



Machine Learning Based Classification for Alzheimer's Disease Using EEG Signal

Abdulfatah Kalaje^{1,◇} and Hasan Demir^{2,¥}

¹Kalacı mühendislik, Çorlu, Tekirdağ, Türkiye.

²Department of Electronics and Communication Engineering, Çorlu Faculty of Engineering, Tekirdağ Namık Kemal University, Türkiye.

To cite this article:

Abdulfatah Kalaje and Hasan Demir. "Machine Learning Based Classification for Alzheimer's Disease Using EEG Signal", *Parana Journal of Science and Education*. Vol. 12, No. 1, 2026, pp. 12-23.

Received: January 9, 2026; **Accepted:** January 25, 2026; **Published:** February 3, 2026.

Abstract

Alzheimer's disease (AD) is neurodegenerative disorder that affects memory, cognition, and behavior, representing the most common cause of dementia with old people. Alzheimer's disease detection was performed using convolutional neural networks (CNNs) applied to electroencephalogram (EEG) signals, where the features were extracted from signal transformation coefficients. The EEG dataset collected from 48 participants, divided into two groups: Alzheimer's disease patients and healthy controls. Several signal transformation techniques were compared, including the fast Fourier transform (FFT), short-time Fourier transform (STFT), synchrosqueezed Fourier transform (SSFT), continuous wavelet transform (CWT), discrete wavelet transform (DWT) for 1D and 2D, and synchrosqueezed wavelet transform (SSWT), to determine the most effective approach for EEG-based classification. Experimental results demonstrated that the STFT method provided the highest performance, achieving superior accuracy, precision, sensitivity, specificity, and F1-score for Alzheimer's disease detection.

Keywords: Alzheimer, Convolutional Neural Network, Electroencephalography (EEG), Machine learning.

[◇] Email: janaaboda19@gmail.com

[¥] Email: hdemir@nku.edu.tr



1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia in older adults, causing memory loss, cognitive decline, and behavioral changes. It primarily affects the hippocampus and cerebral cortex, with early symptoms including forgetfulness, confusion, and difficulty with daily activities, progressing to severe cognitive and language impairments. Traditional diagnostic methods, such as memory and cognitive tests, are often subjective and influenced by factors like anxiety, depression and stress. To overcome these challenges, neuroimaging and electrophysiological techniques have been increasingly explored. Among these, electroencephalography (EEG) demonstrates as promising tools for early and objective detection of AD. EEG shows Slowing of electrical brain activity with AD patients. Machine learning, a branch of artificial intelligence, provides an effective alternative for detecting and classifying AD from medical signal data. By learning patterns from multi-layered structures, machine-learning models can identify symptoms early and improve diagnostic accuracy. The aim of this study is to determine the most effective signal transformation for detecting AD by applying deep learning techniques and comparing the results. There are various studies that used this type of datasets and machine learning methods. In the study conducted by Morteza Amini, Mir Mohsen Pedram, AliReza Moradi, and Mahshad Ouchani [1], EEG signal were utilized to diagnose Alzheimer's disease (AD). The dataset included three subject groups: Alzheimer's disease (AD), mild cognitive impairment (MCI), and healthy control (HC). To improve signal resolution and feature extraction wavelet transform were applied. K-nearest neighbors (KNN), support vector machine (SVM), linear discriminant analysis (LDA), and convolutional neural networks (CNN) were used for classification. The models were evaluated using performance metrics such as sensitivity, precision, area under the curve (AUC), and overall accuracy to determine the most effective classifier. The KNN model achieved sensitivities of 79.7% (MCI), 71.9% (AD), and 62.5% (HC), with corresponding precisions of 63.7%, 75.4%, and 78.4%, an AUC of 0.902, and 71.4% accuracy. The SVM model demonstrated lower performance, with sensitivities of 9.4% (MCI), 32.8% (AD), and 81.3% (HC), precisions of 31.6%, 47.7%, and 40.3%, an AUC of 0.593, and an accuracy of

41.1%. Similarly, the LDA classifier obtained sensitivities of 28.1% (MCI), 53.1% (AD), and 50.0% (HC), precisions of 45.0%, 43.0%, and 48.8%, an AUC of 0.594, and an accuracy of 43.8%. Among all models, the CNN architecture achieved the best performance, with sensitivities of 82.8% (MCI), 89.1% (AD), and 75.5% (HC), precisions of 81.5%, 79.2%, and 87.3%, an AUC of 0.988, and an overall accuracy of 82.3%. In the study conducted by Hanife GÖKER [2], EEG signals were analyzed to detect alzheimer's disease using power spectral density (PSD) features. The multitaper method was applied to compute PSD values across the 1–49 Hz frequency range, resulting in 49 extracted features representing different EEG frequency components. Multiple ensemble learning algorithms were evaluated for classification, including AdaBoostM1, Total Boost, Gentle Boost, Logit Boost, Robust Boost, and Bagging. Among these, the Logit Boost algorithm achieved the best performance, obtaining an accuracy of 93.04%, an F1-score of 93.09%, sensitivity of 92.75%, precision of 93.43%, and specificity of 93.33%. These results demonstrate the effectiveness of ensemble learning methods, particularly Logit Boost, in identifying Alzheimer's disease from EEG-derived frequency features. In the study that was made by Giulia Fiscon, Placido Bramanti, Paola Bertolazzi, Emanuel Weitschek, Alessio Cialini, Giovanni Felici, Simona De Salvo, Alessia Bramanti and Maria Cristina De Cola [3]. They classified Alzheimer's disease (AD) using EEG signals collected from 109 participants, consisting of individuals diagnosed with AD, (MCI), and (HC). The authors implemented a time–frequency signal analysis to extract the features using the discrete Fourier transform and discrete wavelet transform, followed by classification through tree-based supervised machine learning models. Their findings demonstrated that features derived from the wavelet transform were particularly discriminative, yielding classification accuracies of 83% for HC vs AD, 92%, for HC vs MCI and 79% for MCI vs AD. In the study that was made by Zülfikar Aslan [4]. The dataset included EEG recordings from 48 participants (24 AD and 24 healthy controls), acquired from 19 channels. After the noises were removed by preprocessing the raw EEG data, the DWT was applied to extract features from each channel. Three statistical measures Activity, Mobility, and Complexity were computed from the resulting



