Journal of Universal Computer Science, vol. 31, no. 11 (2025), 1196-1221
submitted: 18/2/2025, accepted: 25/4/2025, appeared: 28/9/2025

DeepV-Net: A Deep Learning Technique for Multimodal Biometric Authentication Using EEG Signals and Handwritten Signatures

Ashish Ranjan Mishra
(Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP, 273010, India
https://orcid.org/0009-0000-8635-0447, ashish.mishra@recsonbhadra.ac.in)

Rakesh Kumar
(Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP, 273010, India
https://orcid.org/0000-0003-1125-0031, rkiitr@gmail.com)

Rajkumar Saini
(Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Luleå, Sweden
https://orcid.org/0000-0001-8532-0895, rajkumar.saini@ltu.se)

Abstract: Ensuring secure and reliable person authentication is a critical challenge in modern security systems. Traditional biometric systems relying on physiological traits like fingerprints, iris, and facial recognition often suffer from spoofing vulnerabilities. In contrast, electroencephalogram (EEG) signals, characterized by unique temporal and cognitive patterns, provide a robust authentication mechanism. This paper introduces DeepV-Net, a multimodal fully convolutional neural network that leverages both EEG signals and dynamic handwritten signature data acquired from Wacom devices. The proposed model integrates spatial and temporal features of EEG signals with distinctive movement-based signature patterns through an end-to-end multimodal fusion strategy. Experimental evaluations on benchmark datasets demonstrate that DeepV-Net outperforms unimodal approaches and state-of-the-art authentication methods, achieving a training accuracy of 99.1% and a validation accuracy of 93.3%. These findings highlight the complementary nature of EEG and signature modalities, paving the way for more secure and efficient biometric authentication systems.

Keywords: Person Authentication, EEG, Hand-Written Signature, Multimodal, DeepV-Net
Categories: C.4, D, D.1, I.2, J, L.2, L.4
DOI: 10.3897/jucs.150681

1 Introduction

In the current context of biometric authentication, it is crucial to investigate new and robust ways to address the changing security and user identity requirements. Electroencephalogram (EEG) signals, which record the brain's electrical activity, have become a potential method for person verification because of their inherent uniqueness [Iyer et al., 23 Yang et al., 22]. Furthermore, the human brain is equipped with the neocortex, which is comprised of 10 billion neurons. These coupled neurons are intricately tailored to human emotions and actions [Chaladar et al., 21]. Furthermore, EEG signals that are not linear exhibit causality and efficacy that change over time in a dynamic system [Triggiani et al., 17]. EEG signals may be created based on human cognition, mental capacity, and emotions. However, capturing spontaneous EEG signals within the 2 to 10uV range is challenging [Bertazzoli et al., 21]. Furthermore, the EEG signal often contains intricate noise, including white noise, spikes, and power frequency interference. An artifact is another term used to describe this sort of noise. The frequency domain characteristics of the EEG signal are particularly notable [Guan et al., 22].

User verification proves the asserted identity via contact with the system, while person identification entails creating unique credentials or characteristics within a system. The process of identifying a user from a larger set of biometric data is known as user identification in the biometrics field. Alternatively, user verification focuses on checking a person's stated identification against a single match [Mishra et al., 24]. Biometric systems have been developed using a broad range of physical and behavioral traits. These include; but are not limited to; the retina, iris, sclera, ear, hand, palm shape, fingerprints, face, and a variety of vocal and movement patterns. Signature-based biometric systems are favored for their ease, non-intrusiveness, affordability, widespread use, ability to be canceled, and legal acceptance. Forgery of signatures is a well-known drawback in systems that rely on signatures since proficient persons may replicate someone else's signature with different degrees of precision. EEG-based biometric systems provide strong protection against counterfeiting since it is inherently difficult to replicate an individual's distinct brain activity patterns. EEG waves are challenging to counterfeit, giving it a trustworthy and resilient means of verifying one's identification, unlike signatures or bodily biometrics. Although signatures are distinct to each person, they may still be counterfeited, highlighting the need for a more robust authentication method [Das et al., 21 Fidas et al., 23]. The system seeks to improve identification and verification performance by integrating EEG signals with distinctive characteristics, thus protecting the confidentiality of information and reducing the risk of impersonation. This work utilizes the interconnection between Wacom signature and EEG data to identify and verify users.

