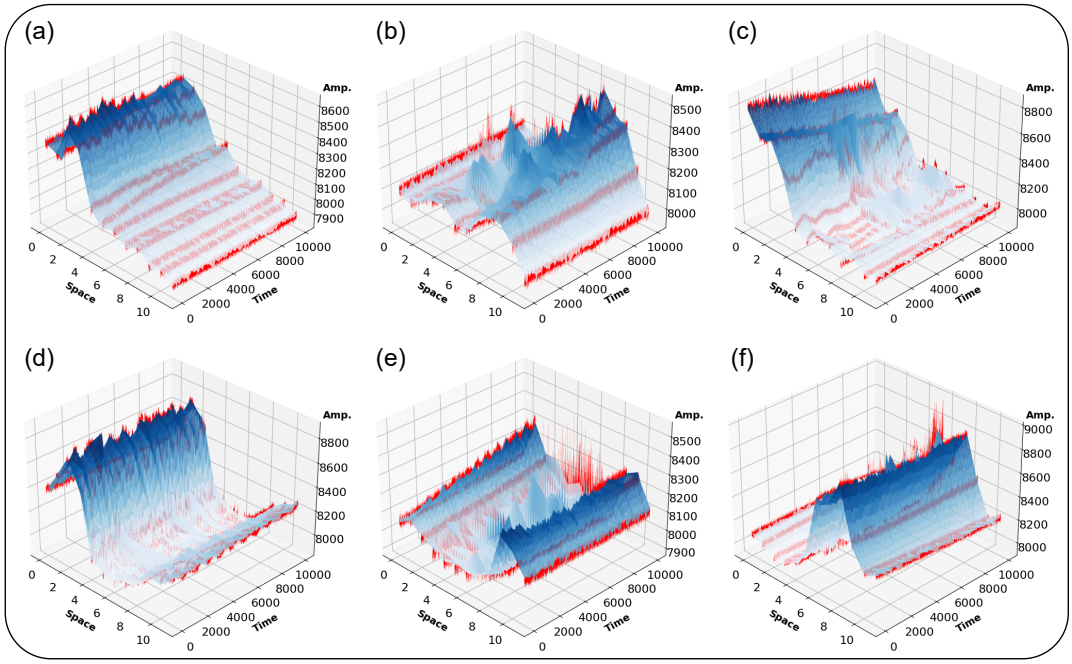

Figure 5.1. This collection of six images displays the interpolated spatial-temporal samples of various events, highlighting the unique features of each scenario. The images correspond to the following events: a) Background, b) Digging, c) Knocking, d) Watering, e) Shaking, and f) Walking. These images show raw data without any preprocessing or normalization, revealing the inherent characteristics of the recorded events. The surfaces are created through interpolation, and the red lines extending from the axes from space to time on the surface indicate the positions of 12 distinct regions on the fiber.

Phase-OTDR data, initially presented as a 1D time series of amplitude information collected over a defined period, can be challenging to analyze directly. Transforming this 1D data into images simplifies analysis and allows for the application of advanced machine learning techniques, representing a significant step towards understanding and interpreting such data (Garcia et al., 2022; Woodward, Kanjo, and Tsanas, 2023). Several techniques, including Gramian Angular Difference Field (GADF), Gramian Angular Summation Field (GASF), and Recurrence Plot (RP), are used to convert the 1D data into image representations (Adib et al., 2023; Zeeshan Ahmad et al., 2021; Batista et al., 2023; J. Wang et al., 2023). Each technique employs unique mathematical approaches to emphasize different data aspects, resulting in a variety of images suitable for in-depth analysis.

Table 5.1. The table shows the distribution of dataset samples for various event types, each associated with a specific label, facilitating easier analysis and classification.

Events	Sample Counts	Labels
Background	3094	0
Digging	2512	1
Knocking	2530	2
Watering	2298	3
Shaking	2728	4
Walking	2450	5
Total	15612	6

In this context, data from 12 spatial points are transformed into individual images using each of these techniques, with each image highlighting distinct data characteristics through specific mathematical methods. These individual images are then arranged into a 3x4 grid, forming a multi-channel representation. This process is applied separately for GADF, GASF, and RP, resulting in three distinct grayscale images, each stored in an 8-bit format. To further enhance analysis, the output of each technique is assigned to a specific color channel – typically red, green, and blue (RGB) – creating color-coded images. These color-coded representations are then combined to ensure compatibility with deep learning models, enabling a comprehensive examination of the data from multiple perspectives.

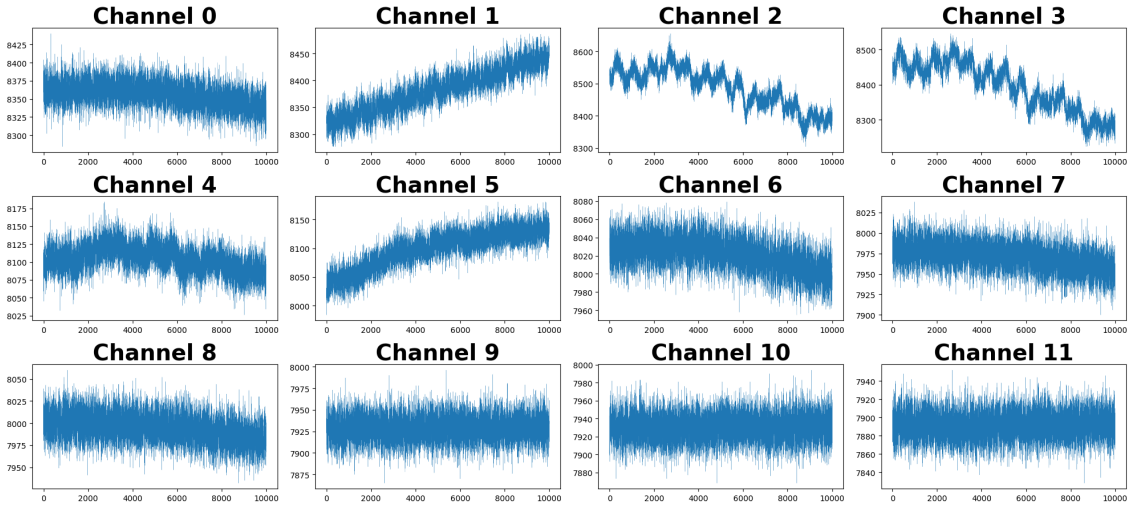

Figure 5.2. Depicts the raw data from 12 spatial regions on the fiber for one event, with each subplot representing a different channel.

Additionally, the dataset originally provided by Cao et al. was composed of .mat files with a total size of 2.03 GB (2,187,235,932 bytes) (Cao et al., 2023). By applying an image-based conversion method, the dataset size was significantly reduced to 180 MB (189,709,203 bytes). This substantial reduction, approximately 11 times smaller, demonstrates the effectiveness and benefits of such a transformation. Beyond the immediate advantage of improved storage efficiency, this downsizing greatly simplifies data management and accessibility. It addresses the challenges of handling large datasets and contributes to faster data processing and easier transmission. The significance of this size reduction, as shown in this paper, highlights the crucial role of innovative data transformation techniques in optimizing data storage and analysis, emphasizing their importance in contemporary data-driven research and applications.

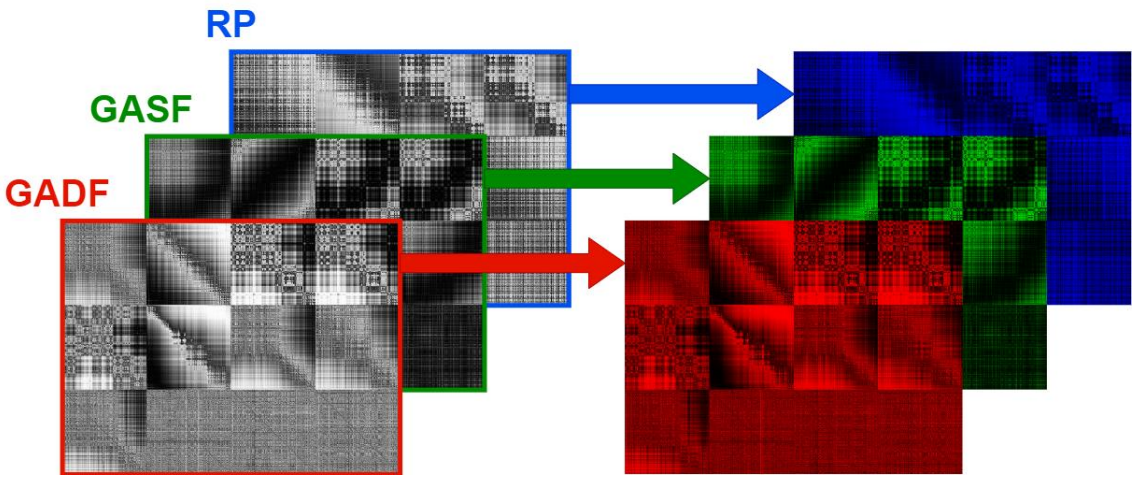

Figure 5.3. The figure shows the assignment of grayscale images, converted from 1D data using Recurrence Plot (RP), Gramian Angular Summation Field (GASF), and Gramian Angular Difference Field (GADF) methods, to RGB channels. The resulting images are organized into a 3x4 grid, with each small image having dimensions of 500x500 pixels, resulting in a combined image of 1500x2000 pixels.

In Figure 5.4 presents examples from the dataset created by combining three distinct color channels (red, green, and blue) and downsampled to 224x224 pixels. This figure contains six images, each representing a different activity. The images exhibit unique patterns corresponding to the specific characteristics of each activity. This dataset can be used for training image-based deep learning models and facilitates the analysis of

different types of events.

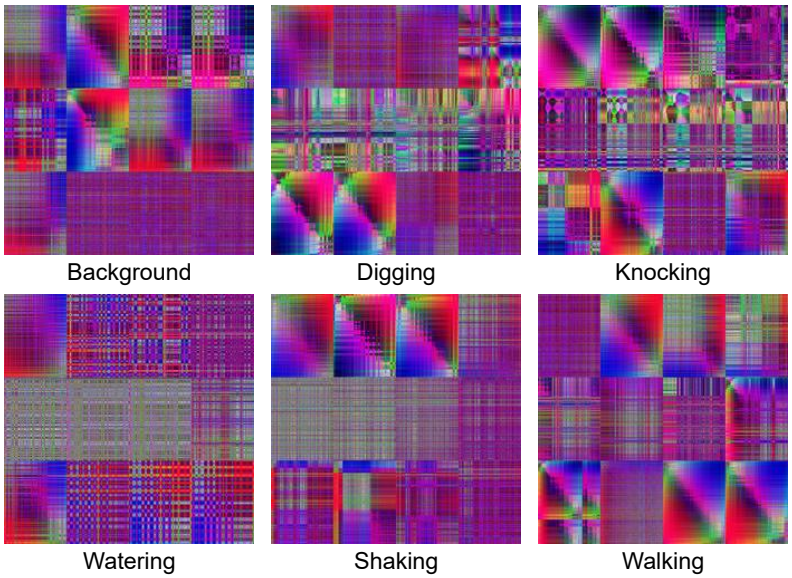

Figure 5.4. Examples from the dataset created by combining color channels. Each image represents a different activity: Background, Digging, Knocking, Watering, Shaking, and Walking.

In conclusion, transforming 1D data into images not only facilitates easier analysis but also offers additional benefits, such as enabling the use of machine learning and reducing dataset size. This approach leads to more efficient data processing and sharper analytical results, underscoring the essential role of image-based analysis.

5.2.2.1. Gramian Angular Field

The Gramian Angular Field (GAF) offers a novel approach to time series data analysis by transforming the data into a visual representation that enhances the understanding of temporal dynamics (Wang and Oates, 2015). This method utilizes polar coordinates for data representation, with angles depicting the interconnections among data points and radial distances indicating time-related data. The primary elements of GAF include the Gram Matrix, GASF Matrix, and GADF Matrix.

The Gram Matrix, also referred to as the Gramian Matrix, serves as a crucial component in GAF. It represents the inner products of vectors within the dataset. For a set of vectors $x_1, x_2, \ldots, x_n$, the Gram Matrix G is formulated as follows:

$$G = X^T X \tag{5.1}$$

$$G = \begin{bmatrix} (x_1, x_1) & (x_1, x_2) & \dots & (x_1, x_n) \\ (x_2, x_1) & (x_2, x_2) & \dots & (x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ (x_n, x_1) & (x_n, x_2) & \dots & (x_n, x_n) \end{bmatrix} \tag{5.2}$$

In this context, X^T denotes the transpose of the data matrix X. The Gram Matrix G encapsulates the pairwise inner products among data points, playing a vital role in the GAF transformation process.

The GASF Matrix, which originates from the Gram Matrix, reflects the cosine values of the summed angles between data points. Each entry in the GASF Matrix represents a pair of data points and is determined as follows:

$$\text{GASF}(x_i, x_j) = \cos(\Theta_i + \Theta_j) \tag{5.3}$$

$$\text{GASF} = \begin{bmatrix} \cos(\Theta_1 + \Theta_1) & \cos(\Theta_1 + \Theta_2) & \dots & \cos(\Theta_1 + \Theta_j) \\ \cos(\Theta_2 + \Theta_1) & \cos(\Theta_2 + \Theta_2) & \dots & \cos(\Theta_2 + \Theta_j) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\Theta_i + \Theta_1) & \cos(\Theta_i + \Theta_2) & \dots & \cos(\Theta_i + \Theta_j) \end{bmatrix} \tag{5.4}$$

where Θ_i and Θ_j represent the angular values corresponding to two distinct data points. The GASF Matrix highlights the aggregate patterns and interactions among data points, offering a perspective on their collective temporal development.