sub-band signals. The extracted features were then classified using the KNN algorithm. Classification performance varied across frequency bands, achieving accuracies of 67.1% Gamma, 63.6% Beta, 83.9% Alpha, 83.5% Theta, 88.3% Delta, and a highest accuracy of 91.1% when combining features from all sub-bands. The study demonstrates the effectiveness of wavelet-based statistical features in distinguishing AD from healthy subjects. In the study that prepared by Caroline L. Alves, Christiane Thielemann, Aruane M. Pineda, Francisco A. Rodrigues and Kirstin Roster [5], EEG signals were used for diagnosis of Alzheimer's disease and schizophrenia. The dataset consisted of EEG recorded from 48 participants for Alzheimer's disease (24 AD and 24 healthy controls) collected from 19 channels. They used EEG signal, which are free of artifacts to construct the matrices of connections. The connection strengths between brain regions were quantified using three different methods, Granger, causality and Pearson correlation, and Spearman correlation. Two CNN architectures were used: one with a hyperparameter tuning procedure (CNNtuned) and one without tuning (CNNuntuned). Hyperparameter tuning techniques including random search, hyperband, and Bayesian optimization were applied to enhance model performance. CNNuntuned demonstrated lower computational complexity

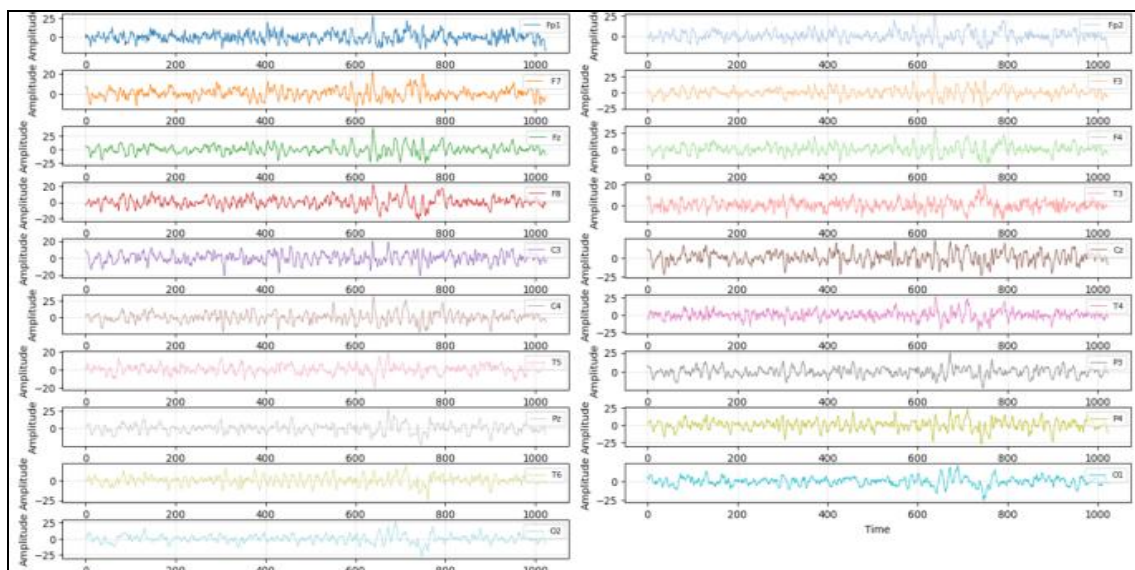
due to having fewer layers. The highest performance was achieved using Pearson correlation features with the hyperband-tuned CNN, resulting in 86% training accuracy and 100% testing accuracy. In comparison, the best CNNuntuned model achieved 98% training accuracy and 92% testing accuracy using Pearson correlation.

2. Material and Methods

2.1 Dataset

The dataset contain EEG time-series signal. The data was recorded from 19 channels (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, T5, T6, Pz, P3, P4, O1, O2) with sampling frequency 128 Hz. The channel nomenclature corresponds to the cerebral lobes: Frontal (F), Central (C), Parietal (P), Occipital (O), and Temporal (T). The data is divided into two sets. The first one contain 24 healthy old individuals (control group; aged 72 ± 11) who do not have any neurological disorders. The second one contain 24 old individuals with AD (aged 69 ± 16) diagnosed by the National Institute of Neurological and Communicative Disorders and Stroke [5].

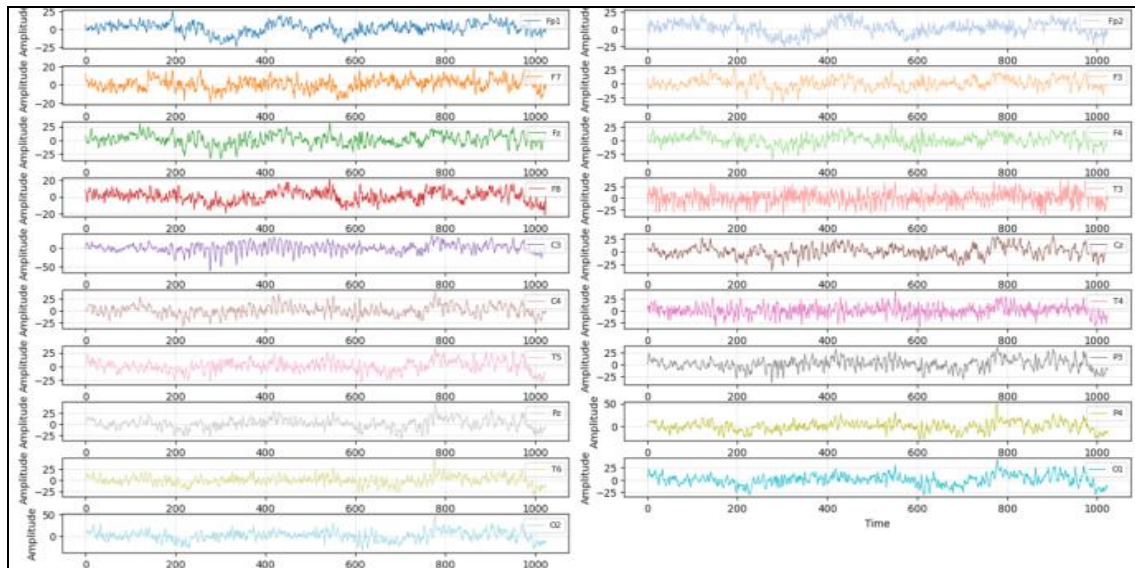
Figure 1: AD person EEG signal.



Source: Authors.



Figure 2: Healthy person EEG signal.



Source: Authors.