1.1 Motivation

Authentication mechanisms are at the core of a modern security system and can be challenged in many ways, attacking the integrity and reliability of the system and user convenience. Conventional unimodal biometric technology like fingerprint, facial recognition, or iris scanning is becoming more susceptible to spoofing attacks, synthetic identity fraud, and sensor tampering. Additionally, these systems tend to underperform in dynamic or uncontrolled settings, where factors like lighting, positioning, or user behavior may compromise precision. Though appealing, behavioural biometrics are also limited by high intra-user variability and ease of imitation. These challenges highlight the importance of adopting a multimodal strategy, incorporating several independent characteristics to improve the resilience and accuracy of the system. Combining EEG signals, which are inherently unique and difficult to forge with dynamic handwritten signatures is a promising solution in this context. In this context, we present DeepV-Net, a novel deep learning-based framework that integrates cognitive and behavioral biometric data to overcome the current systems' security limitations and lay the ground for more secure, scalable and user-friendly identity authentication solutions.

This research presents DeepV-Net, an innovative neural network structure developed for the purpose of human verification. It utilizes both unimodal EEG and Handwritten Wacom signature data and combines them in a multimodal framework. DeepV-Net expands on the V-Net framework, which has shown impressive achievements in medical picture segmentation assignments, and applies it to the field of biometric verification. DeepV-Net intends to develop a solid and accurate authentication system using EEG data's spatial and temporal features and the dynamic signature patterns recorded by Handwritten Wacom devices. The goal is to distinguish between persons [Saichand et al., 21 Li et al., 22] efficiently. This research has been performed on a dataset of seventy people captured using a 5-channel acquisition device [Mishra et al., 24].

The major contribution of this paper is as follows:

- For the purpose of person verification using EEG and Hand-Written Wacom signature data, the suggested DeepV-Net is a completely convolutional V-Net architecture.
- Develop and implement DeepV-Net models to automatically extract relevant features from EEG and Hand-Written Wacom data, aiming to improve the overall accuracy of person authentication compared to existing methods.
- The deep learning approaches are being developed to combine information from EEG and Hand-Written Wacom signature modalities in a multimodal authentication configuration.
- The performance of DeepV-Net on benchmark datasets is evaluated, showing its superiority over baseline approaches and emphasizing the advantages of multimodal authentication.
- Find the best model for EEG-based person authentication by investigating and comparing several deep learning models, such as Optimized Convolutional Memory Fusion (OCMC), Multimodal Mutual Information Fusion (MMIF), Multimodal Siamese Neural Network (mSNN), M-LSTM, and NeuroWave-Net.

The rest of the paper is organized as follows: Section 2 represents the comprehensive analysis of the literature; Section 3 delineates the materials and methodologies used. Section 4 examines experimental findings and their corresponding outcomes. Section 5 delineates the conclusion and prospective research.

2 Literature Review

There has been much exploration of BCI technologies in recent years, primarily because they provide intrinsic benefits over other techniques. Specifically, the acquisition of non-invasive EEG measurements has made it easier to capture brain-related [TajDini et al., 23 Yap et al., 23 Tatar et al., 23]. Physiological signals used in biometrics resist spoofing due to their independence from, or at least not exclusive reliance on, behavior. Research has shown that user authentication based on EEG may surpass offline signature-based methods [Gorur et al., 23 Kingsy et al., 22 Yadav et al., 22]. EEG characteristics from several wavebands (alpha, beta, and gamma) based on Shannon entropy have been used for user verification. This approach achieved a classification accuracy of 97.1% [Phung et al., 14]. EEG studies have efficiently used spatial and statistical parameters, such as mean, standard deviation, kurtosis, and skewness, respectively [Jayarathne et al., 16].