The GADF Matrix, another integral part of the GAF transformation, also stems from the Gram Matrix. It computes the sine of the differences in angular values between data points. Each component of the GADF Matrix is computed as:

$$\text{GADF}(x_i, x_j) = \cos(\Theta_i - \Theta_j) \tag{5.5}$$

$$\text{GADF} = \begin{bmatrix} \sin(\Theta_1 - \Theta_1) & \sin(\Theta_1 - \Theta_2) & \dots & \sin(\Theta_1 - \Theta_j) \\ \sin(\Theta_2 - \Theta_1) & \sin(\Theta_2 - \Theta_2) & \dots & \sin(\Theta_2 - \Theta_j) \\ \vdots & \vdots & \ddots & \vdots \\ \sin(\Theta_i - \Theta_1) & \sin(\Theta_i - \Theta_2) & \dots & \sin(\Theta_i - \Theta_j) \end{bmatrix} \tag{5.6}$$

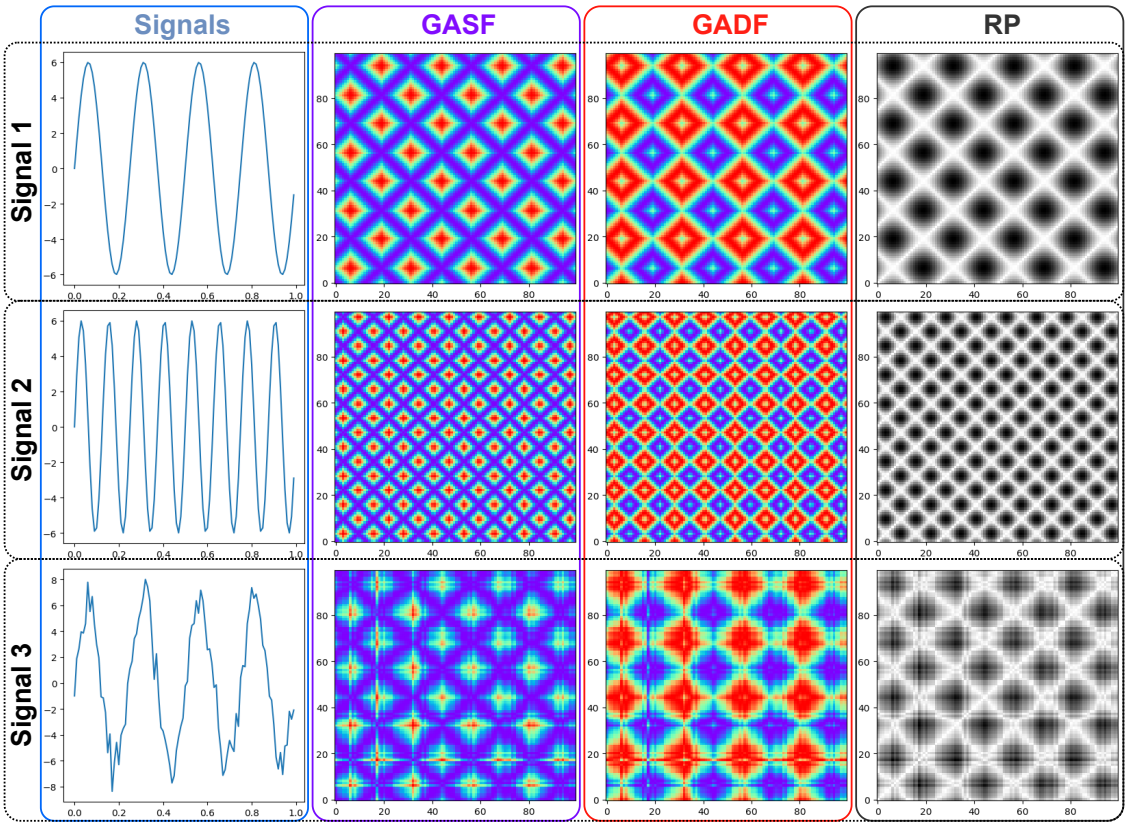

Figure 5.5. Investigating transformations in signals: The methods of GASF, GADF, and RP image encoding reveal distinct patterns in various sinusoidal signals. Signal 1 consists of a pure sinusoidal wave with an amplitude of 6 and a frequency of 4 Hz. Signal 2 is another sinusoidal wave, identical in amplitude at 6 but with a higher frequency of 8 Hz. Signal 3 resembles Signal 1 but includes random noise, resulting in a sinusoidal wave with an amplitude of 6 and a frequency of 4 Hz.

5.2.2.2. Recurrence Plot

The Recurrence Plot (RP), developed by Eckmann and colleagues in 1987, serves as a tool for visualizing and analyzing recurring patterns and structures in dynamic systems and time series data (Eckmann, Kamphorst, and Ruelle, 1987). It has become extensively utilized, especially in the analysis of complex systems and the identification of specific behaviors. The underlying mathematical principle of the RP is based on assessing the similarity or closeness between two points in time. If the distance between two points is less than a predetermined threshold, they are deemed to be close. Mathematically, the distance $d(i, j)$ between two time points i and j is determined as follows:

$$d(i, j) = ||x(i) - x(j)|| \tag{5.7}$$

Here, $x(i)$ and $x(j)$ represent the values of time series data at indices i and j, respectively.

This procedure is applied to every data point within the time series, culminating in the creation of a matrix. This matrix is fundamental to the RP. Each element in the i-th row and j-th column of the matrix indicates the distance between the time points i and j. If this distance is below a predetermined threshold (ϵ), the matrix element is marked as 1; if not, it is marked as 0. The mathematical formulation of the RP is described as follows:

$$R_{i,j} = \begin{cases} 1 & \text{if } d(i, j) < \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{5.8}$$

Here, $R_{i,j}$ is an element of the matrix that indicates the similarity between the i-th and j-th time points. ϵ serves as a threshold, determining how close two points in time need to be. This matrix constitutes the RP.

5.2.3. Transfer Learning

In this study, transfer learning methods were used to analyze Phase-OTDR data converted into images. Transfer learning involves adapting pre-trained deep learning models on large datasets for a new task. The models used in this study are DenseNet121, EfficientNetB0, InceptionResNetV2, InceptionV3, MobileNet, ResNet50, VGG16, and VGG19. These models were utilized in two different ways:

1. **Fine-Tuning**: All layers of the models were retrained. This method allowed the models to better learn the features specific to Phase-OTDR data. Fine-tuned models achieved higher accuracy in classifying Phase-OTDR data. During this process, all layers of the model were unfrozen and retrained to better capture the features specific to Phase-OTDR data.

2. **Feature Extractor**: The pre-trained layers of the models were frozen, and only the final layers were retrained for the classification task specific to Phase-OTDR data. This method is faster and requires less computational power. Models used as feature extractors retained the general features of the Phase-OTDR data while being

adapted for the classification task.

These methods and models were compared to determine which was more effective in classifying Phase-OTDR data converted into images.

5.2.3.1. Validation Techniques

In this study, two validation techniques were employed to evaluate the performance of models on the image-based Phase-OTDR dataset: holdout validation and 5-fold cross-validation.

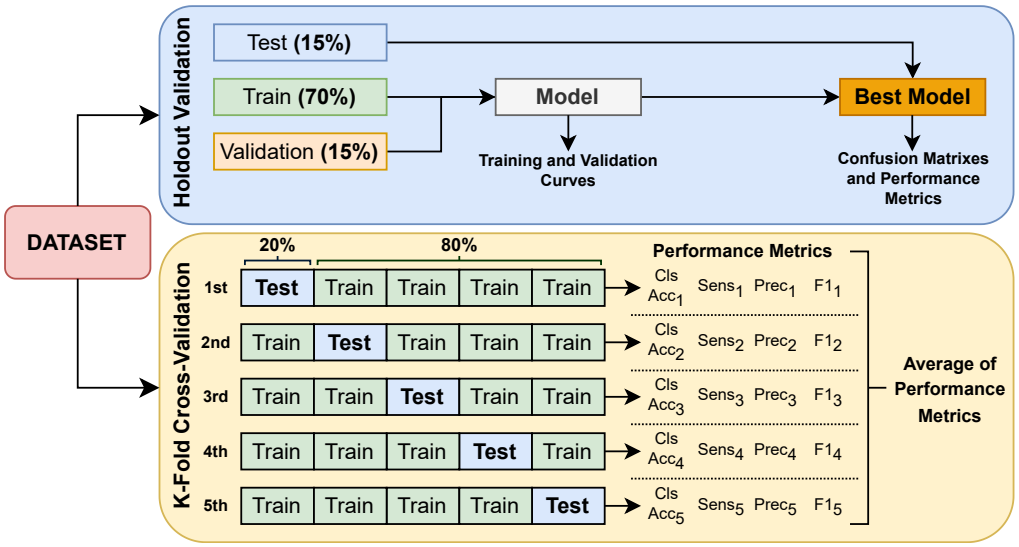

Figure 5.6. The model evaluation procedures are illustrated using two distinct methods. The first method employs the holdout validation technique, which splits the dataset into training, test, and validation sets. The second method utilizes a 5-fold cross-validation technique to assess and validate the models.

Holdout Validation divides the dataset into three subsets: a training set (70%), a test set (15%), and a validation set (15%). This method ensures a significant portion for model learning in the training set, while the test set evaluates the model's performance on unseen instances. The validation set is used to tune the model's hyperparameters and prevent overfitting. The model with the minimum validation loss is saved as the best model, and its performance is assessed using confusion matrices and various performance metrics.

5-Fold Cross-Validation splits the dataset into five equal parts or folds, allocating 80% for training and 20% for testing in each fold. By iteratively using one fold as the test set and the remaining folds for training, this method aims for a more comprehensive assessment. Its goal is to enhance the model's generalization ability by learning from diverse data subsets and averaging performance metrics across iterations. This approach provides a robust evaluation by ensuring that each data point is used for both training and testing.

5.3. Results and Discussion

The ability to monitor phase changes along an optical fiber as a function of position and time opens up numerous possibilities for event detection using Phase-OTDR. However, interpreting the phase trace is challenging due to noise sources and the similarity in strain characteristics experienced by the fiber during different events. Consequently, significant efforts have been made to recognize event types from phase traces. Traditionally, widely used signal processing techniques such as wavelet transform, Fourier analysis, and statistical methods played a key role in this recognition process. However, recent advancements have shifted the focus towards machine learning methods, which leverage data-driven approaches to effectively discern complex patterns, marking a significant evolution in event detection methodologies.

The novelty of this approach lies in transforming 1D Phase-OTDR data into images, presenting a different method from common Phase-OTDR data analysis techniques. This innovative method uses mathematical transformations, including Gramian Angular Difference Field (GADF), Gramian Angular Summation Field (GASF), and Recurrence Plot (RP), to convert raw data into visually interpretable images. This transformation simplifies the complex process of analyzing 1D data and offers a unique way to gain profound insights into monitored events. By converting 1D data into images using three different methods, a multi-channel RGB image is created, with each channel representing a different method. This approach allows for a deeper and more comprehensive analysis by adding an innovative layer. Extracting meaningful features from the RGB image data via transfer learning model architectures has enhanced the classification accuracy of fiber optic sensor data. In summary, the key innovation in this approach is its image-based analysis, which may positively impact Phase-OTDR data analysis, providing improved capabilities and opening new avenues for research and application.

The results presented in Figure 5.7 are obtained through transfer learning with trainable layers set to "True" and utilizing holdout validation. The graph illustrates

the training and validation outcomes of three transfer learning models: DenseNet121, EfficientNetB0, and MobileNet. The initial accuracy of the DenseNet121 model during training started at 58.94%, reaching a maximum value of 99.54%. Validation accuracy began at 79.96% and peaked at 98.72%. The EfficientNetB0 model commenced its training with an initial accuracy of 66.87%, eventually reaching a peak accuracy of 99.61%. Initially, the validation accuracy stood at 93.53% and later reached a maximum of 99.31%. The MobileNet model initiated its training with an initial accuracy of 69.06%, culminating in a peak accuracy of 99.58%. Initially, the validation accuracy was recorded at 92.51% and later reached a maximum of 98.68%.

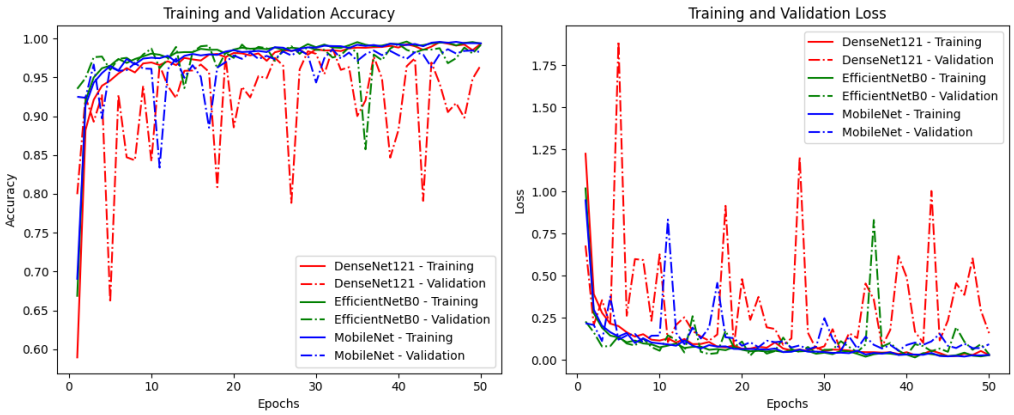

Figure 5.7. The figure displays the training and validation results for the transfer learning fine-tuning model, where the trainable layers were set to "True" for each model.