2.2 Signal Transformation

The raw EEG signal is a time-domain signal. However, time-domain representations often do not provide sufficient information for accurate analysis because they may contain noise, which can adversely affect the results. Signal transformations convert the signal into the frequency domain and time frequency domain, making it easier to identify and reduce noise while revealing additional information about the underlying signal characteristics.

2.2.1 Fast Fourier Transform (FFT)

FFT is a fundamental computational technique widely used for analysing the frequency of the signals. It efficiently computes the discrete frequency components of a signal, thereby producing its corresponding frequency spectrum. This transformation enables the extraction of meaningful frequency-domain features that are valuable for understanding the signal. Mathematically, the FFT formula shown in Equation (1) [6].

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi kn}{N}} \quad (1)$$

2.2.2 Short-time Fourier Transform (STFT)

STFT is a widely used technique for time frequency analysis of non-stationary signals. It addresses the limitation of conventional Fourier analysis, which assumes that the signal's frequency content remains constant over time, by providing a dynamic representation of how frequencies evolve throughout the signal's duration. The STFT operates by segmenting the original signal into overlapping windows using a time-sliding window function. For each window, FFT is applied to compute the local frequency spectrum. By continuously shifting the analysis window across the entire signal, a time-varying spectral representation is obtained. Mathematically, the STFT is expressed as shown in Equation (2) [7].

$$X(\tau, \omega) = \int_{-\infty}^{\infty} x(t) \cdot \omega(t - \tau) \cdot e^{-i\omega t} dt \quad (2)$$

2.2.3 Synchrosqueezed Fourier Transform (SSFT)

SSFT is an advanced time-frequency analysis technique that provides higher time-frequency resolution compared to the STFT. SSFT compress the time-frequency coefficients of the STFT more efficiently. That led to obtain signal with non-stationary properties such as good



localization and time-varying frequencies. SSFT uses a synchronization step to reassign the STFT frequency values. This synchronization step is calculated by reassigning the STFT frequency values using the estimated instantaneous frequency $\omega_f(\eta, t)$. Mathematically, the SSFT is shown in Equation (3-5) [8].

$$Vf(\eta, t) \int_{-\infty}^{\infty} f(x) \cdot g(x - t) \cdot e^{-i2\pi\eta(x-t)} dx \quad (3)$$

$$\omega_f(\eta, t) = \frac{1}{2j\pi} \frac{\partial_t Vf(\eta, t)}{Vf(\eta, t)} \quad (4)$$

$$Tf(\omega, t) = \frac{1}{g(0)} \int_{-\infty}^{\infty} Vf(\eta, t) \cdot \delta(\omega - \hat{\omega}_f(\eta, t)) d\eta \quad (5)$$

2.2.4 Continuous Wavelet Transform (CWT)

CWT is a powerful technique for analyzing the time–frequency characteristics of non-stationary signals. Unlike traditional Fourier-based approaches, the CWT provides simultaneous localization in both the time and frequency domains, enabling a more detailed examination of transient and time-varying signal components. By applying scalable and translatable wavelet functions, the CWT allows multi-resolution analysis, offering insights into the signal's structure at different scales and temporal positions. The mathematical formulation of the CWT is expressed in Equation (6) [9].

$$Wx(a, b) = \langle x(t), \psi_{a, b}(t) \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \psi^*\left(\frac{t-b}{a}\right) dt \quad (6)$$

2.2.5 Discrete Wavelet Transform (DWT)

DWT analyse signals with time-varying characteristics. The DWT applies high-pass and low-pass filters to the input signal, followed by subsampling. The high-pass filter captures the high-frequency components of the signal, isolating fine details and rapid changes. On the other hand, the low-pass filter captures the low-frequency components, preserving the overall signal trend and slower changes. After filtering, the signal is subsampled, which reduces the number of data points by preserving only each sample. This process effectively compresses signal data at different scales. Filtering and subsampling create multi-resolution analysis. The

discrete wavelet transform is mathematically illustrated in Equation (7) representing the detail coefficients of the signal, and Equation (8) representing the approximation coefficients [10].

$$D_j(n) = \langle x(t), \psi_{j, n}(t) \rangle = \sum_{-\infty}^{\infty} x(t) g(2t - n) \quad (7)$$

$$A_j(n) = \langle x(t), \phi_{j, n}(t) \rangle = \sum_{-\infty}^{\infty} x(t) h(2t - n) \quad (8)$$

2.2.6 Synchrosqueezed Wavelet Transform (SSWT)

SSWT addresses the limitations of the traditional wavelet transform by employing a reassignment technique based on the instantaneous frequency, denoted as $\omega_x(a, b)$ to achieve a more precise localization of frequency components in time. This refinement enhances the concentration of the time–frequency representation, providing a significant advantage in the analysis of non-stationary signals. The SSWT produces a sharper and more compact depiction of the signal's time–frequency characteristics. The mathematical formulation of the SSWT is presented in Equation (9-10) [11].

$$\omega_x(a, b) = -j (Wx(a, b))^{-1} \frac{\partial Wx(a, b)}{\partial b} \quad (9)$$

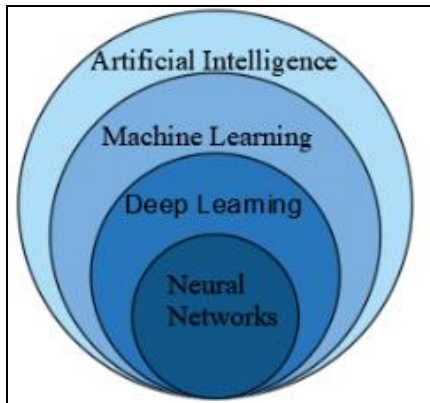
$$Tf(\omega, b) = \int Wx(a, b) \cdot a^{-\frac{3}{2}} \cdot \delta(\omega - \omega_f(a, b)) da \quad (10)$$

2.3 Machine Learning (ML)

ML methods are widely employed for classification tasks and for generating optimal predictions based on input data. In ML, algorithms learn underlying patterns or features from labeled or unlabeled datasets to predict unknown outputs. ML represents a fundamental subfield of Artificial Intelligence (AI) is a system capable of performing tasks that require human intelligence. The relationship between AI and ML is illustrated in Figure 3.



Figure 3: Comparisons of artificial intelligence, machine learning, deep learning and neural network.