[Pham et al., 14] developed a person verification system that utilizes electroencephalogram (EEG) data and incorporates many factors such as age and gender. The system employs the Power Spectral Density (PSD) characteristic of EEG and a Support Vector Machine (SVM) classifier. They obtained classification accuracies of 97.1% and 96.7% when considering the characteristics of gender and age, respectively. [Mu et al., 16] created a method for verifying a person's identity using fuzzy entropy EEG signal characteristics and an ANN classifier. [Kumar et al., 17] developed a framework that utilizes statistical aspects of EEG to enhance the security of mobile devices. They found that the system achieved a global Half of Total Error Rate (HTER) of 25% and a local HTER of 2.01% when tested with 50 users.

[Khademi et al., 22] proposed a hybrid model that combines three modules: Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and a replica of CBI based on Mutual Information (MI). The three modules executed the ResNet-50 layer on the IV dataset and obtained accuracy values of 85.9%, 90%, and 91.7%. [Zhang et al., 18] suggested a hybrid model that combines the Auto-Regressive (AR) and Empirical Mode Decomposition (EMD) techniques to accurately categorize the levels of stress and emotion experienced by individuals. At first, the EEG signals are obtained from a standard database called the DEAP dataset, and then a process of extracting 2-dimensional features is performed. The emotional states are divided based on the retrieved characteristics into four categories: low arousal high valence, high arousal high valence, low arousal high valence, and low arousal low valence. However, the process of categorizing the emotion groups requires a more extended amount of time. The identification of epileptic seizures primarily relies on examining EEG signals to identify various bodily artifacts.

For use in real-time scenarios, [Hussein et al., 19] presented a model of a Deep Neural Network (DNN). Here, the LSTM is used to extract the top-level model from the EEG data. The most important elements from each EEG pattern are then extracted using a fully connected architecture and sent into the output softmax layer. [Gao et al., 20] created a Genetic Particle Swarm Optimization (GPSO) system that uses binary coding and an optimized Convolutional Neural Network (CNN) framework to detect emotional states accurately. In this study, we concurrently forecast three specific emotional states based on the EEG pattern. To enhance classification accuracy, the fitness function in the CNN layer is modified via swarm optimization. Furthermore, the primary objective of the emotion identification task is to provide accurate classification results. However, it is not well-suited for tasks that involve making many predictions.

[Nandy et al., 22] proposed the Intelligent Agent-based Bag of Neural Networks (IABoNN) model, which is designed to include both machine learning (ML) and deep learning (DL) networks in a reversed manner. This model has the ability to categorize brain signals based on individuals who are impacted by stress. In this case, the remote system has an artificial intelligence module that regularly monitors the EEG signal. The experimental examination yielded accuracy values ranging from 91% to 95%. With their colleagues, Omer Kasim and Mustafa Tosun have devised a machine learning-based technique to remove artifacts from EEG signal segmentation efficiently. This strategy has been successfully implemented and proven effective. Additionally, a Butterworth filter eliminates low-frequency noise from the EEG data. As a result, the performance of Mean Square Error (MSE) and Normalized Mean Square Error (NMSE) was negatively impacted without a filtered EEG input. The method of filtering the artifact signal segment may be readily conducted.

The current protocol fusion method incorporates user-specific and generic patterns into a single dataset by combining EEG data acquired from several elicitation procedures [Debie et al., 21]. Several EEG biometrics research studies have employed this technique to address group variance, particularly when using supervised learning models. An essential part of these techniques is training the model using a training dataset. The training procedure and the quality of the training dataset are as important as the model's inherent capabilities in determining the model's efficacy [Sahoo et al., 23 Landau et al., 20]. Table 1 represents the comparison of recent multimodal biometrics authentication approaches.