In Figure 5.8, the results were obtained under the condition of trainable layers set to "False". The DenseNet121 model commenced its training phase with an initial training accuracy of 57.10%, eventually achieving a maximum accuracy of 93.16%. In terms of validation accuracy, it started at 81.24% and reached a peak of 94.48%. EfficientNetB0 embarked on its training with an initial training accuracy of 58.85%, reaching an impressive maximum accuracy of 96.58%. Throughout the validation, it started at 82.14% and achieved a peak accuracy of 95.59%. The MobileNet model initiated its training phase with an initial training accuracy of 52.49%, achieving an impressive maximum accuracy of 97.71%. During validation, it started at 77.04% and reached a peak accuracy of 88.91%.

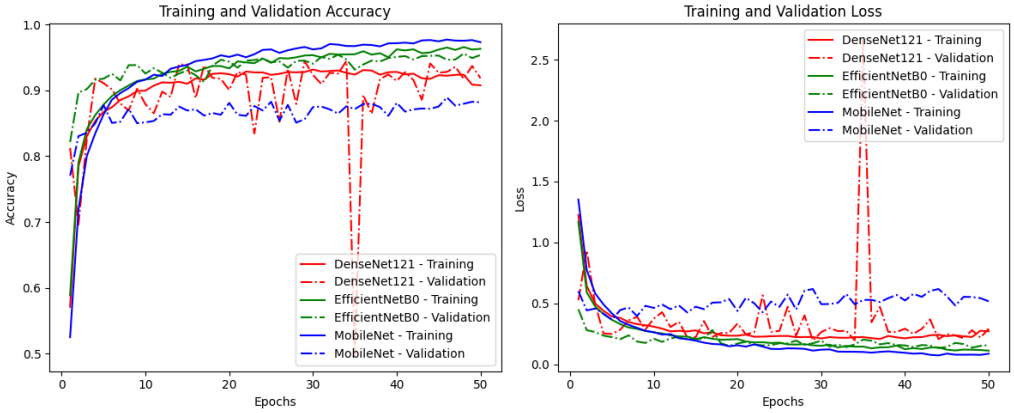

Figure 5.8. The figure illustrates the results for the transfer learning model used as a feature extractor with trainable layers set to "False".

The matrices shown in Figure 5.9 and 5.10 display the results derived from three selected models used to evaluate the holdout validation method on the test dataset. Experiments involving adjustments to the trainable parameter between true and false consistently demonstrated superior performance when the parameter was set to true. This underscores the effectiveness of our approach in converting 1D data into image format, facilitating seamless integration within transfer learning models. Furthermore, our comparative analysis distinctly highlighted the superiority of fine-tuning, indicated by "Trainable: True", compared to the feature extractor mode, denoted by "Trainable: False".

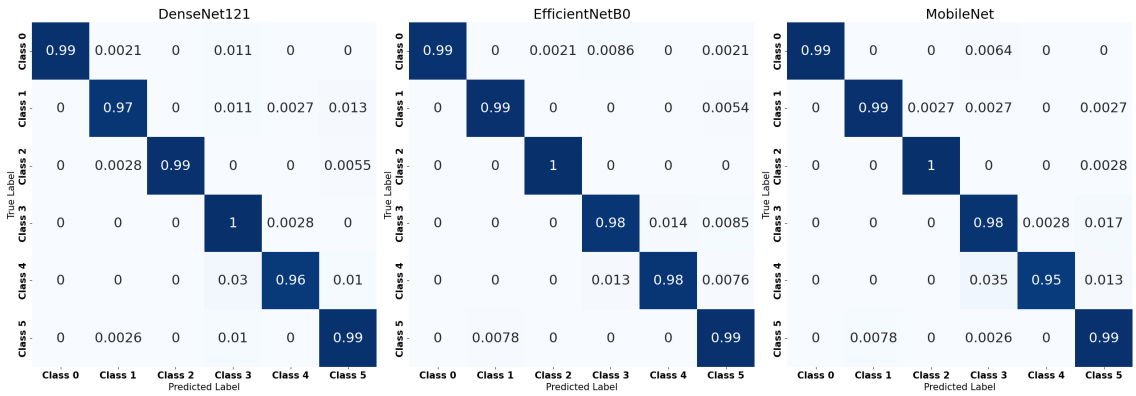

Figure 5.9. The figure demonstrates the classification performance of the selected DenseNet121, EfficientNetB0, and MobileNet models under "Trainable: True" setting.

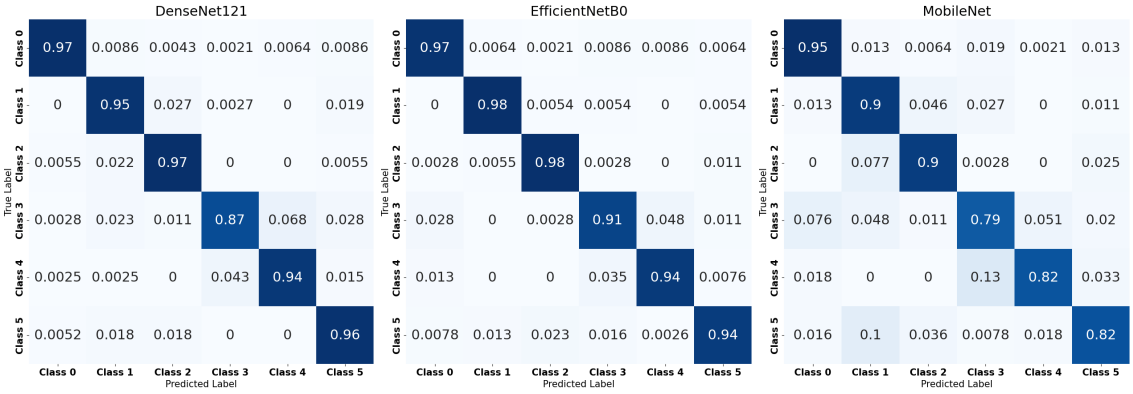

Figure 5.10. The figure demonstrates the classification performance of the selected DenseNet121, EfficientNetB0, and MobileNet models under "Trainable: False" setting.

Fine-tuning allowed for dynamic adjustments in specific weights or layers, especially within transfer learning, leading to better alignment with the dataset and more efficient feature extraction. In contrast, the feature extractor mode primarily relied on pre-trained layers without further refinement, thereby limiting its adaptability to specific datasets and tasks while emphasizing the use of predetermined features.

Table 5.2. Comparison of performance metrics for different models with varying trainable parameters using holdout validation.

Model	**Fine Tuning**				**Feature Extractor**			
	Cls Acc	***Sens***	***Prec***	***F1***	***Cls Acc***	***Sens***	***Prec***	***F1***
DenseNet121	0.9824	0.9826	0.9820	0.9821	0.9435	0.9418	0.9421	0.9416
EfficientNetB0	0.9884	0.9885	0.9880	0.9882	0.9542	0.9535	0.9537	0.9536
InceptionResNetV2	0.9833	0.9835	0.9831	0.9830	0.4788	0.4925	0.5432	0.4709
InceptionV3	0.9790	0.9785	0.9788	0.9786	0.8904	0.8890	0.8927	0.8898
MobileNet	0.9842	0.9841	0.9836	0.9837	0.8664	0.8633	0.8665	0.8634
ResNet50	0.9709	0.9711	0.9700	0.9705	0.9666	0.9665	0.9660	0.9661
VGG16	0.9722	0.9718	0.9713	0.9715	0.9456	0.9440	0.9456	0.9442
VGG19	0.9533	0.9529	0.9521	0.9521	0.9289	0.9287	0.9283	0.9281

Based on the analysis shown in Table 5.2, the performance of deep learning models in the context of image classification was evaluated using holdout validation with varying configurations of trainable parameters. When the trainable parameters were enabled ("True"), models such as DenseNet121, EfficientNetB0, InceptionResNetV2, InceptionV3, MobileNet, and ResNet50 consistently achieved exceptional classification

accuracies, surpassing 97%, with EfficientNetB0 reaching an impressive accuracy exceeding 99%. However, when the trainable parameters were disabled ("False"), a noticeable decline in performance was observed. In particular, InceptionResNetV2 exhibited a significant reduction in classification accuracy, while InceptionV3, MobileNet, and VGG19 also experienced performance degradation. The notable performance difference between the configurations can be attributed to the models' adaptability to new data, which is prominently demonstrated during fine-tuning procedures.

Using t-SNE, we visualized feature representations generated by three deep learning models, as shown in Figure 5.11. This analysis focuses on six distinct classes: background, digging, knocking, watering, shaking, and walking, revealing how these models differentiate the training dataset as the learning process concludes. Our findings demonstrate that both trainable parameter configurations result in noticeable class separations. However, comparative analysis indicates that the "True" configuration offers superior discriminative power, highlighting its effectiveness in capturing and distinguishing underlying data patterns. This emphasizes the significance of parameter settings in optimizing deep learning models for specific tasks, with the "True" configuration exhibiting exceptional separation capabilities.

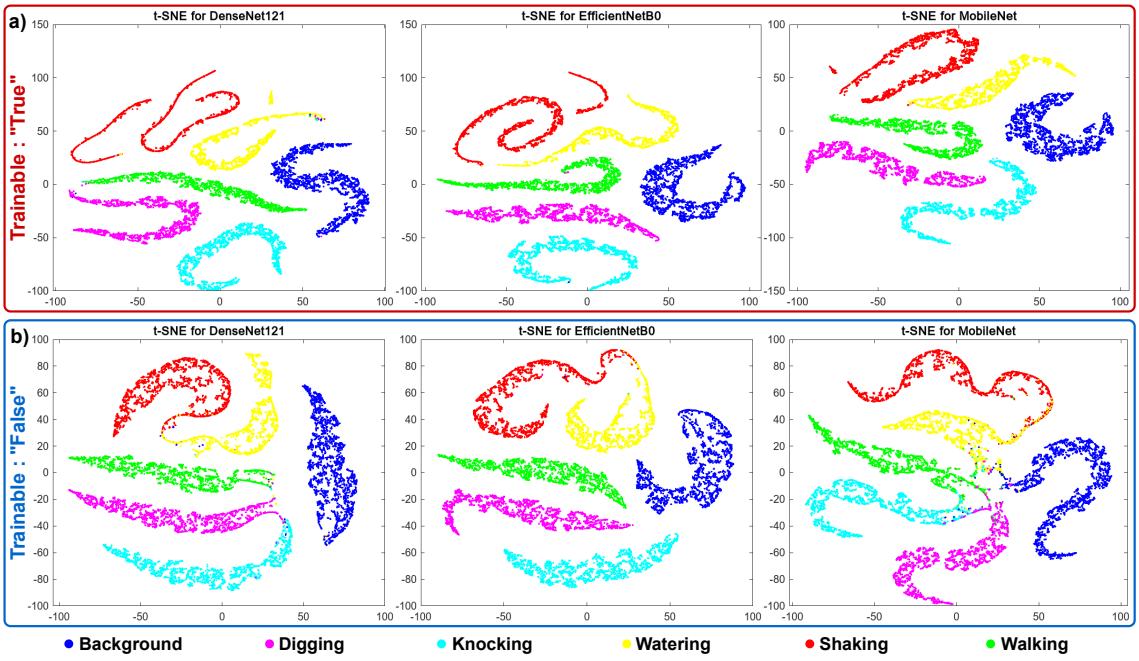

Figure 5.11. t-SNE visualizations illustrate the feature space separation for DenseNet121, EfficientNetB0, and MobileNet models under different trainable parameter settings. The plots show the clustering of training data as interpreted by each trained model, with the x and y axes representing the distribution of data points in a two-dimensional space that approximates their high-dimensional relationships.

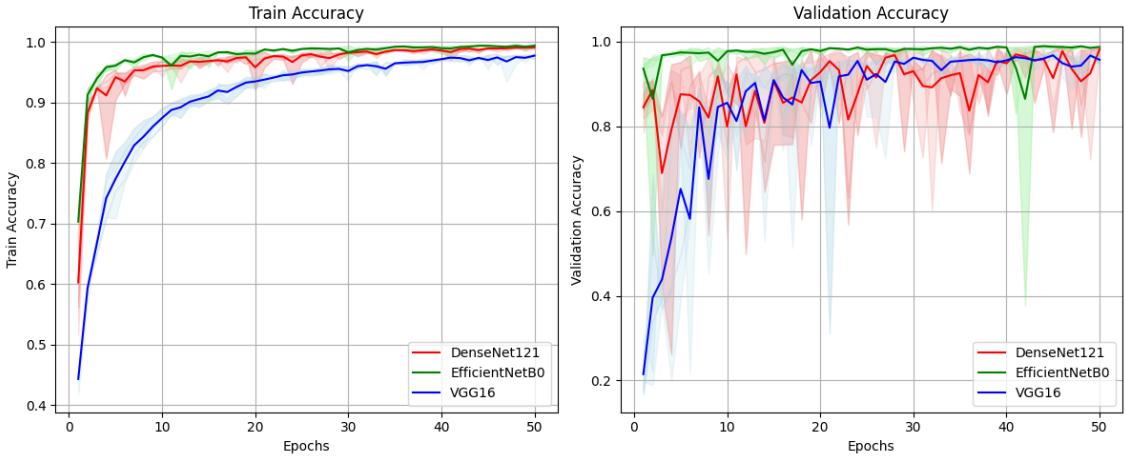

Figure 5.12. Depicts the 5-fold cross-validation analysis showing training and validation curves for three selected models with trainable parameters set to true. The colored lines indicate the mean performance metrics across folds, while the shaded areas emphasize the variations.