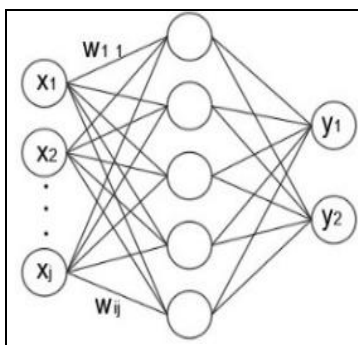


Source: Authors.

2.3.1 Artificial Neural Networks (ANN)

ANN is computational models inspired by the structure and function of the human neuron. As shown in Figures 4, an ANN consists of an input layer, one or more hidden layers, and an output layer. Each neuron processes inputs by applying weights, biases, and an activation function to produce an output. During training, the network adjusts these parameters iteratively by using the backpropagation to minimize the error between the predicted and actual outputs. This enables ANNs to learn patterns, make predictions, and effectively solve complex nonlinear problems. The mathematical formula of the model is provided in Equation (11) [12].

Figure 4: Artificial neural networks.



Source: Authors.

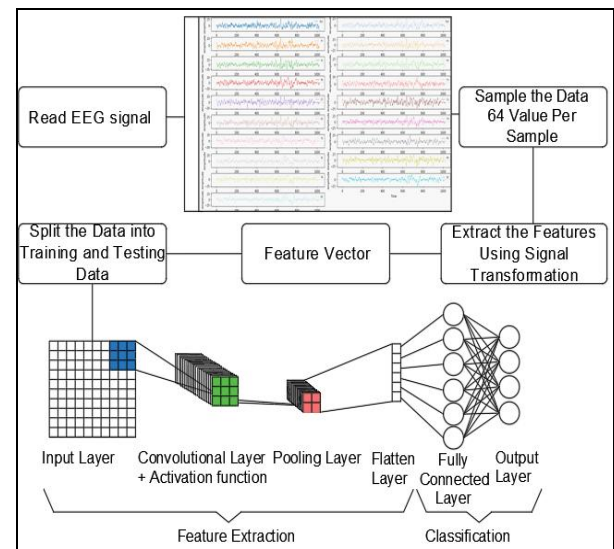
$$y_i = \sum_{j=1}^n x_j \cdot w_{ij} + b_i \quad (11)$$

2.3.2 Convolutional Neural Networks (CNN)

CNN is a deep learning model specialized for analyzing structured grid data such as images and

signals. It uses convolutional layers with sliding filters (kernels) to automatically detect spatial features. As shown in Figures 5 CNN diagram includes convolutional, pooling, and fully connected layers that together produce the final output. Initially, the input data are represented as a matrix, denoted by I . A kernel matrix K is then convolved with the input to generate the output matrix Y . The mathematical formula of this operation is presented in Equation (12) [13]

Figure 5: Convolutional neural networks.



Source: Authors.

$$Y(m, n) = (I * K)[m, n] = \sum_j \sum_t K(j, t) \cdot I(m - j, n - t) \quad (12)$$

3. Results and Discussion

After extracting features from EEG signal by using signal transformation techniques, CNN model was used for the diagnosis of Alzheimer's disease. The model begins with an input vector of size $64 \times \text{EEG channel number}$ and includes two convolutional layers: the first containing 64 neurons and the second 32 neurons. Each layer employs a 3×3 filter, the ReLU activation function, and 2×2 max pooling to reduce dimensionality and enhance feature extraction. The architecture concludes with a 32-neuron fully connected layer followed by a 2 neuron output layer for binary classification (AD vs. Healthy). The dataset was divided into 80% for training and 20% for testing. Model performance was evaluated by using a confusion matrix along with metrics such as accuracy sensitivity, specificity,



precision, and F1-score. The classification results are shown in Table 1, while the corresponding confusion matrices are illustrated in Figures 6-13. Additionally, the comparison of training and validation accuracy as well as training and validation loss across the models is presented in Figures 4-21. These metrics collectively provide a comprehensive assessment of the CNN model's performance in recognizing Alzheimer's disease from EEG signals. For the computer simulations,

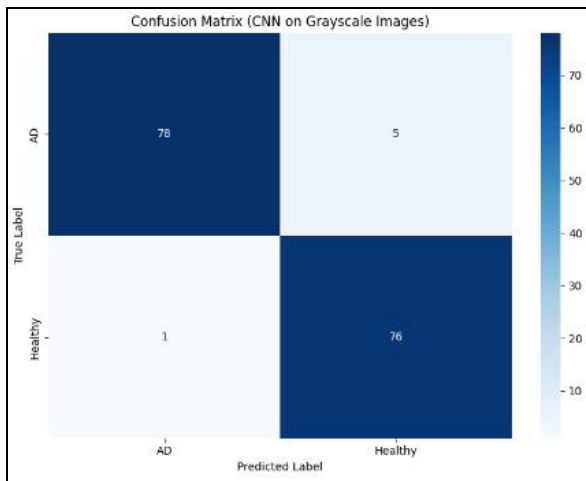
the Python programming language was used along with several scientific and machine-learning libraries, including pywavelets [14], ssqueezepy [15], matplotlib [16], scipy [17], pandas [18], seikit-learn [19], numpy [20], seaborn [21], and tensorflow [22].

Table 1: Classification result.

	Accuracy	Sensitivity	Specificity	Precision	F1score
Raw Data	96.25	96.34	96.34	96.28	96.2
FFT	91.46	91.35	91.35	91.50	91.42
STFT	99.39	99.35	99.35	99.43	99.39
CWT	98.12	98.19	98.19	87.78	87.02
DWT 1D	87.2	86.81	86.81	87.78	87.02
DWT 2D	98.78	98.77	98.77	98.77	98.77
SSFT	96.88	96.8	96.8	97	96.86
SSWT	95.73	95.68	96.68	95.75	95.71

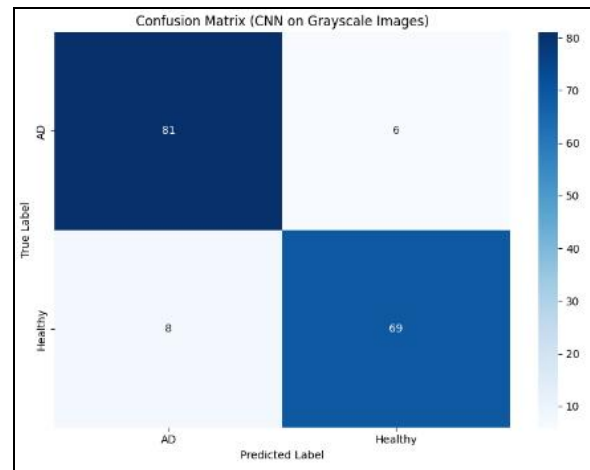
Source: Authors.