References	**Multimodal**	**Methods**	**Metric**	**Advantages**	**Limitations**
A. El_Rahman et al., 24	EEG and Fingerprint	Parallel and Sequential Fusion	Accuracy, F1-Score	High AUC with sequential fusion, robust to spoofing	Required separate acquisition for hardware ECG
Salturk et al., 24	Static and Dynamic (Face and Air Signature)	Feature Level Fusion using CNN, LSTM, GRU	Accuracy, F1-score, AUC, and Sensitivity	Cost effective, no special sensors, work with simple webcams	Limited dataset size, complex environmental variations
Khademi et al., 22	Motor Imagery and MI-based EEG	CNN and LSTM Hybrid Method	Accuracy, F1-score	Good temporal modelling	Limited to motor imagery EEG

Byeon et al., 24	Unspecified Real Biometrics	Deep Neural Network Intelligent Method	Classification Accuracy	Highly flexible fusion level and scalable,	Relies on virtual datasets, need validation on real biometrics data
Vatchala et al., 25	Face and voice signature	Hybrid CNN and RNN Method	Emotion classification accuracy	Balanced trade-off between shared and modality specific layers.	Unspecified dataset details, possible complexity in model training
Sumalatha et al., 25	ECG, Fingerprint, and Finger Knuckle Print	Siamese Neural Network Method	Accuracy	High accuracy, robust against spoofing	Require multiple sensors, system may be complex in real time implementation
Proposed Method	EEG and Handwritten Signature	Modified V-Net with SE & AG	Accuracy, Precision, Sensitivity, Specificity	End-to-end fusion, high robustness, scalability	Requires multimodal hardware setup

Table 1: Comparison of recent multimodal biometrics authentication approaches

3 Materials and Methods

Our methodology proposes a DeepV-Net model for person authentication using unimodal and multimodal data, encompassing EEG signals and Hand-drawn Wacom signature data. For unimodal and multimodal, data processing undergoes three activities: (a)Thinking of the signature image by closing the eye or mental imagery: Mental imagery, or the construction of a sensory experience without actual stimuli, is central to this paradigm. People practice mental imaging when they shut their eyes and visualize the signature. This paradigm is the easiest to grasp because it demands the least mental and physical processing power. The primary emphasis is placed on creating a mental representation in the imagination. (b) Drawing the signature trajectory mentally by closing the eye or motor imagery: Visualization is similarly fundamental to this paradigm, but it calls for more advanced mental operations like motor planning and spatial reasoning. The individual has to see the signature in their mind's eye before they can practice drawing it. Possible increased activity in regions of the brain responsible for motor control and planning. Looking at this paradigm via a more conventional lens would reveal it to be a motor imagery job. (c) Gaze openly while physically creating a signature. In this model, visual perception and real motor execution are crucial. A combination of visual perception, motor planning, and execution is required to create the signature. Additional areas of the brain that are involved in visual perception, feedback regulation, and sensorimotor integration may be activated by this paradigm. The paradigm calls for integrating visual and motor systems and using higher-order cognitive functions, including planning, attention, and working memory.

3.1 Dataset Acquisition

This section comprehensively describes the data-collecting procedure, including the equipment, people, recording setup, experimental paradigm, and acquisition technique. The dataset [Mishra et al., 24] creation process included 70 enrolled people, and their identities were anonymized. The study included a total of 70 volunteers who were in good health. Out of these, there were five females and seventy men. The age range of the participants was between seventeen and thirty-six years. None of the participants had any neurological diseases or had undergone any brain-related drugs or surgery. All of the volunteers were inexperienced in EEG. Following the data collection, the participants individually assessed their level of satisfaction, boredom, unpleasantness, and tranquillity in relation to their experience. The average ratings (on a scale of 1 to 10) were 9.56, 1.43, 1.17, and 9.50, respectively. Every participant provided their informed written permission papers. Every individual involved in the activity was fluent in Hindi from birth. None of the participants had prior experience with EEG experiments or recordings, and they participated in a recording session lasting around one hour. The individuals in this study are denoted by aliases "subject-01" to "subject-70". Before donning the EEG Emotiv Insight 2.0 headset, the user provided their signature on a written permission form. Figure 1 depicts the primary experimental configuration used for data collection.