In the 5-fold cross-validation analysis, as shown in Figure 5.12 and 5.13, the training accuracy curve is displayed on the left, while the validation curve for the test dataset is on the right. The graphical representation tracks the performance of three distinct models, each identified by unique color coding. To enhance the interpretability of our results, we applied color-based background shading to highlight variations across different folds.

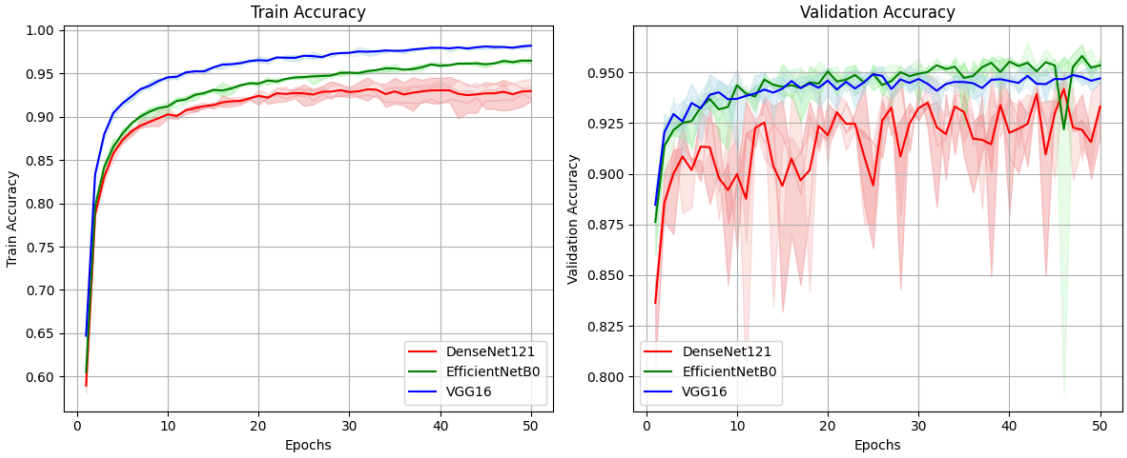

Figure 5.13. Depicts the 5-fold cross-validation analysis showing training and validation curves for three selected models with trainable parameters set to false. The colored lines indicate the mean performance metrics across folds, while the shaded areas emphasize the variations.

The solid lines in the chart represent the mean performance metrics across cross-validation folds for each model. Notably, during this analysis with "Trainable: True" parameters in a 5-fold cross-validation framework, consistently higher accuracy values were achieved.

The results shown in Table 5.3 are derived from a detailed 5-fold cross-validation process, thoroughly assessing the performance of three selected models. This comprehensive evaluation examines model performances across different trainable parameter settings. The "Mean Acc" column displays the average performance metrics over the 5-fold cross-validation. Notably, DenseNet121, EfficientNetB0, and VGG16 consistently exhibit strong performance across these cross-validation folds. In the "Trainable: True" configuration, these models achieve high classification accuracy: DenseNet121 with a mean accuracy of approximately 98.68%, EfficientNetB0 with around 99.07%, and VGG16 with approximately 96.86%. These metrics consistently surpass the 97% benchmark, indicating the reliability of these models in maintaining high classification accuracy. Additionally, the standard deviation values in the "Trainable: True" configuration are relatively low, reflecting the models' stability and minimal variance in their performance. Conversely, when the trainable parameters are set to "False", the models show slightly lower classification accuracy and marginally higher variance. For example, DenseNet121 and EfficientNetB0, while still performing well, experience a reduction in classification accuracy to the range of 94.53% to 95.87%, and an increase in standard deviation, indicating reduced adaptability and a greater reliance on pre-trained features.

Table 5.3. Comparison of performance metrics for three selected models with different trainable parameter settings using 5-fold cross-validation.

Model	**Statistic Type**	**Fine Tuning**				**Feature Extractor**			
		Cls Acc	***Sens***	***Prec***	***F1***	***Cls Acc***	***Sens***	***Prec***	***F1***
DenseNet121	Mean Acc	0.9868	0.9863	0.9863	0.9862	0.9453	0.9434	0.9444	0.9435
	Standard Deviation	0.0012	0.0016	0.0012	0.0014	0.0051	0.0058	0.0054	0.0058
EfficientNetB0	Mean Acc	0.9907	0.9903	0.9903	0.9903	0.9587	0.9579	0.9578	0.9577
	Standard Deviation	0.0024	0.0026	0.0023	0.0025	0.0032	0.0031	0.0031	0.0031
VGG16	Mean Acc	0.9686	0.9671	0.9675	0.9672	0.9478	0.9460	0.9467	0.9462
	Standard Deviation	0.0063	0.0066	0.0066	0.0066	0.0027	0.0025	0.0026	0.0024

Our study presents a novel image-based approach for analyzing Phase-OTDR data for event detection. The multi-channel RGB images, created using mathematical transformations such as GADF, GASF, and RP, significantly enhance the depth of analysis. In the fine-tuning scenario (with trainable layers set to "True"), holdout validation results

show that EfficientNetB0 and DenseNet121 achieved classification accuracies of approximately 98.8% and 98.2%, respectively. Similarly, in the 5-fold cross-validation process, these models demonstrated high accuracy rates of around 99.1% and 98.7%, respectively. These results highlight the consistency and generalizability of our model, representing a significant advancement in Phase-OTDR data analysis and offering substantial potential across various application domains.

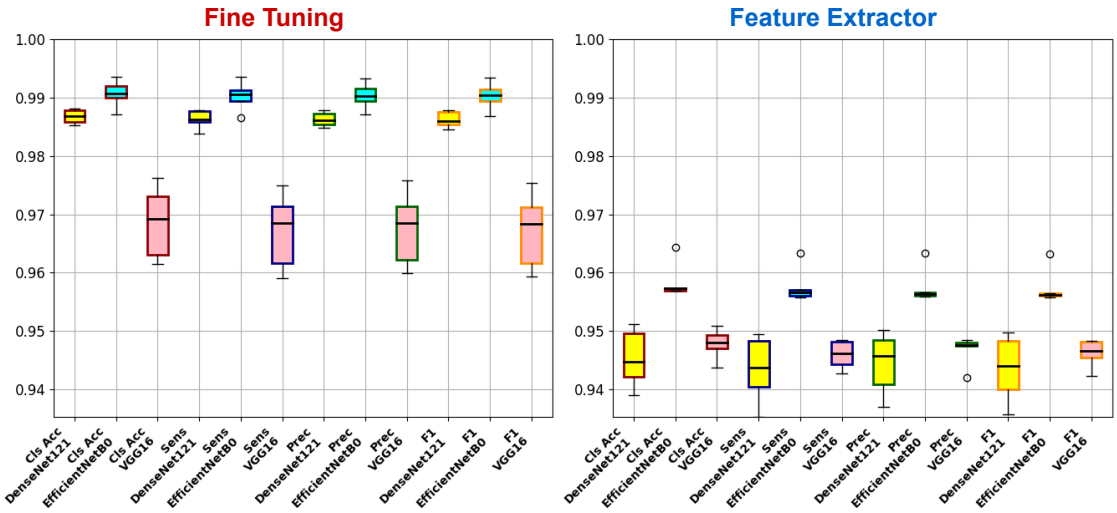

Figure 5.14. Comparative boxplots illustrate the performance of transfer learning models with 'Trainable: True' (left) and 'Trainable: False' (right) settings over 5-fold cross-validation. These visualizations compare key performance metrics, including accuracy (Acc), sensitivity (Sens), precision (Prec), and F1 scores, highlighting the variability and central tendencies in each configuration.

5.4. Conclusion

In conclusion, this pioneering study demonstrates the transformative potential of image-based classification techniques in analyzing Phase-OTDR data. By employing a publicly available dataset and innovative data transformation methods, the research showcases the effectiveness of converting raw data into grayscale images and multi-channel RGB representations. The study utilized EfficientNetB0 and DenseNet121 models, which exhibited outstanding performance in classifying various disturbance events along optical fibers. The EfficientNetB0 model achieved an impressive accuracy of 98.8% in holdout validation and 99.1% in 5-fold cross-validation, while the DenseNet121 model attained accuracies of 98.2% and 98.7% in holdout validation and 5-fold cross-validation, respec-

tively. These high accuracies underscore the reliability and robustness of the image-based classification approach in accurately identifying and categorizing Phase-OTDR events.

Moreover, the study highlights the significance of model adaptability in enhancing classification performance. The models that were fine-tuned with trainable parameters set to "True" consistently outperformed the feature extractor models with trainable parameters set to "False". This finding emphasizes the importance of allowing the models to adapt and learn from the specific characteristics of the Phase-OTDR data, leading to improved classification accuracy.

In addition to the impressive classification results, the methodology also demonstrates its efficacy in data storage and analysis. The transformation of raw data into grayscale images and multi-channel RGB representations resulted in a significant reduction in dataset size from 2.03 GB to 180 MB. This compression not only facilitates more efficient data storage but also enables faster processing and analysis of the Phase-OTDR data.

This groundbreaking research paves the way for future advancements in fiber optic sensing data analysis and has far-reaching implications for domains relying on accurate fiber optic monitoring systems. The successful application of image-based classification techniques in this study serves as a foundation for further exploration and practical applications in the field of Phase-OTDR data analysis. As such, this research represents a significant step forward in understanding and leveraging the potential of image-based methods in extracting valuable insights from complex fiber optic sensing data.

CHAPTER 6

CONCLUSION

In this thesis, we have explored the application of deep learning (DL) techniques across various domains, demonstrating significant improvements in the detection, classification, and analysis of complex data. The integration of DL models has showcased their versatility and robustness, providing enhanced solutions to traditional analytical challenges. This work emphasizes the integrity of the thesis by highlighting the opportunity to deal with different data from diverse fields and perform analyses using various techniques.

In Chapter 2, we focused on electrochemical analysis, developing a DL-based approach for the detection and classification of CD36 using an immuno-biosensor. Traditional techniques often fall short in sensitivity and rapid analysis, particularly at low analyte concentrations. By integrating DL models, specifically 1D-CNN and a hybrid 1D-CNN – LSTM network, we significantly enhanced the sensitivity and specificity of the biosensor. The hybrid model outperformed the standalone 1D-CNN model, particularly in classifying lower analyte concentrations, achieving an accuracy of 95.93%, precision of 97.01%, recall of 95.89%, and F1-score of 95.73% for the CV method. For the DPV method, the hybrid model achieved an accuracy of 99.23%, precision of 99.41%, recall of 99.21%, and F1-score of 99.28%. This integration of DL techniques with electrochemical immuno-biosensors presents a powerful approach for sensitive and accurate detection, with significant potential for point-of-care diagnostics and real-time data processing.

In Chapter 3, we explored the biomedical domain, integrating surface electromyography (sEMG) with advanced time-frequency analysis (TFA) and various Vision Transformer (ViT) models for hand gesture classification. By combining sEMG data with TFA methods such as STFT and CWT, and utilizing ViT models including base ViT, Swin Transformer, and MaxViT, we achieved near-perfect accuracy in gesture recognition. The MaxViT model, in particular, exhibited the highest performance, achieving an accuracy of 98.40% with a 500 ms window length for both STFT and CWT methods. This study underscores the potential of combining sEMG with advanced TFA and ViT models to develop intuitive and responsive control systems for applications such as prosthetic devices and human-computer interaction systems.

In Chapter 4, we introduced DL methods to improve the efficiency of Phase-OTDR-based current sensing systems. We applied 1D-CNN, 1D-CNN – LSTM, and 1D-CNN – Bi-LSTM models to classify current values based on Phase-OTDR measurements. The results highlighted the effectiveness of these models in accurately distinguishing between

current levels, even in the presence of noise. The study demonstrated high accuracy rates, achieving 100% accuracy with a 20 A buffer zone between current levels, and maintaining high accuracy rates of 97% and 89% for current levels of 15 A and 10 A, respectively. This research presents an opportunity to enhance the performance of Phase-OTDR-based current sensing systems and explore new avenues for improving contactless electrical current measurements.