Figure 6: Raw data confusion matrix.



Source: Authors.

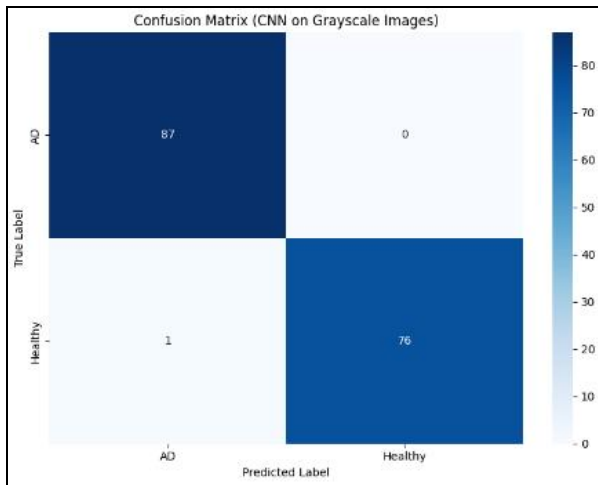
Figure 7: FFT confusion matrix.



Source: Authors.

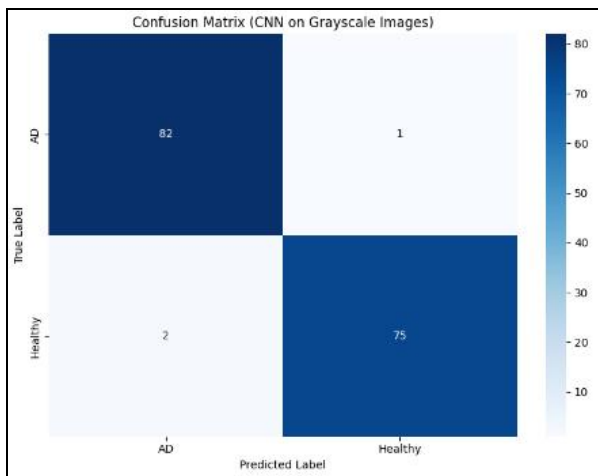


Figure 8: STFT confusion matrix.



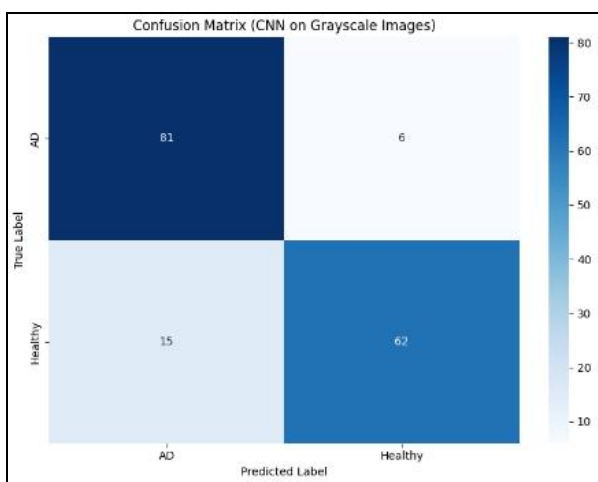
Source: Authors.

Figure 9: CWT confusion matrix.



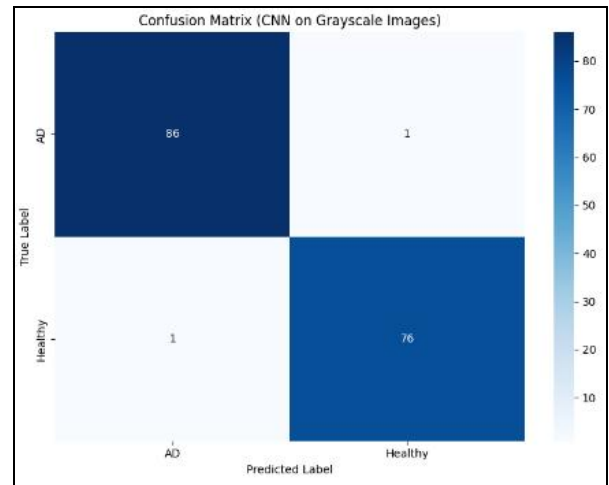
Source: Authors.

Figure 10: DWT 1D confusion matrix.



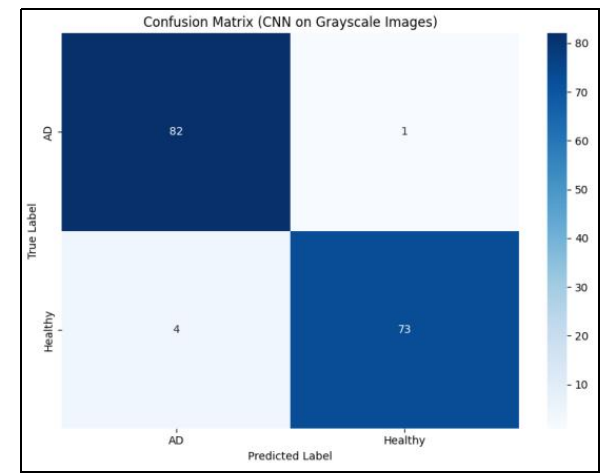
Source: Authors.

Figure 11: DWT 2D confusion matrix.



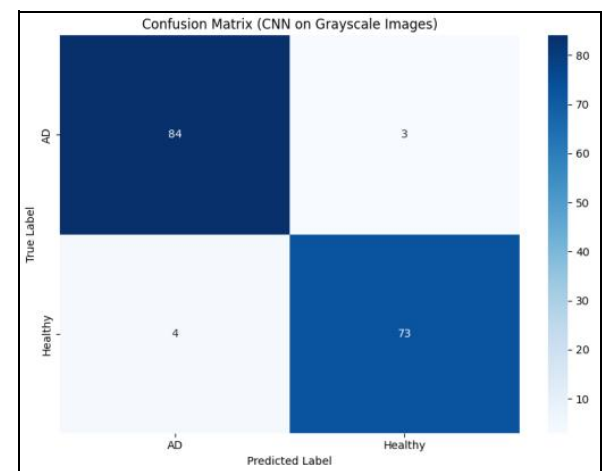
Source: Authors.

Figure 12: SSFT confusion matrix.



Source: Authors.