Although this study used this dataset to investigate the performance aspects of multimodal biometric authentication systems, it does show some biases and limitations that must be addressed. Most participants were young healthy males (65 out of 70 participants), aged 17 to 36. Hence, the generalizability of the findings to different genders, ages, and those with neurodegenerative or cognitive impairment was limited. Moreover, all subjects were native Hindi speakers and had not previously participated in EEG-based studies, which is expected to impact the uniformity of cognitive engagement during data collection. In addition, data collection occurred in a controlled, quiet environment, which, although ideal for minimizing noise, may not accurately reflect the real-world operational environments with distractions, movement, and differing environmental conditions. This limitation should be considered while interpreting the findings of this study, as future studies include diverse demographic and environmental settings to reinforce the validity and generalizability of the proposed authentication framework.

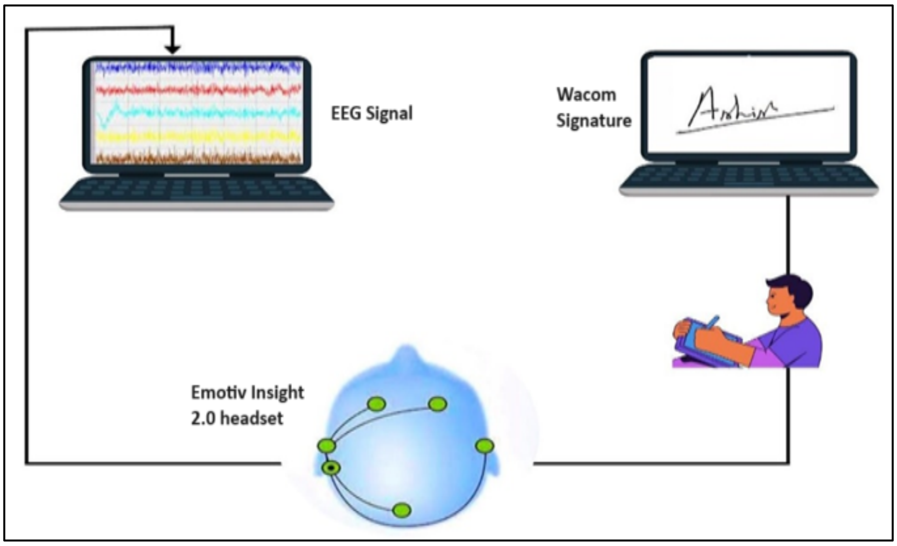

Figure 1: Data collection method architecture

The experimental setup to collect multimodal biometric data is shown in Figure 1. The study utilized an Emotiv Insight 2.0 electroencephalograph (EEG) headset worn by participants to capture real-time brain activity as they engaged in four tasks. A Wacom tablet also captured dynamic signature data (stroke pressure, direction, and pen position). The devices receive precisely timed auditory instructions guiding the participants between mental imagery and physical signature tasks. This also confirms the sequential recording of both EEG and corresponding signature data.

3.2 Data Preprocessing

The EEG data was pre-processed using EEGLAB, following the EEG data preparation pipeline shown in Figure 2. Epoching is the process of dividing the continuous EEG data into smaller parts in order to facilitate further analysis. The EEG data collected using the acquisition methodology described in Section 2.5 underwent further processing before being analyzed for signal processing. Around 30 incidents, including real and fake attempts, were documented for each individual. EEG activity associated with rest and signatures was isolated by detecting inserted markers with a marker value of 4. The length of each EEG activity associated with the first six markers was 10 seconds, which is comparable to 1280 frames, using the EEG headset's sampling rate of 128 frames per second. The length of the previous EEG activity was contingent upon the timing of the subject's completion of the signature on the Wacom tablet.