In Chapter 5, we demonstrated the transformative potential of image-based classification techniques in analyzing Phase-OTDR data. By converting raw data into grayscale images and multi-channel RGB representations, and utilizing EfficientNetB0 and DenseNet121 models, we achieved outstanding performance in classifying various disturbance events along optical fibers. The EfficientNetB0 model achieved an impressive accuracy of 98.8% in holdout validation and 99.1% in 5-fold cross-validation, while the DenseNet121 model attained accuracies of 98.2% and 98.7% in holdout validation and 5-fold cross-validation, respectively. The image-based classification approach not only improved classification accuracy but also facilitated more efficient data storage and analysis, reducing the dataset size from 2.03 GB to 180 MB. This methodology paves the way for future advancements in fiber optic sensing data analysis and has far-reaching implications for domains relying on accurate fiber optic monitoring systems.

Overall, this thesis highlights the significant advancements that can be achieved by integrating deep learning techniques with various analytical methods across different domains. The studies presented here demonstrate the potential of DL to enhance the performance of traditional techniques, providing more accurate, sensitive, and efficient solutions. By working with diverse datasets and employing different DL models, we have shown the versatility and robustness of these techniques in addressing complex analytical challenges. The ability to work with different types of data and apply various DL techniques has provided valuable insights and opened new possibilities for future research. The methodologies developed in this thesis can be extended to other biomarkers, signal types, and analytical challenges, further refining the models to improve their robustness and generalizability. Future work could explore the application of these approaches to other domains, enhancing the capabilities of DL in real-world applications.

In conclusion, the findings of this thesis contribute to the growing body of knowledge in DL applications, offering new perspectives and solutions for complex analytical problems. The continued development and refinement of these techniques hold significant promise for future advancements in various fields, ultimately enhancing the quality of life and technological progress.

REFERENCES

Adib, Edmonmd, Amanda S. Fernandez, Fatemeh Afghah, and John J. Prevost. 2023. "Synthetic ECG Signal Generation Using Probabilistic Diffusion Models." *IEEE Access* 11:75818–75828. ISSN: 2169-3536. https://doi.org/10.1109/ACCESS.2023.3296542.

Aerssens, M., A. Gusarov, P. Moreau, P. Malard, V. Massaut, P. Mégret, and M. Wuilpart. 2012. "Development of a Jones vector based model for the measurement of a plasma current in a thermonuclear fusion reactor with a POTDR setup." In *Optical Sensing and Detection II,* edited by Francis Berghmans, Anna G. Mignani, and Piet De Moor, 8439:84390D. SPIE, April. https://doi.org/10.1117/12.922714. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.922714.

Ahmad, Zeeshan, Anika Tabassum, Ling Guan, and Naimul Mefraz Khan. 2021. "ECG Heartbeat Classification Using Multimodal Fusion." *IEEE Access* 9:100615–100626. ISSN: 2169-3536. https://doi.org/10.1109/ACCESS.2021.3097614.

Ahmad, Zishan, Raghav Jindal, Mukuntha N.S., Asif Ekbal, and Pushpak Bhattacharyya. 2022. "Multi-modality helps in crisis management: An attention-based deep learning approach of leveraging text for image classification." *Expert Systems with Applications* 195:116626. ISSN: 0957-4174. https://doi.org/https://doi.org/10.1016/j.eswa.2022.116626. https://www.sciencedirect.com/science/article/pii/S0957417422001166.

Alibabaei, Khadijeh, Pedro D. Gaspar, Tânia M. Lima, Rebeca M. Campos, Inês Girão, Jorge Monteiro, and Carlos M. Lopes. 2022. "A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities." *Remote Sensing* 14 (3): 638. ISSN: 2072-4292. https://doi.org/10.3390/rs14030638. https://www.mdpi.com/2072-4292/14/3/638.

Alzubaidi, Laith, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. 2021. "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." *Journal of Big Data* 8 (1): 53. ISSN: 2196-1115. https://doi.org/10.1186/s40537-021-00444-8. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8.

Armbruster, D., and T. Pry. 2008. "Limit of blank, limit of detection and limit of quantitation." *The Clinical biochemist. Reviews* 29 Suppl 1:S49–52.

Atila, Ümit, and Furkan Sabaz. 2022. "Turkish lip-reading using Bi-LSTM and deep learning models." *Engineering Science and Technology, an International Journal* 35 (November): 101206. ISSN: 22150986. https://doi.org/10.1016/j.jestch.2022.101206. https://linkinghub.elsevier.com/retrieve/pii/S221509862200115X.

Bahadır, E., and M. Sezgintürk. 2015. "Applications of electrochemical immunosensors for early clinical diagnostics." *Talanta* 132:162–174. https://doi.org/10.1016/J.TALANTA.2014.08.063.

Bai, Yu-Xin, Ting-Ting Lin, and Zhi-Cheng Zhong. 2021. "Noise Reduction Method of Φ-OTDR System Based on EMD-TFPF Algorithm." *IEEE Sensors Journal* 21 (21): 24084–24089. ISSN: 1530-437X. https://doi.org/10.1109/JSEN.2021.3107039. https://ieeexplore.ieee.org/document/9521223/.

Baranova, I., R. Kurlander, A. Bocharov, T. Vishnyakova, Zhigang Chen, A. Remaley, G. Csako, A. Patterson, and T. Eggerman. 2008. "Role of Human CD36 in Bacterial Recognition, Phagocytosis, and Pathogen-Induced JNK-Mediated Signaling1." *The Journal of Immunology* 181:7147–7156. https://doi.org/10.4049/jimmunol.181.10.7147.

Barantsov, Ivan A., Alexey B. Pnev, Kirill I. Koshelev, Vadim S. Tynchenko, Vladimir A. Nelyub, and Aleksey S. Borodulin. 2023. "Classification of Acoustic Influences Reg-

istered with Phase-Sensitive OTDR Using Pattern Recognition Methods." *Sensors* 23 (2). ISSN: 1424-8220. https://doi.org/10.3390/s23020582. https://www.mdpi.com/1424-8220/23/2/582.

Bashford, J., Kerry R. Mills, and Chris Shaw. 2019. "The evolving role of surface electromyography in amyotrophic lateral sclerosis: A systematic review." *Clinical Neurophysiology* 131:942–950. https://doi.org/10.1016/j.clinph.2019.12.007.

Batista, Antonio, Xianbin Sun, Mohan Wang, Bo Zhan, Yuanyuan Xiong, and Wei Yu. 2023. "An Intelligent Diagnostic Method for Multisource Coupling Faults of Complex Mechanical Systems." *Shock and Vibration* 2023:6928871. ISSN: 1070-9622. https://doi.org/10.1155/2023/6928871. https://doi.org/10.1155/2023/6928871.

Becattini, Federico, Federico Palai, and Alberto Del Bimbo. 2022. "Understanding Human Reactions Looking at Facial Microexpressions With an Event Camera." *IEEE Transactions on Industrial Informatics* 18 (12): 9112–9121. https://doi.org/10.1109/TII.2022.3195063.

Belhadi, Asma, Youcef Djenouri, Gautam Srivastava, Djamel Djenouri, Jerry Chun-Wei Lin, and Giancarlo Fortino. 2021. "Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection." *Information Fusion* 65 (January): 13–20. ISSN: 15662535. https://doi.org/10.1016/j.inffus.2020.08.003. https://linkinghub.elsevier.com/retrieve/pii/S1566253520303316.

Beniczky, Sándor, Philippa Karoly, Ewan Nurse, Philippe Ryvlin, and Mark Cook. 2021. "Machine learning and wearable devices of the future." *Epilepsia* 62 (S2): S116–S124. https://doi.org/https://doi.org/10.1111/epi.16555. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/epi.16555. https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.16555.

Bonet-Solà, Daniel, Ester Vidaña-Vila, and Rosa Ma Alsina-Pagès. 2023. "Analysis and Acoustic Event Classification of Environmental Data Collected in a Citizen Science

Project." *International Journal of Environmental Research and Public Health* 20 (4). ISSN: 1660-4601. https://doi.org/10.3390/ijerph20043683. https://www.mdpi.com/1660-4601/20/4/3683.

Bora, Merve, Ali Yalçin, Numan Bulut, Öznur Yilmaz, Ayşe Karaduman, Semra Topuz, and İpek Alemdaroğlu-Gürbüz. 2021. "Investigation of surface electromyography amplitude values during stair climbing task in children with Duchenne muscular dystrophy." *Neurological Sciences,* https://doi.org/10.1007/s10072-021-05643-y.

Buelvas, Hebert Elias Palmera, Juan Diego Trujillo Montaña, and Ruthber Rodriguez Serrezuela. 2023. "Hand Gesture Classification using Deep learning and CWT images based on multi-channel surface EMG signals." In *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME),* 1–7. https://doi.org/10.1109/ICECCME57830.2023.10253296.

Cao, Xiaomin, Yunsheng Su, Zhiyan Jin, and Kuanglu Yu. 2023. "An open dataset of φ-OTDR events with two classification models as baselines." *Results in Optics* 10:100372. ISSN: 2666-9501. https://doi.org/https://doi.org/10.1016/j.rio.2023.100372. https://www.sciencedirect.com/science/article/pii/S266695012300024X.

Chen, Xiaojuan, Cheng Yang, Haoyu Yu, and Guangwei Hou. 2023. "Research on Pattern Recognition Method for phi;-OTDR System Based on Dendrite Net." *Electronics* 12 (18). ISSN: 2079-9292. https://doi.org/10.3390/electronics12183757. https://www.mdpi.com/2079-9292/12/18/3757.

Chen, Yiqiang, Xin Qin, Jindong Wang, Chaohui Yu, and Wen Gao. 2020. "FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare." *IEEE Intelligent Systems* 35, no. 4 (July): 83–93. ISSN: 1941-1294. https://doi.org/10.1109/MIS.2020.2988604.

Chiou, Shao-Hua, and Wen-Teng Wu. 2004. "Immobilization of Candida rugosa lipase on chitosan with activation of the hydroxyl groups." *Biomaterials* 25 (2): 197–204.

ISSN: 0142-9612. https://doi.org/https://doi.org/10.1016/S0142-9612(03)00482-4. https://www.sciencedirect.com/science/article/pii/S0142961203004824.

Cho, Soo-Yeon, Youhan Lee, Sangwon Lee, Hohyung Kang, Jaehoon Kim, Junghoon Choi, Jin Ryu, Heeeun Joo, Hee-Tae Jung, and Jihan Kim. 2020. "Finding Hidden Signals in Chemical Sensors Using Deep Learning." *Analytical chemistry* 92 (9): 6529–6537. ISSN: 1520-6882. https://doi.org/10.1021/acs.analchem.0c00137. http://www.ncbi.nlm.nih.gov/pubmed/32286053.

Ciobanu, C. A., I. Ionita, A. Mihailescu, and A.-M. Ciobanu. 2023. "Cyclic voltammetry as a measuring tool in Parkinson's disease and associated psychiatric commorbidities." *European Psychiatry* 66:S757–S758. https://doi.org/10.1192/j.eurpsy.2023.1597.

Cruz, J L, M V Andres, and M A Hernandez. 1996. "Faraday effect in standard optical fibers: dispersion of the effective Verdet constant." *Applied optics* 35 (6): 922–7. ISSN: 1559-128X. https://doi.org/10.1364/AO.35.000922. http://www.ncbi.nlm.nih.gov/pubmed/21069090.

Ditria, Ellen M., Sebastian Lopez-Marcano, Michael Sievers, Eric L. Jinks, Christopher J. Brown, and Rod M. Connolly. 2020. "Automating the Analysis of Fish Abundance Using Object Detection: Optimizing Animal Ecology With Deep Learning." *Frontiers in Marine Science* 7 (June): 538334. ISSN: 2296-7745. https://doi.org/10.3389/fmars.2020.00429. https://www.frontiersin.org/article/10.3389/fmars.2020.00429/full.

Dong, Hongwei, Lamei Zhang, and B. Zou. 2022. "Exploring Vision Transformers for Polarimetric SAR Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 60:1–15. https://doi.org/10.1109/TGRS.2021.3137383.

Dzik, W., C. Cserti-Gazdewich, I. Ssewanyana, M. Delelys, and F. Preffer. 2009. "When monocytes and platelets compete: The effect of platelet count on the flow cytometric measurement of monocyte CD36." *Cytometry Part B: Clinical Cytometry* 78B. https://doi.org/10.1002/cyto.b.20504.

Eckmann, J.-P., S. Oliffson Kamphorst, and D. Ruelle. 1987. "Recurrence Plots of Dynamical Systems." *Europhysics Letters* 4, no. 9 (November): 973. https://doi.org/10.1209/0295-5075/4/9/004. https://dx.doi.org/10.1209/0295-5075/4/9/004.

Elbeshbeshy, Ahmed M., Muhammad A. Rushdi, and Shereen M. El-Metwally. 2021. "Electromyography Signal Analysis and Classification using Time-Frequency Representations and Deep Learning." In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC),* 661–664. https://doi.org/10.1109/EMBC46164.2021.9630815.

Electrical Engineers, P. Institution of, Institution of Engineering, and Technology. 1985. *Electronics letters.* 21:226–228. IET Digital Library, March. https://digital-library.theiet.org/content/journals/10.1049/el_19850161.

Er, Simge, and Dilek Odaci Demirkol. 2022. "Graphene oxide incorporated polystyrene electrospun nanofibers for immunosensing of CD36 as a marker of diabetic plasma." *Bioelectrochemistry* 145:108083. ISSN: 1567-5394. https://doi.org/https://doi.org/10.1016/j.bioelechem.2022.108083. https://www.sciencedirect.com/science/article/pii/S1567539422000342.