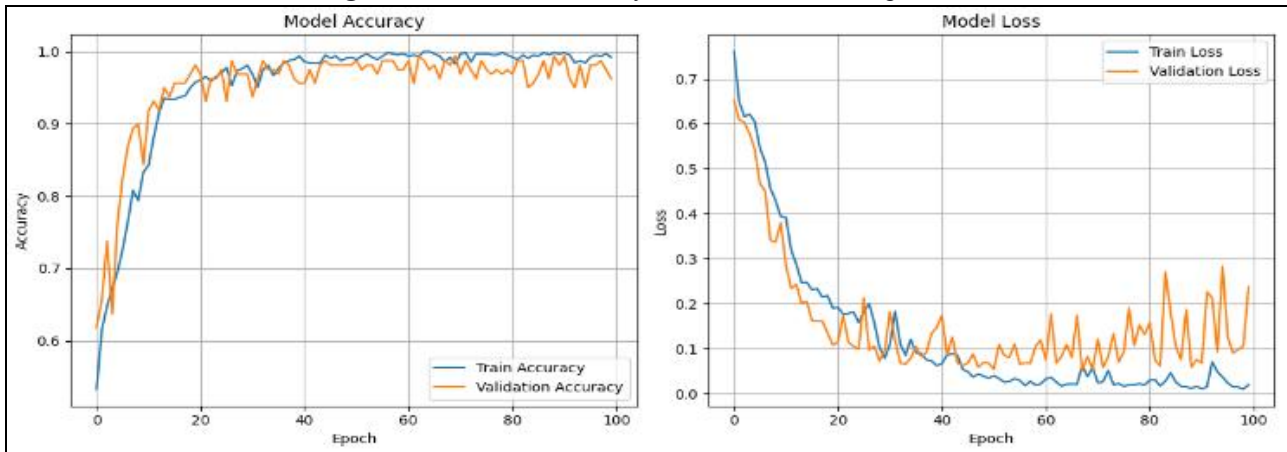
Figure 13: SSWT confusion matrix.



Source: Authors.

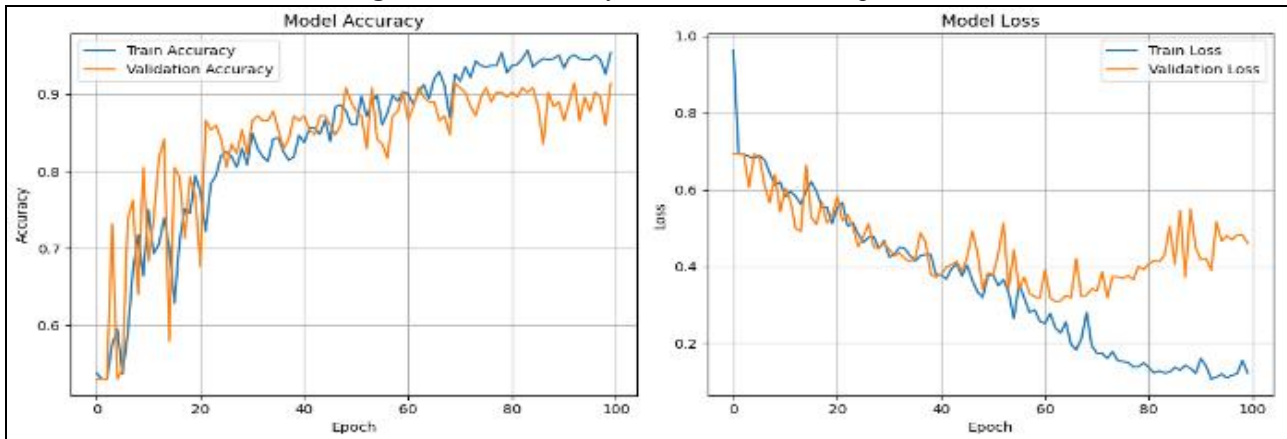


Figure 14: Raw data accuracy validation values comparison.



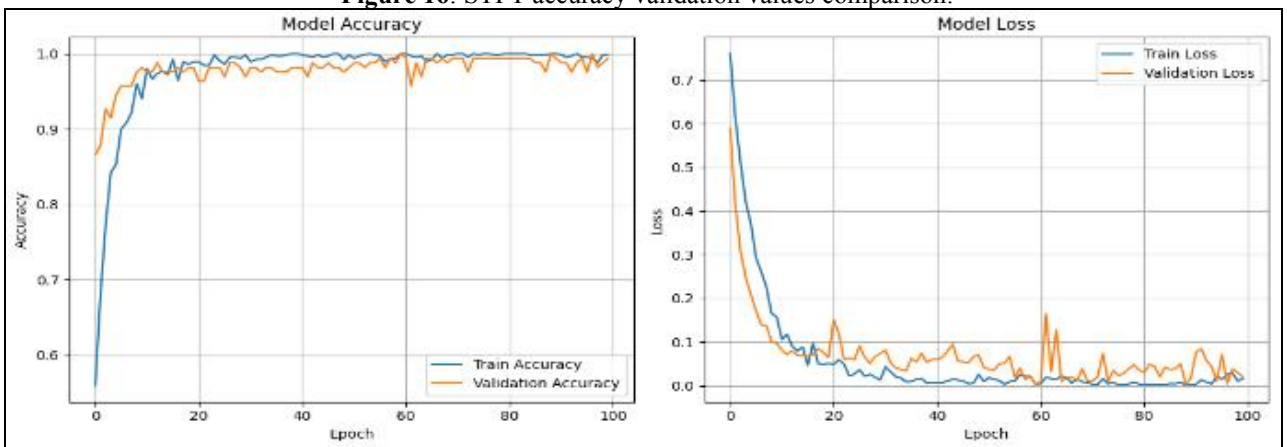
Source: Authors.

Figure 15: FFT accuracy validation values comparison.



Source: Authors.