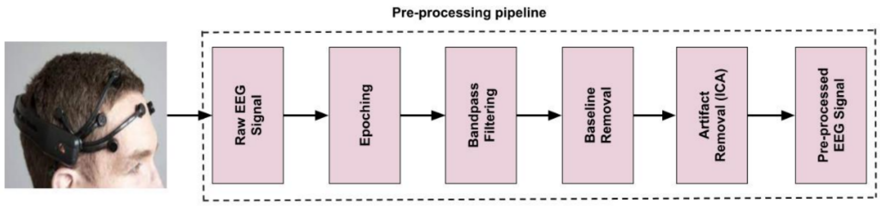

Figure 2: Architecture of EEG preprocessing pipeline

The EEG signal data entails a sequence of crucial processes to guarantee the integrity and quality of the recorded signals. At first, the unprocessed EEG data is collected from the recording device, following precise electrode positioning and signal capture procedures. Afterward, the ongoing EEG signal is divided into segments of a predetermined length using the technique of epoching, which allows for the separation of specific temporal events for study. After dividing the data into epochs, a bandpass filter reduces the strength of frequency components that are not within the required range while keeping the critical brain activity intact. Subsequently, baseline removal procedures eradicate any direct current offset or gradual change in the signal, facilitating a more precise depiction of event-related potentials. ICA is used to eliminate artifacts in the EEG data. The decomposed EEG epochs are carefully analyzed to detect and eliminate sources of noise, such as eye blinks, eye movements, or muscular activity. The EEG data is efficiently purified and made ready for further analysis by undergoing these preprocessing procedures, guaranteeing strong and dependable outcomes in the following processing tasks.

3.3 Model Architecture

For person authentication utilizing both unimodal and multimodal EEG and Wacom signature data, the proposed DeepV-Net architecture is designed to integrate information from multiple modalities. DeepV-Net comprises separate pathways for processing unimodal EEG and Hand-Written Wacom signature data, each consisting of convolutional layers followed by a multimodal fusion layer. This fusion mechanism effectively combines complementary information from both modalities, enhancing the discriminative models for authentic tasks. Considerations include the complexity of the model, computational efficiency, and the ability to handle multimodal inputs.

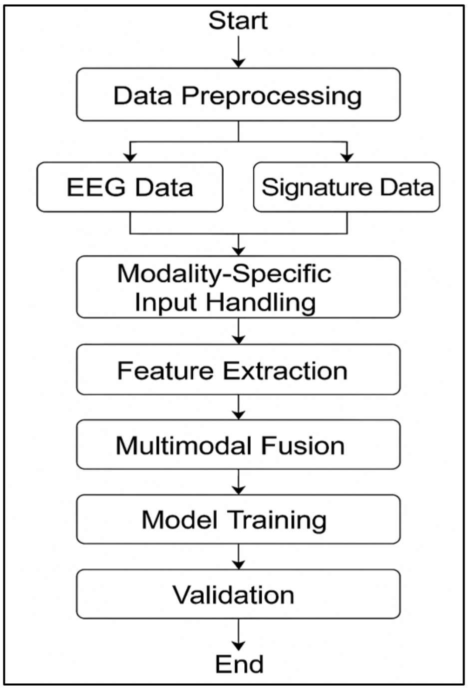

Figure 3: Flowchart of the training and validation process

Figure 3 provides an overview of the entire training and validation pipeline of the proposed DeepV-Net for multimodal biometric authentication. The process starts by data preprocessing where the raw EEG signals and signature date are filtered, cleaned, and segmented to eliminate noise and artifacts. The preprocessed data are split into two branches: EEG data and signature data. Then, the model extracts spatial-temporal patterns of the EEG signals and dynamic characteristics of the signatures with the deep convolutional layers to the network. Then the extracted feature is fused with a multimodal fusion strategy to learns the complex intermodal relations. The fused representation is brought to the model training stage and the classification objectives are used to optimize the network. Afterwards, our trained model is evaluated by the validation phase with respect to its performance in accuracy, precision, sensitivity, and specificity metrics. This defined pipeline provides a complex yet a black box model for user access control using heterogeneous biometrics.