Fang, Wei, Yupeng Chen, and Qiongying Xue. 2021. "Survey on Research of RNN-Based Spatio-Temporal Sequence Prediction Algorithms." *Journal on Big Data* 3 (3): 97–110. ISSN: 2579-0056. https://doi.org/10.32604/jbd.2021.016993. https://www.techscience.com/jbd/v3n3/45671.

Farsi, Leila, Siuly Siuly, Enamul Kabir, and Hua Wang. 2021. "Classification of Alcoholic EEG Signals Using a Deep Learning Method." *IEEE Sensors Journal* 21 (3): 3552–3560. ISSN: 1530-437X. https://doi.org/10.1109/JSEN.2020.3026830. https://ieeexplore.ieee.org/document/9207939/.

Frosio, Guido, and René Dändliker. 1994. "Reciprocal reflection interferometer for a fiber-optic Faraday current sensor." *Applied Optics* 33 (25): 6111. ISSN: 0003-6935. https:

//doi.org/10.1364/AO.33.006111. https://opg.optica.org/abstract.cfm?URI=ao-33-25-6111.

Garcia, Gabriel Rodriguez, Gabriel Michau, Mélanie Ducoffe, Jayant Sen Gupta, and Olga Fink. 2022. "Temporal signals to images: Monitoring the condition of industrial assets with deep learning image processing algorithms." *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 236 (4): 617–627. https://doi.org/10.1177/1748006X21994446. eprint: https://doi.org/10.1177/1748006X21994446. https://doi.org/10.1177/1748006X21994446.

Giffhorn, F. 2000. "Fungal pyranose oxidases: occurrence, properties and biotechnical applications in carbohydrate chemistry." *Applied Microbiology and Biotechnology* 54:727–740. https://doi.org/10.1007/s002530000446.

Glatz, J., M. Nabben, and J. Luiken. 2022. "CD36 (SR-B2) as master regulator of cellular fatty acid homeostasis." *Current Opinion in Lipidology* 33:103–111. https://doi.org/10.1097/MOL.0000000000000819.

Godoy, Ricardo V., Anany Dwivedi, Bonnie Guan, Amber Turner, Dasha Shieff, and Minas V. Liarokapis. 2022. "On EMG Based Dexterous Robotic Telemanipulation: Assessing Machine Learning Techniques, Feature Extraction Methods, and Shared Control Schemes." *IEEE Access* 10:99661–99674. https://doi.org/10.1109/ACCESS.2022.3206436.

González, J., and J.A. Sequí-Castellano. 2021. "Electrochemical determination of kinetic parameters of surface confined redox probes in presence of intermolecular interactions by means of Cyclic Voltammetry. Application to TEMPO monolayers in gold and platinum electrodes." *Electrochimica Acta* 365:137331. https://doi.org/10.1016/j.electacta.2020.137331.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep learning*. 775. MIT Press. ISBN: 0262035618.

Graves, Alex. 2012. “Long Short-Term Memory,” 37–45. https://doi.org/10.1007/978-3-642-24797-2_4. http://link.springer.com/10.1007/978-3-642-24797-2_4.

He, Fan, Meitao Wang, and Peng Zhou. 2022. “Evaluation of market risk and resource allocation ability of green credit business by deep learning under internet of things.” Edited by Stefan Cristian Gherghina. *PLOS ONE* 17 (4): e0266674. ISSN: 1932-6203. https://doi.org/10.1371/journal.pone.0266674. https://dx.plos.org/10.1371/journal.pone.0266674.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. “Deep Residual Learning for Image Recognition.” In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 770–778. https://doi.org/10.1109/CVPR.2016.90.

Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.” *arXiv e-prints* (April): arXiv:1704.04861. https://doi.org/10.48550/arXiv.1704.04861. arXiv: 1704.04861 [cs.CV].

Huan, Sha, Limei Wu, Man Zhang, Zhaoyue Wang, and Chao Yang. 2023. “Radar Human Activity Recognition with an Attention-Based Deep Learning Network.” *Sensors* 23 (6): 3185. ISSN: 1424-8220. https://doi.org/10.3390/s23063185. https://www.mdpi.com/1424-8220/23/6/3185.

Huang, Gao, Zhuang Liu, Geoff Pleiss, Laurens van der Maaten, and Kilian Q. Weinberger. 2022. “Convolutional Networks with Dense Connectivity.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, no. 12 (December): 8704–8716. ISSN: 1939-3539. https://doi.org/10.1109/TPAMI.2019.2918284.

Huo, Yingzi, Kai Jin, Jiahong Cai, Huixuan Xiong, and Jiacheng Pang. 2023. “Vision Transformer (ViT)-based Applications in Image Classification.” *2023 IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference*

on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), 135–140. https://doi.org/10.1109/BigDataSecurity-HPSC-IDS58521.2023.00033.

Ibrahim, A. F. T., V. R. Gannapathy, L. Chong, and Ida Syafiza M. Isa. 2016. "Analysis of Electromyography (EMG) Signal for Human Arm Muscle: A Review," 567–575. https://doi.org/10.1007/978-3-319-24584-3_49.

Ionescu, Radu Tudor, Fahad Shahbaz Khan, Mariana-Iuliana Georgescu, and Ling Shao. 2019. "Object-centric auto-encoders and dummy anomalies for abnormal event detection in video." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,* 7842–7851.

Jelodar, Hamed, Yongli Wang, Rita Orji, and Shucheng Huang. 2020. "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach." *IEEE journal of biomedical and health informatics* 24 (10): 2733–2742. ISSN: 2168-2208. https://doi.org/10.1109/JBHI.2020.3001216. http://www.ncbi.nlm.nih.gov/pubmed/32750931.

Kandamali, Deus F., Xiaomin Cao, Manling Tian, Zhiyan Jin, Hui Dong, and Kuanglu Yu. 2022. "Machine learning methods for identification and classification of events in Φ-OTDR systems: a review." *Applied optics* 61 11:2975–2997. https://doi.org/10.1364/ao.444811.

Karimi, Masoud, Mehrdad Majidi, Hamed MirSaeedi, Mohammad Mehdi Arefi, and Mohammad Oskuoee. 2020. "A Novel Application of Deep Belief Networks in Learning Partial Discharge Patterns for Classifying Corona, Surface, and Internal Discharges." *IEEE Transactions on Industrial Electronics* 67 (4): 3277–3287. ISSN: 0278-0046. https://doi.org/10.1109/TIE.2019.2908580. https://ieeexplore.ieee.org/document/8683990/.

Kayali, Devrim, Nemah Abu Shama, Suleyman Asir, and Kamil Dimililer. 2023. "Machine learning-based models for the qualitative classification of potassium ferrocyanide using electrochemical methods." *Journal of Supercomputing* 79 (11): 12472–12491. ISSN: 15730484. https://doi.org/10.1007/S11227-023-05137-Y/TABLES/8. https://link.springer.com/article/10.1007/s11227-023-05137-y.

Kersey, A., and D. Jackson. 1986. "Current sensing utilizing heterodyne detection of the Faraday effect in single-mode optical fiber." *Journal of Lightwave Technology* 4 (6): 640–644. ISSN: 0733-8724. https://doi.org/10.1109/JLT.1986.1074778. http://ieeexplore.ieee.org/document/1074778/.

Kim, Jae-yeol, and Hyuk-Yoon Kwon. 2022. "Threat classification model for security information event management focusing on model efficiency." *Computers & Security* 120:102789. ISSN: 0167-4048. https://doi.org/https://doi.org/10.1016/j.cose.2022.102789. https://www.sciencedirect.com/science/article/pii/S0167404822001821.

Kim, Jinwoo, Kyungjun Min, Minhyuk Jung, and Seokho Chi. 2020. "Occupant behavior monitoring and emergency event detection in single-person households using deep learning-based sound recognition." *Building and Environment* 181:107092. ISSN: 0360-1323. https://doi.org/https://doi.org/10.1016/j.buildenv.2020.107092. https://www.sciencedirect.com/science/article/pii/S0360132320304686.

Kiranyaz, S., Onur Avcı, Osama Abdeljaber, T. Ince, M. Gabbouj, and D. Inman. 2019. "1D Convolutional Neural Networks and Applications: A Survey." *ArXiv* abs/1905.03554. https://doi.org/10.1016/j.ymssp.2020.107398.

Kissinger, P., and W. Heineman. 2005. "Cyclic voltammetry," https://doi.org/10.1021/ed060p702.

Krohling, Breno, and R. Krohling. 2023. "1D Convolutional neural networks and machine learning algorithms for spectral data classification with a case study for Covid-19." *ArXiv* abs/2301.10746. https://doi.org/10.48550/arXiv.2301.10746.

Kumar, Saurabh, and Ashish Kalkal. 2021. “3 - Electrochemical detection: Cyclic voltammetry/differential pulse voltammetry/impedance spectroscopy.” In *Nanotechnology in Cancer Management,* edited by Kamil Reza Khondakar and Ajeet Kumar Kaushik, 43–71. Elsevier. ISBN: 978-0-12-818154-6. https://doi.org/https://doi.org/10.1016/B978-0-12-818154-6.00008-1. https://www.sciencedirect.com/science/article/pii/B9780128181546000081.

Kurosawa, Kiyoshi. 2014. “Development of fiber-optic current sensing technique and its applications in electric power systems.” *Photonic Sensors* 4 (1): 12–20. ISSN: 1674-9251. https://doi.org/10.1007/s13320-013-0138-z. http://link.springer.com/10.1007/s13320-013-0138-z.

Ladányi, A., A. Ladányi, A. Mukherjee, H. Kenny, Alyssa Johnson, A. Mitra, Sinju Sundaresan, et al. 2018. “Adipocyte-induced CD36 expression drives ovarian cancer progression and metastasis.” *Oncogene* 37:2285–2301. https://doi.org/10.1038/s41388-017-0093-z.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep learning.” *Nature* 521 (7553): 436–444. ISSN: 0028-0836. https://doi.org/10.1038/nature14539. https://www.nature.com/articles/nature14539.

Lee, Kwang-Sig, Hyun-Joon Park, Ji Eon Kim, Hee Jung Kim, Sangil Chon, Sangkyu Kim, Jaesung Jang, et al. 2022. “Compressed Deep Learning to Classify Arrhythmia in an Embedded Wearable Device.” *Sensors* 22 (5). ISSN: 1424-8220. https://doi.org/10.3390/s22051776. https://www.mdpi.com/1424-8220/22/5/1776.

Li, Ji, and Yongfei Shen. 2017. “Image describing based on bidirectional LSTM and improved sequence sampling.” In *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA),* 735–739. IEEE, March. ISBN: 978-1-5090-3618-9. https://doi.org/10.1109/ICBDA.2017.8078733. http://ieeexplore.ieee.org/document/8078733/.

Li, Yao, Zhengcai Zhao, Yucan Fu, and Qingliang Chen. 2024. “A novel approach for tool condition monitoring based on transfer learning of deep neural networks using time–frequency images.” *Journal of Intelligent Manufacturing* 35, no. 3 (March): 1159–1171. ISSN: 1572-8145. https://doi.org/10.1007/s10845-023-02099-z. https://doi.org/10.1007/s10845-023-02099-z.

Li, Yaonian, Xiaozhou Huang, Erin Witherspoon, Zhe Wang, Pei Dong, and Qiliang Li. 2024. “Intelligent Electrochemical Sensors for Precise Identification of Volatile Organic Compounds Enabled by Neural Network Analysis.” *IEEE Sensors Journal* 24 (9): 15011–15022. https://doi.org/10.1109/JSEN.2024.3374354.

Li, Zewen, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. 2022. “A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects.” *IEEE transactions on neural networks and learning systems* 33 (12): 6999–7019. ISSN: 2162-2388. https://doi.org/10.1109/TNNLS.2021.3084827. http://www.ncbi.nlm.nih.gov/pubmed/34111009.

Liao, Xinzhi, Sheng Yan, Jialin Li, Chengming Jiang, Sigen Huang, Shengyi Liu, X. Zou, Guoxi Zhang, Junrong Zou, and Quanliang Liu. 2022. “CD36 and Its Role in Regulating the Tumor Microenvironment.” *Current Oncology* 29:8133–8145. https://doi.org/10.3390/curroncol29110642.

Lin, Haojia, Zhilu Yuan, Biao He, Xi Kuai, Xiaoming Li, and Renzhong Guo. 2022. “A Deep Learning Framework for Video-Based Vehicle Counting.” *Frontiers in Physics* 10 (February): 829734. ISSN: 2296-424X. https://doi.org/10.3389/fphy.2022.829734. https://www.frontiersin.org/articles/10.3389/fphy.2022.829734/full.

Liu, Lei, Weiqi Song, Chao Zeng, and Xiaohui Yang. 2021. “Microseismic event detection and classification based on convolutional neural network.” *Journal of Applied Geophysics* 192:104380. ISSN: 0926-9851. https://doi.org/https://doi.org/10.1016/j.jappgeo.2021.104380. https://www.sciencedirect.com/science/article/pii/S0926985121001270.

Liu, Shing-Hong, Chuan-Bi Lin, Ying Chen, Wenxi Chen, Tai-Shen Huang, and Chi-Yueh Hsu. 2019. "An EMG Patch for the Real-Time Monitoring of Muscle-Fatigue Conditions During Exercise." *Sensors (Basel, Switzerland)* 19. https://doi.org/10.3390/s19143108.