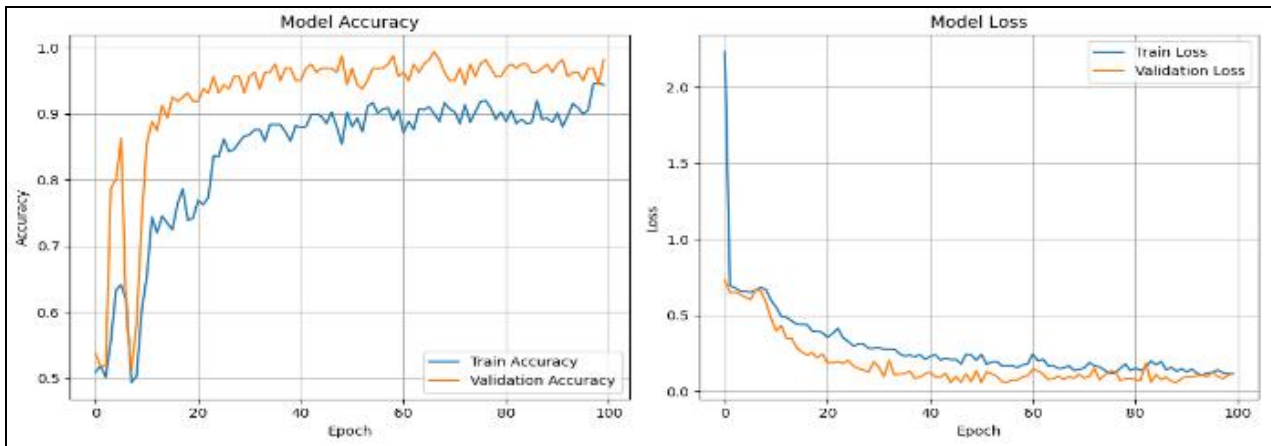
Figure 16: STFT accuracy validation values comparison.



Source: Authors.

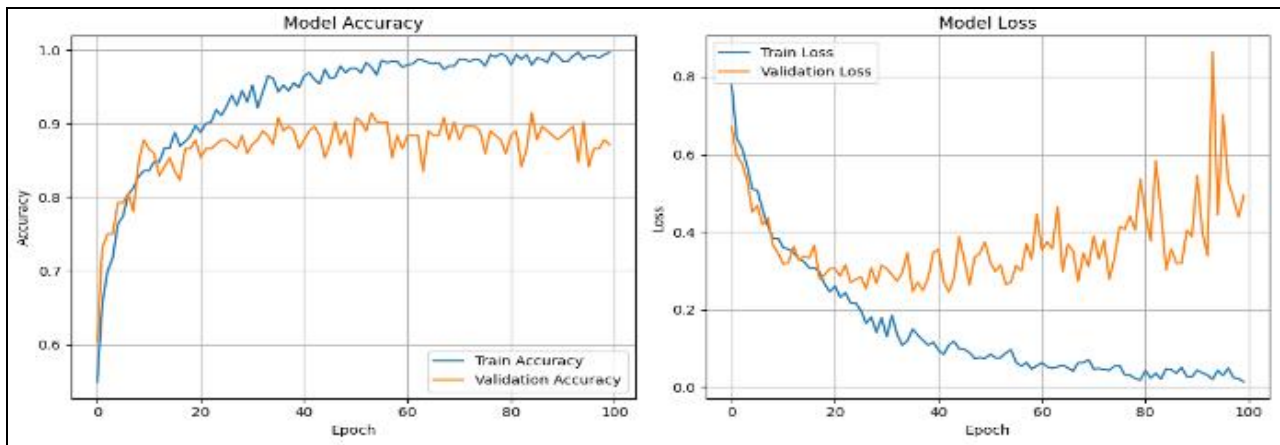


Figure 17: CWT accuracy validation values comparison.



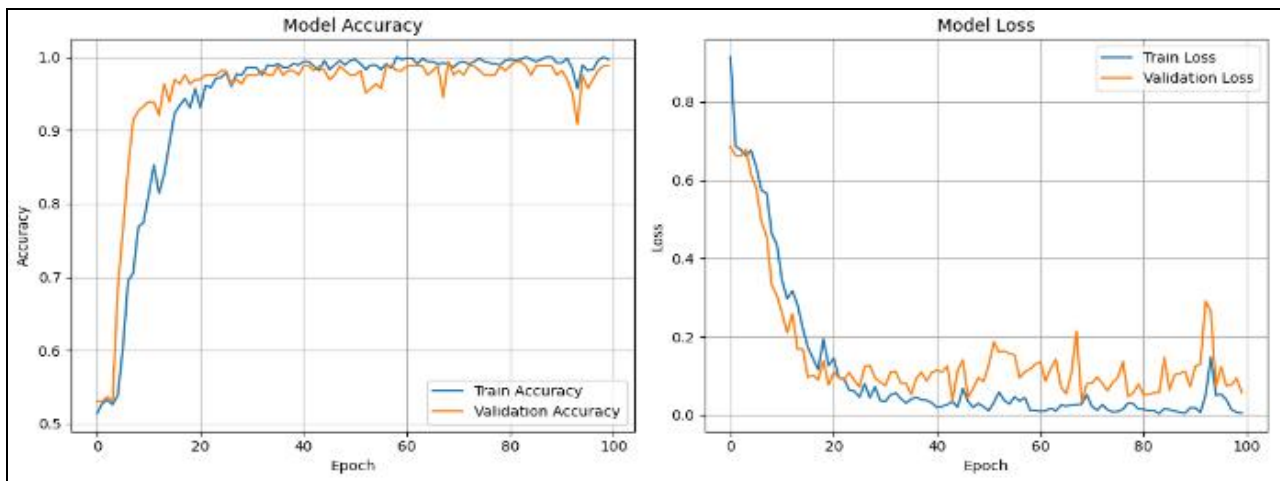
Source: Authors.

Figure 18: DWT 1D accuracy validation values comparison.



Source: Authors.

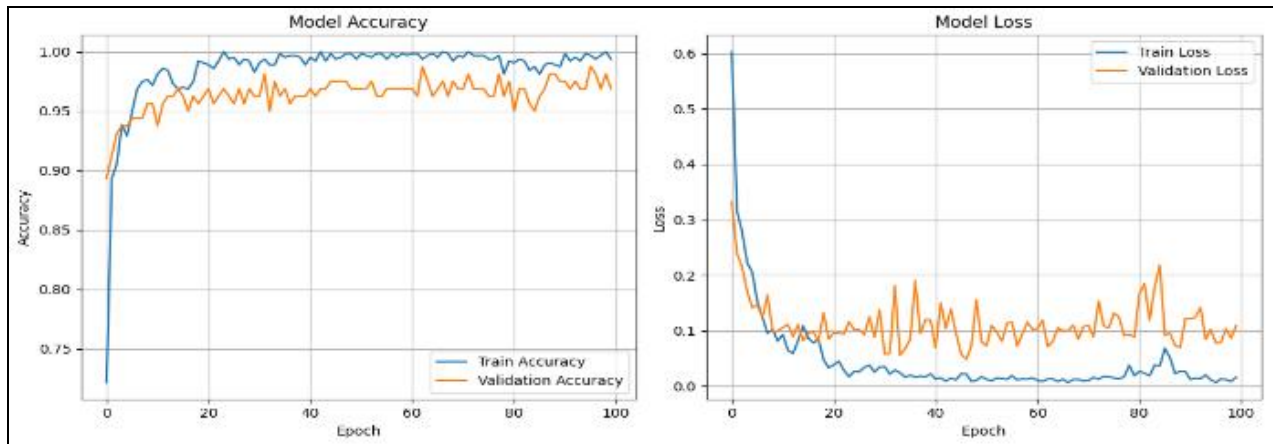
Figure 19: DWT 2D accuracy validation values comparison.