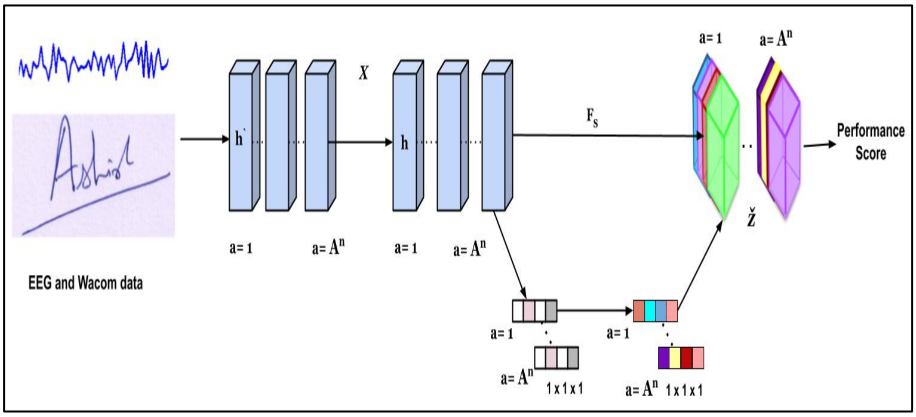

Figure 4: Proposed architecture of DeepV-Net with Squeeze-Excitation module

V-Net [Li et al., 20], integrated with the Squeeze-Excitation module, can improve it by emphasizing informative features, as shown in Figure 4. In the Squeeze phase, global spatial information is extracted from the feature maps. For each feature map, global average pooling is applied to generate a channel-wise descriptor, compressing spatial dimensions to a single value per channel. The excitation phase involves learning channel-wise dependencies to recalibrate feature responses. It is accomplished by evidently modelling the dependency between the channels.

$$\mathcal{O}_a = \sum_{i=1}^{\tilde{a}} S_a^i \times X_s^i \qquad (1)$$

According to Equation (1), the input is X, where $X \in R^{\tilde{\mathcal{D}}\times\tilde{\mathcal{W}}\times\tilde{\hbar}\times\tilde{a}}$.The units of measurement for depth, height, width, and number of channels are $\mathcal{D}$, $\hbar$, $\mathcal{W}$, and A, respectively. The result is represented as U, where $\mathcal{O}$ is a subset of $R^{\mathcal{D}\times\mathcal{W}\times\hbar\times a}$. The S_c^i operation is a spatial convolution in three dimensions, where the S_c denotes that each channel influences the feature of the associated channel.

The Squeeze operation is the first one that the feature U goes through, as seen in Equation (2). The steps include reducing the feature sets across the spatial gradient along with merging their resulting feature maps into a single set with $\mathcal{W} \times \hbar$ dimensions. It is from these feature mappings that the feature descriptor is formed. Currently, the actual number is more closely aligned with the global receptive field to some extent. This occurs because each feature channel in three-dimensional space is converted into a scalar value, which then responds to the overall distribution throughout the feature channel. As a result of this action, the input of $\mathcal{W} \times \hbar \times A$ is converted into an output of $1 \times 1 \times A$.

$$d_a = \sum_{i=1}^{\mathcal{D}} \sum_{j=1}^{\mathcal{W}} \sum_{k=1}^{\hbar} \mathcal{O}(i,j,k) \qquad (2)$$

As seen in Equation (3), a parameterized gating mechanism using two completely linked layers is used to mitigate the complexity and enhance the specificity of the model. It also illustrates the operation of a fully linked layer, where $\mathcal{W}_1, d$ is the mathematical representation of this action. The dimension of $\mathcal{W}_1$ is given by the

Cartesian product of the set of complex numbers C and the set of complex numbers raised to the power of m. The variable "m" represents a scaling parameter in this context. The value of m was determined as 4 by empirical means. The objective of the parameter seeks to decrease the number of channels, hence resulting in a reduction in computational operations. Subsequently, by passing the input via a Rectified Linear Unit (ReLU) layer, the output dimension stays unaltered, followed by a multiplication operation with the weight matrix $\mathcal{W}_2$. The act of multiplying with $\mathcal{W}_2$ may be classified as a completely linked layer action. Subsequently, the parameter s is produced by using the Sigmoid function.

$$\mathcal{K} = \varphi(\mathcal{W}_2\theta(\mathcal{W}_1 d)) \quad (3)$$

The combined sequence is recalibrated and the final output is achieved using Equation (4).

$$\tilde{z}_c = F_S\ (\mathcal{O}_a, \mathcal{K}_a) \quad (4)$$

Among the variables under consideration, X represents a collection denoted as $[x_1, x_2, . . , x_a]$, whereas $F_S\ (\mathcal{O}_a, \mathcal{K}_a)$represents the corresponding channel between the scalar and the feature map.