Liu, Ze, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and B. Guo. 2021. "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows." *2021 IEEE/CVF International Conference on Computer Vision (ICCV),* 9992–10002. https://doi.org/10.1109/ICCV48922.2021.00986.

Lomax, Anthony, Alberto Michelini, and Dario Jozinović. 2019. "An Investigation of Rapid Earthquake Characterization Using Single-Station Waveforms and a Convolutional Neural Network." *Seismological Research Letters* 90, no. 2A (February): 517–529. ISSN: 0895-0695. https://doi.org/10.1785/0220180311. eprint: https://pubs.geoscienceworld.org/ssa/srl/article-pdf/90/2A/517/4655400/srl-2018311.1.pdf. https://doi.org/10.1785/0220180311.

Mahmoud, Seedahmed S. 2023. "Practical Aspects of Perimeter Intrusion Detection and Nuisance Suppression for Distributed Fiber-Optic Sensors." *IEEE Transactions on Instrumentation and Measurement* 72:1–11. https://doi.org/10.1109/TIM.2023.3284133.

Maity, Sourav, and Karan Veer. 2023. "A Generalized Review Of Human-Computer Interaction Using Electromyogram Signals." *Recent Patents on Engineering* 17 (4): 16–25. ISSN: 1872-2121/2212-4047. https://doi.org/10.2174/1872212116666220518122621. http://www.eurekaselect.com/article/123727.

Mamun, M. Al, Y. Abdul Wahab, M. Hossain, A. Hashem, N. Hamizi, Z. Chowdhury, S. F. Wan Muhamad Hatta, I. Badruddin, S. Kamangar, and M. Johan. 2023. "Differential Pulse Voltammetric Tuning of the Screen-Printed Carbon Electrode Surface to Enhance the Electrochemical Performance and Multiplex Detection." *Journal of The Electrochemical Society,* https://doi.org/10.1149/1945-7111/ace7f8.

Masoudi, Ali, and Trevor P. Newson. 2017a. “Analysis of distributed optical fibre acoustic sensors through numerical modelling.” *Optics Express* 25 (25): 32021. ISSN: 1094-4087. https://doi.org/10.1364/OE.25.032021. https://opg.optica.org/abstract.cfm?URI=oe-25-25-32021.

Masoudi, Ali, and Trevor P. Newson. 2017b. “High spatial resolution distributed optical fiber dynamic strain sensor with enhanced frequency and strain resolution.” *Optics Letters* 42 (2): 290. ISSN: 0146-9592. https://doi.org/10.1364/OL.42.000290. https://opg.optica.org/abstract.cfm?URI=ol-42-2-290.

Mcmanus, Lara, G. de Vito, and M. Lowery. 2020. “Analysis and Biophysics of Surface EMG for Physiotherapists and Kinesiologists: Toward a Common Language With Rehabilitation Engineers.” *Frontiers in Neurology* 11. https://doi.org/10.3389/fneur.2020.576729.

Meliboev, Aziz, Jumabek Alikhanov, and Wooseong Kim. 2020. “1D CNN based network intrusion detection with normalization on imbalanced data.” *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC),* 218–224. https://doi.org/10.1109/ICAIIC48513.2020.9064976.

Miao, W., E. Vasile, William Arbuthnot Sir Lane, and J. Lawler. 2001. “CD36 associates with CD9 and integrins on human blood platelets.” *Blood* 97 6:1689–96. https://doi.org/10.1182/BLOOD.V97.6.1689.

Mills, K. 2005. “The basics of electromyography.” *Journal of Neurology, Neurosurgery & Psychiatry* 76:ii32–ii35. https://doi.org/10.1136/jnnp.2005.069211.

Mo, Wenjing, Chuanjin Han, and Yuan Wang. 2023. “A Baseline Modeling and Anomaly Detection Method for Inertial Navigation System Based on Integrated 1DCNN.” *2023 IEEE 16th International Conference on Electronic Measurement & Instruments (ICEMI),* 142–147. https://doi.org/10.1109/ICEMI59194.2023.10270760.

Mohapatra, Amlan Deep, Amol Aggarwal, and Rajesh Kumar Tripathy. 2024. "Automated Recognition of Hand Gestures From Multichannel EMG Sensor Data Using Time–Frequency Domain Deep Learning for IoT Applications." *IEEE Sensors Letters* 8 (6): 1–4. https://doi.org/10.1109/LSENS.2024.3400388.

Molinara, M., R. Cancelliere, A. Tinno, L. Ferrigno, M. Shuba, P. Kuzhir, A. Maffucci, and L. Micheli. 2022. "A Deep Learning Approach to Organic Pollutants Classification Using Voltammetry." *Sensors (Basel, Switzerland)* 22. https://doi.org/10.3390/s22208032.

Mredula, Motahara Sabah, Noyon Dey, Md. Sazzadur Rahman, Imtiaz Mahmud, and You-Ze Cho. 2022. "A Review on the Trends in Event Detection by Analyzing Social Media Platforms." *Sensors* 22 (12). ISSN: 1424-8220. https://doi.org/10.3390/s22124531. https://www.mdpi.com/1424-8220/22/12/4531.

Nahid, Nazmun, Arafat Rahman, and M.A.R. Ahad. 2020. "Deep Learning Based Surface EMG Hand Gesture Classification for Low-Cost Myoelectric Prosthetic Hand." In *2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR),* 1–8. https://doi.org/10.1109/ICIEVicIVPR48672.2020.9306613.

Nguyen, Truong X., Jay J. Ely, and George N. Szatkowski. 2015. "A fiber-optic current sensor for lightning measurement applications." In *Fiber Optic Sensors and Applications XII,* edited by Gary Pickrell, Eric Udd, and Henry H. Du, 94800X. May. https://doi.org/10.1117/12.2179195. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2179195.

Ni, Wenhao, Liqin Hu, and Kuanglu Yu. 2023. "A ϕ-OTDR event identification based on data augmentation with TSG." In *Asia-Pacific Optical Sensors Conference (APOS 2023),* edited by Tiegen Liu, Yuwen Qin, Yunjiang Rao, and Zuyuan He, 12813:1281303. International Society for Optics and Photonics, SPIE. https://doi.org/10.1117/12.2691540. https://doi.org/10.1117/12.2691540.

Nicati, P A, and P Robert. 1988. “Stabilised current sensor using Sagnac interferometer.” *Journal of Physics E: Scientific Instruments* 21 (8): 791–796. ISSN: 0022-3735. https://doi.org/10.1088/0022-3735/21/8/010. https://iopscience.iop.org/article/10.1088/0022-3735/21/8/010.

Noureldin, Hind A. M., Ali M. Abdel-Aziz, M. Mabrouk, Amira H. K. Saad, and I. Badr. 2023. “Green and cost-effective voltammetric assay for spiramycin based on activated glassy carbon electrode and its applications to urine and milk samples.” *RSC Advances* 13:844–852. https://doi.org/10.1039/d2ra06768d.

Odusami, Modupe, Rytis Maskeliūnas, and Robertas Damaševičius. 2022. “An Intelligent System for Early Recognition of Alzheimer’s Disease Using Neuroimaging.” *Sensors* 22 (3): 740. ISSN: 1424-8220. https://doi.org/10.3390/s22030740. https://www.mdpi.com/1424-8220/22/3/740.

Ozdemir, Mehmet Akif, Ozlem Karabiber Cura, and Aydin Akan. 2021. “Epileptic EEG Classification by Using Time-Frequency Images for Deep Learning.” PMID: 34039254, *International Journal of Neural Systems* 31 (08): 2150026. https://doi.org/10.1142/S012906572150026X. eprint: https://doi.org/10.1142/S012906572150026X. https://doi.org/10.1142/S012906572150026X.

Ozdemir, Mehmet Akif, Deniz Hande Kisa, Onan Guren, and Aydin Akan. 2022a. “Dataset for multi-channel surface electromyography (sEMG) signals of hand gestures.” *Data in Brief* 41:107921. ISSN: 2352-3409. https://doi.org/https://doi.org/10.1016/j.dib.2022.107921. https://www.sciencedirect.com/science/article/pii/S2352340922001330.

Ozdemir, Mehmet Akif, Deniz Hande Kisa, Onan Guren, and Aydin Akan. 2022b. “Hand gesture classification using time–frequency images and transfer learning based on CNN.” *Biomedical Signal Processing and Control* 77:103787. ISSN: 1746-8094. https://doi.org/https://doi.org/10.1016/j.bspc.2022.103787. https://www.sciencedirect.com/science/article/pii/S1746809422003093.

Pepino, M., O. Kuda, D. Samovski, and N. Abumrad. 2014. "Structure-function of CD36 and importance of fatty acid signal transduction in fat metabolism." *Annual review of nutrition* 34:281–303. https://doi.org/10.1146/annurev-nutr-071812-161220.

Porumb, Mihaela, Saverio Stranges, Antonio Pescapè, and Leandro Pecchia. 2020. "Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ECG." *Scientific reports* 10 (1): 170.

Qi, Jinxian, Guozhang Jiang, Gongfa Li, Ying Sun, and Bo Tao. 2019. "Intelligent Human-Computer Interaction Based on Surface EMG Gesture Recognition." *IEEE Access* 7:61378–61387. https://doi.org/10.1109/ACCESS.2019.2914728.

Rao, Yunjiang, Zinan Wang, Huijuan Wu, Zengling Ran, and Bing Han. 2021. "Recent Advances in Phase-Sensitive Optical Time Domain Reflectometry (Φ-OTDR)." *Photonic Sensors* 11 (1): 1–30. ISSN: 1674-9251. https://doi.org/10.1007/s13320-021-0619-4. http://link.springer.com/10.1007/s13320-021-0619-4.

Rehman, Anwar Ur, Ahmad Kamran Malik, Basit Raza, and Waqar Ali. 2019. "A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis." *Multimedia Tools and Applications* 78 (18): 26597–26613. ISSN: 1380-7501. https://doi.org/10.1007/s11042-019-07788-7. http://link.springer.com/10.1007/s11042-019-07788-7.

Saad, Omar M., Ali G. Hafez, and M. Sami Soliman. 2021. "Deep Learning Approach for Earthquake Parameters Classification in Earthquake Early Warning System." *IEEE Geoscience and Remote Sensing Letters* 18 (7): 1293–1297. https://doi.org/10.1109/LGRS.2020.2998580.

Samuel, O. W., M. G. Asogbon, Yanjuan Geng, Ali H. Al-timemy, Sandeep Pirbhulal, Ning Ji, Shixiong Chen, Peng Fang, and Guanglin Li. 2019. "Intelligent EMG Pattern Recognition Control Method for Upper-Limb Multifunctional Prostheses: Advances,

Current Challenges, and Future Prospects." *IEEE Access* 7:10150–10165. https://doi.org/10.1109/ACCESS.2019.2891350.

Sathi, Khaleda Akhter, Md Kamal Hosain, Md. Azad Hossain, and Abbas Z. Kouzani. 2023. "Attention-assisted hybrid 1D CNN-BiLSTM model for predicting electric field induced by transcranial magnetic stimulation coil." *Scientific Reports* 13 (1): 2494. ISSN: 2045-2322. https://doi.org/10.1038/s41598-023-29695-6. https://www.nature.com/articles/s41598-023-29695-6.

Shewalkar, Apeksha, Deepika Nyavanandi, and Simone A. Ludwig. 2019. "Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU." *Journal of Artificial Intelligence and Soft Computing Research* 9 (4): 235–245. ISSN: 2083-2567. https://doi.org/10.2478/jaiscr-2019-0006. https://www.sciendo.com/article/10.2478/jaiscr-2019-0006.

Shi, Shuangshuang, Kefei Zhang, Suqin Wu, Jiaqi Shi, Andong Hu, Huajing Wu, and Yu Li. 2022. "An Investigation of Ionospheric TEC Prediction Maps Over China Using Bidirectional Long Short-Term Memory Method." *Space Weather* 20 (6): e2022SW003103. ISSN: 1542-7390. https://doi.org/10.1029/2022SW003103. https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022SW003103.

Shi, Yi, Yinghuan Li, Yingchao Zhang, Zhemin Zhuang, and Tao Jiang. 2021. "An Easy Access Method for Event Recognition of Φ-OTDR Sensing System Based on Transfer Learning." *Journal of Lightwave Technology, Vol. 39, Issue 13, pp. 4548-4555* 39 (13): 4548–4555. https://opg.optica.org/jlt/abstract.cfm?uri=jlt-39-13-4548.

Shi, Yi, Yuanye Wang, Liyuan Wang, Lei Zhao, and Zhun Fan. 2020. "Multi-event classification for Φ-OTDR distributed optical fiber sensing system using deep learning and support vector machine." *Optik* 221 (November): 165373. ISSN: 00304026. https://doi.org/10.1016/j.ijleo.2020.165373. https://linkinghub.elsevier.com/retrieve/pii/S0030402620312092.

Sikdar, Shirsendu, Dianzi Liu, and Abhishek Kundu. 2022. "Acoustic emission data based deep learning approach for classification and detection of damage-sources in a composite panel." *Composites Part B: Engineering* 228 (January): 109450. ISSN: 13598368. https://doi.org/10.1016/j.compositesb.2021.109450. https://linkinghub.elsevier.com/retrieve/pii/S1359836821008179.