Source: Authors.

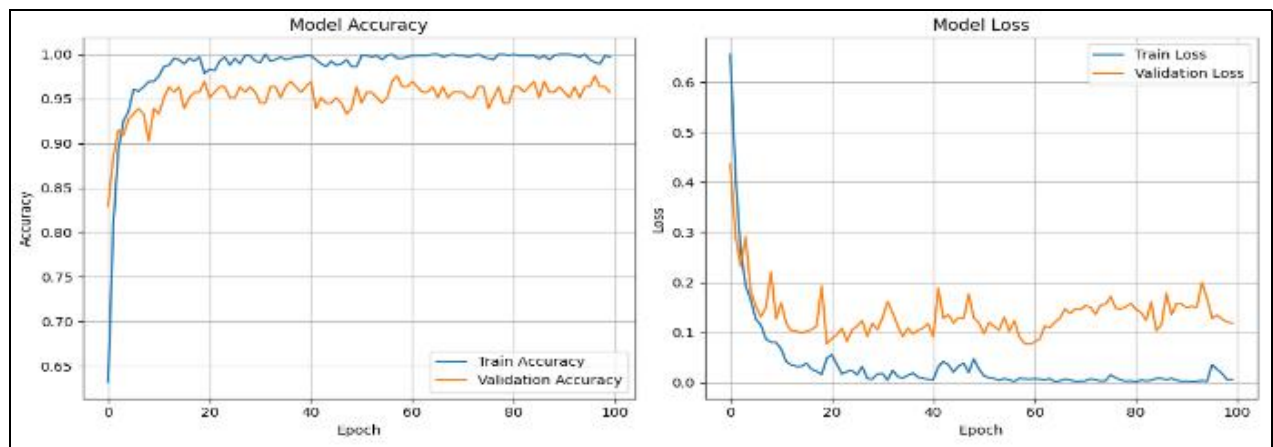


Figure 20: SSFT accuracy validation values comparison.



Source: Authors.

Figure 21: SSWT accuracy validation values comparison.



Source: Authors.

4. Conclusion

Alzheimer's disease detection was achieved by extracting features from EEG signals using a signal transforms and classifying the result with a CNN model. In this study FFT, STFT, CWT, DWT, SSFT and SSWT were evaluated. Consequently, STFT appears to be the most effective algorithm for Alzheimer's disease detection, with the highest training and testing accuracy, precision, sensitivity, specificity, and F1-score.



References

- [1] M. Amini, M. M. Pedram, A. Moradi, and M. Ouchani, "Diagnosis of alzheimer's disease by time-dependent power spectrum descriptors and convolutional neural network using eeg signal," *Computational and Mathematical Methods in Medicine*, vol. 2021, no. 1, p. 5511922, 2021.
- [2] H. Göker, "Detection of alzheimer's disease from electroencephalography (eeg) signals using multitaper and ensemble learning methods," *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, vol. 28, no. 1, pp. 141–152, 2023.
- [3] G. Fiscon, E. Weitschek, A. Cialini, et al., "Combining eeg signal processing with supervised methods for alzheimer's patients classification," *BMC medical informatics and decision making*, vol. 18, no. 1, p. 35, 2018.
- [4] Z. Aslan, "Eeg sinyallerini kullanarak alzheimer hastalığının otomatik tespiti için bilgisayar destekli tanı sistemi," *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, vol. 13, no. 2, pp. 213–220, 2022.
- [5] C. L. Alves, A. M. Pineda, K. Roster, C. Thielemann, and F. A. Rodrigues, "Eeg functional connectivity and deep learning for automatic diagnosis of brain disorders: Alzheimer's disease and schizophrenia," *Journal of Physics: complexity*, vol. 3, no. 2, p. 025001, 2022.
- [6] M. Murugappan and S. Murugappan, "Human emotion recognition through short time electroencephalogram (eeg) signals using fast fourier transform (fft)," in *2013 IEEE 9th International Colloquium on Signal Processing and its Applications*, IEEE, 2013, pp. 289–294.
- [7] A. Zabidi, W. Mansor, Y. Lee, and C. C. W. Fadzal, "Short-time fourier transform analysis of eeg signal generated during imagined writing," in *2012 international conference on system engineering and technology (ICSET)*, IEEE, 2012, pp. 1–4.
- [8] G. Thakur and H.-T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM Journal on Mathematical Analysis*, vol. 43, no. 5, pp. 2078–2095, 2011.
- [9] N. Bajaj, "Wavelets for eeg analysis," *Wavelet theory*, vol. 11, 2020.
- [10] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.
- [11] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and computational harmonic analysis*, vol. 30, no. 2, pp. 243–261, 2011.
- [12] C. Gershenson, "Artificial neural networks for beginners," *arXiv preprint cs/0308031*, 2003.
- [13] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network-a deep learning approach," *Procedia computer science*, vol. 132, pp. 679–688, 2018.
- [14] G. R. Lee, R. Gommers, F. Wasilewski, K. Wohlfahrt, and A. O'Leary, "Pywavelets: A python package for wavelet analysis," *Journal of Open Source Software*, vol. 4, no. 36, p. 1237, 2019.
- [15] J. Muradeli, "Ssqueezepy," *GitHub Note* <https://github.com/OverLordGoldDragon/ssqueezepy/>, 2020.
- [16] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
- [17] P. Virtanen, R. Gommers, T. E. Oliphant, et al., "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [18] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, S. van der Walt and J. Millman, Eds., 2010, pp. 56–61.
- [19] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [20] C. R. Harris, K. J. Millman, S. J. van der Walt, et al., "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020.
- [21] M. L. Waskom, "Seaborn: Statistical data visualization," *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021.
- [22] Martín Abadi, Ashish Agarwal, Paul Barham, et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, Software available from tensorflow.org, 2015.