Silverstein, R., and M. Febbraio. 2009. "CD36, a Scavenger Receptor Involved in Immunity, Metabolism, Angiogenesis, and Behavior." *Science Signaling* 2:re3–re3. https://doi.org/10.1126/scisignal.272re3.

Simonyan, Karen, and Andrew Zisserman. 2014. "Very Deep Convolutional Networks for Large-Scale Image Recognition." *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (September). https://arxiv.org/abs/1409.1556v6.

Sirin, Samil. 2022. "Fiber optic current sensor interrogated by phase-sensitive optical time-domain reflectometer." Master's thesis, Izmir Institute of Technology, July. https://gcris.iyte.edu.tr/handle/11147/12628.

Sirin, Samil, Kivilcim Yuksel Aldogan, and Marc Wuilpart. 2022. "Current sensing using a Phase-Sensitive Optical Time Domain Reflectometer: Feasibility study." *Optical Fiber Technology* 74 (December): 103084. ISSN: 10685200. https://doi.org/10.1016/j.yofte.2022.103084. https://linkinghub.elsevier.com/retrieve/pii/S106852002200267X.

Su, Ziyi, Qingchao Liu, Chunxia Zhao, and Fengming Sun. 2022. "A Traffic Event Detection Method Based on Random Forest and Permutation Importance." *Mathematics* 10 (6). ISSN: 2227-7390. https://doi.org/10.3390/math10060873. https://www.mdpi.com/2227-7390/10/6/873.

Sun, Min, and Nian Fang. 2023. "Explainable feature selection by SHAP for signal recognition of ϕ-OTDR system." In *Fourteenth International Conference on Information*

Optics and Photonics (CIOP 2023), edited by Yue Yang, 12935:1293540. International Society for Optics and Photonics, SPIE. https://doi.org/10.1117/12.3007857. https://doi.org/10.1117/12.3007857.

Taiar, Redha, Rukhma Qasim, Waqas Haider Bangyal, Mohammed A. Alqarni, and Abdulwahab Ali Almazroi. 2022. "A Fine-Tuned BERT-Based Transfer Learning Approach for Text Classification." *Journal of Healthcare Engineering* 2022:3498123. ISSN: 2040-2295. https://doi.org/10.1155/2022/3498123. https://doi.org/10.1155/2022/3498123.

Tan, Mingxing, and Quoc Le. 2019. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." In *Proceedings of the 36th International Conference on Machine Learning,* edited by Kamalika Chaudhuri and Ruslan Salakhutdinov, 97:6105–6114. Proceedings of Machine Learning Research. PMLR, June. https://proceedings.mlr.press/v97/tan19a.html.

Tang, Jianbin, Rima El Atrache, Shuang Yu, Umar Asif, Michele Jackson, Subhrajit Roy, Mahtab Mirmomeni, et al. 2021. "Seizure detection using wearable sensors and machine learning: Setting a benchmark." *Epilepsia* 62 (8): 1807–1819. https://doi.org/https://doi.org/10.1111/epi.16967. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/epi.16967. https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.16967.

Tatsunami, Yuki, and Masato Taki. 2022. "Sequencer: Deep LSTM for Image Classification." *Advances in Neural Information Processing Systems* 35 (December): 38204–38217. https://doi.org/10.48550/arXiv.2205.01972. https://proceedings.neurips.cc/paper_files/paper/2022/hash/f9d7d6c695bc983fcfb5b70a5fbdfd2f-Abstract-Conference.html.

Tu, Zhengzhong, Hossein Talebi, Han Zhang, Feng Yang, P. Milanfar, A. Bovik, and Yinxiao Li. 2022. "MaxViT: Multi-Axis Vision Transformer," 459–479. https://doi.org/10.48550/arXiv.2204.01697.

Ullah, Safi, and K. Iqbal. 2020. “A Preliminary Review on EMG Signals for Assessment of Diabetic Peripheral Neuropathy Disorder.” *2020 7th International Conference on Electrical and Electronics Engineering (ICEEE),* 42–46. https://doi.org/10.1109/ICEEE49618.2020.9102488.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. “Attention is All you Need.” In *Advances in Neural Information Processing Systems,* edited by I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Vidaña-Vila, Ester, Joan Navarro, Dan Stowell, and Rosa Ma Alsina-Pagès. 2021. “Multilabel Acoustic Event Classification Using Real-World Urban Data and Physical Redundancy of Sensors.” *Sensors* 21 (22). ISSN: 1424-8220. https://doi.org/10.3390/s21227470. https://www.mdpi.com/1424-8220/21/22/7470.

Wang, Jianping, Jian Ma, Dean Meng, Xuan Zhao, and Kai Zhang. 2023. “Fault Diagnosis of PMSMs Based on Image Features of Multi-Sensor Fusion.” *Sensors* 23 (20). ISSN: 1424-8220. https://doi.org/10.3390/s23208592. https://www.mdpi.com/1424-8220/23/20/8592.

Wang, Jin-lei, Bo-chao Li, Zi-jun Li, Ke-feng Ren, Lie-jiang Jin, Shi-miao Zhang, Hao Chang, Yi-xin Sun, and Jian Ji. 2014. “Electropolymerization of dopamine for surface modification of complex-shaped cardiovascular stents.” *Biomaterials* 35 (27): 7679–7689. ISSN: 0142-9612. https://doi.org/https://doi.org/10.1016/j.biomaterials.2014.05.047. https://www.sciencedirect.com/science/article/pii/S0142961214006048.

Wang, Ming, Hao Feng, Dunzhe Qi, Lipu Du, and Zhou Sha. 2023. “Φ-OTDR pattern recognition based on CNN-LSTM.” *Optik* 272 (February): 170380. ISSN: 00304026. https://doi.org/10.1016/j.ijleo.2022.170380. https://linkinghub.elsevier.com/retrieve/pii/S0030402622016382.

Wang, Wei, Ming Zhu, Jinlin Wang, Xuewen Zeng, and Zhongzhen Yang. 2017. "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," 43–48. https://doi.org/10.1109/ISI.2017.8004872.

Wang, Zhiguang, and Tim Oates. 2015. "Spatially Encoding Temporal Correlations to Classify Temporal Data Using Convolutional Neural Networks." *arXiv e-prints* (September): arXiv:1509.07481. https://doi.org/10.48550/arXiv.1509.07481. arXiv: 1509.07481 [cs.LG].

Woodward, Kieran, Eiman Kanjo, and Athanasios Tsanas. 2023. "Combining Deep Learning with Signal-Image Encoding for Multi-Modal Mental Wellbeing Classification." Just Accepted, *ACM Trans. Comput. Healthcare* (New York, NY, USA) (November). https://doi.org/10.1145/3631618. https://doi.org/10.1145/3631618.

Wu, Huijuan, Jiping Chen, Xiangrong Liu, Yao Xiao, Mengjiao Wang, Yi Zheng, and Yunjiang Rao. 2019. "One-Dimensional CNN-Based Intelligent Recognition of Vibrations in Pipeline Monitoring With DAS." *Journal of Lightwave Technology* 37 (17): 4359–4366. https://doi.org/10.1109/JLT.2019.2923839.

Wu, Huijuan, Shunkun Xiao, Xiaoyu Li, Zinan Wang, Jiwei Xu, and Yunjiang Rao. 2015. "Separation and Determination of the Disturbing Signals in Phase-Sensitive Optical Time Domain Reflectometry (Φ-OTDR)." *Journal of Lightwave Technology, Vol. 33, Issue 15, pp. 3156-3162* 33 (15): 3156–3162. https://opg.optica.org/jlt/abstract.cfm?uri=jlt-33-15-3156.

Xie, Lang, Zhaojie Li, Yiwei Zhou, Weiming Xiang, Yu Wu, and Yunjiang Rao. 2023. "Railway Track Online Detection Based on Optical Fiber Distributed Large-Range Acoustic Sensing." *IEEE Internet of Things Journal,* 1–1. https://doi.org/10.1109/JIOT.2023.3311173.

Xu, Guixian, Yueting Meng, Xiaoyu Qiu, Ziheng Yu, and Xu Wu. 2019. "Sentiment Analysis of Comment Texts Based on BiLSTM." *IEEE Access* 7:51522–51532. ISSN:

2169-3536. https://doi.org/10.1109/ACCESS.2019.2909919. https://ieeexplore.ieee.org/document/8684825/.

Xu, Shaoyi, Wei Li, Yuqiao Wang, and Fangfang Xing. 2014. "Stray current sensor with cylindrical twisted fiber." *Applied optics* 53 (24): 5486–92. ISSN: 1539-4522. https://doi.org/10.1364/AO.53.005486. http://www.ncbi.nlm.nih.gov/pubmed/25321123.

Yang, Nachuan, Yongjun Zhao, and Jinyang Chen. 2022. "Real-Time Phi;-OTDR Vibration Event Recognition Based on Image Target Detection." *Sensors* 22 (3). ISSN: 1424-8220. https://doi.org/10.3390/s22031127. https://www.mdpi.com/1424-8220/22/3/1127.

Yang, Nachuan, Yongjun Zhao, Jinyang Chen, and Fuqiang Wang. 2023. "Real-time classification for φ-OTDR vibration events in the case of small sample size datasets." *Optical Fiber Technology* 76:103217. ISSN: 1068-5200. https://doi.org/https://doi.org/10.1016/j.yofte.2022.103217. https://www.sciencedirect.com/science/article/pii/S1068520022004023.

Yang, Yonghua, Rani A Sarkis, Rima El Atrache, Tobias Loddenkemper, and Christian Meisel. 2021. "Video-Based Detection of Generalized Tonic-Clonic Seizures Using Deep Learning." *IEEE journal of biomedical and health informatics* 25 (8): 2997–3008. ISSN: 2168-2208. https://doi.org/10.1109/JBHI.2021.3049649. http://www.ncbi.nlm.nih.gov/pubmed/33406048.

Yin, Xiaochun, Zengguang Liu, Deyong Liu, and Xiaojun Ren. 2022. "A Novel CNN-based Bi-LSTM parallel model with attention mechanism for human activity recognition with noisy data." *Scientific Reports* 12 (1): 7878. ISSN: 2045-2322. https://doi.org/10.1038/s41598-022-11880-8. https://www.nature.com/articles/s41598-022-11880-8.

Yolchuyev, Orkhan, and Z. Aydoğmuş. 2023. "Electrochemical behavior and differential pulse voltammetric determination of budesonide in suspension ampoules." *İstanbul Journal of Pharmacy,* https://doi.org/10.26650/istanbuljpharm.2023.1093821.

Yu, Yantao, Xin Zeng, Xiaoping Xue, and Jingxiao Ma. 2022. "LSTM-Based Intrusion Detection System for VANETs: A Time Series Classification Approach to False Message Detection." *IEEE Transactions on Intelligent Transportation Systems* 23 (12): 23906–23918. ISSN: 1524-9050. https://doi.org/10.1109/TITS.2022.3190432. https://ieeexplore.ieee.org/document/9832002/.

Yu, Yong, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. 2019. "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures." *Neural Computation* 31 (7): 1235–1270. ISSN: 0899-7667. https://doi.org/10.1162/neco_a_01199. https://direct.mit.edu/neco/article/31/7/1235-1270/8500.

Zeybekler, Simge Er, and Dilek Odaci. 2023. "Carbon Nanotube-Incorporated Nanofibers for Immunosensor Preparation against CD36." *ACS Omega* 8:5776–5786. https://doi.org/10.1021/acsomega.2c07458.

Zhao, Xiaoting, Hongbin Sun, Bo Lin, Hongmin Zhao, Yingli Niu, Xiang Zhong, Yidan Wang, et al. 2022. "Markov Transition Fields and Deep Learning-Based Event-Classification and Vibration-Frequency Measurement for φ-OTDR." *IEEE Sensors Journal* 22, no. 4 (February): 3348–3357. ISSN: 1558-1748. https://doi.org/10.1109/JSEN.2021.3137006.

Zhao, Zehui, Laith Alzubaidi, Jinglan Zhang, Ye Duan, and Yuantong Gu. 2024. "A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations." *Expert Systems with Applications* 242:122807. ISSN: 0957-4174. https://doi.org/https://doi.org/10.1016/j.eswa.2023.122807. https://www.sciencedirect.com/science/article/pii/S0957417423033092.

Zhou, Xuanyi, Wen Qi, S. E. Ovur, Longbin Zhang, Yingbai Hu, Hang Su, G. Ferrigno, and E. Momi. 2020. "A novel muscle-computer interface for hand gesture recognition using depth vision." *Journal of Ambient Intelligence and Humanized Computing* 11:5569–5580. https://doi.org/10.1007/s12652-020-01913-3.

Zhou, Zhongzeng, Luojun Wang, Jing Wang, Conghui Liu, Tailin Xu, and Xueji Zhang. 2022. “Machine Learning with Neural Networks to Enhance Selectivity of Nonenzymatic Electrochemical Biosensors in Multianalyte Mixtures.” *ACS applied materials & interfaces,* https://doi.org/10.1021/acsami.2c17593.

Zhuang, Fuzhen, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. 2021. “A Comprehensive Survey on Transfer Learning.” *Proceedings of the IEEE* 109, no. 1 (January): 43–76. ISSN: 1558-2256. https://doi.org/10.1109/JPROC.2020.3004555.