The attention block combines Attention Guided Filter (AG) module, the guided EEG signal, and Wacom signature filtering. Feature maps of different resolutions may have their spatial and structural information integrated using the Attention Guided Filter. The feature maps with lower resolution and those with more excellent solutions are filtered out to achieve this. Figure 4 illustrates the schematic of the Attention Block and the attention map that was generated throughout the calculation, with x and y standing for the inputs to the attention-guided filter. When it comes to this strategy, Attention Block is of the utmost importance. It effectively addresses the issue of the backdrop affecting the foreground and also brings the foreground into sharp focus while reducing the background. The feature maps x and y are subjected to a linear transformation using convolution with a channel size of 1 × 1 × 1. After that, the two feature maps that have been converted are combined with the ReLU layer using element addition. Finally, a 1 × 1 × 1 is used. There is another linear transformation applied to the convolution, and the sigmoid is the one that is used most often to activate the final attention feature map. The diagrammatic representation of the outcomes of the AG module may be seen in Figure 5. The y and $\tilde{O}$ represent the guided and filtered feature map make up the input, as mentioned in Equation (5). In contrast, the high-resolution feature map Ọ is the end product of the combined $x$and y operations.

$$Output\ (\tilde{O}) = \sum_{i \in \mathcal{W}_m} \mathcal{W}_{ij}\ (x).\, y_j \quad (5)$$

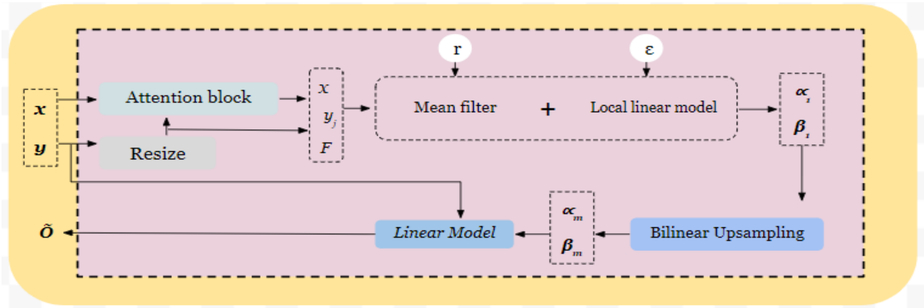

Figure 5: Architecture of Attention Guided Filter

In the first step of the process, the guided feature map I is down-sampled to generate a low-resolution feature map y_i. This feature map y_i is comparable to the map y same size. Afterward, to acquire the coefficients α_1 and β_1 of the attention-directed filter, it is necessary to reduce the reconstruction error of x and y. After that, the high resolution formed by the attention filter Feature map $\tilde{O}$ may eventually be obtained by down-sampling α_m and β_m. The attention filter is one of them, and it is simply a particular window wk with a radius of r. Specifically, the attention guide filter will build a square window denoted by wk, and the radius of k at each place will be denoted by r. As a result of the final segmentation performance, we experimentally determined that r = 15 and ε = 0.1 should be used in our research. (α_m, β_m) is also the sole definite constant coefficient, as shown in Equation (6), which is utilized to compute the least reconstruction error. Ridge regression with a standard term is employed in this calculation.

$$\min_{\alpha_m,\beta_m} Error(\alpha_m, \beta_m) = \sum_{i\in w_m}(F_i^2(\alpha_m y_i + \beta_m - x_i)^2 + \varepsilon\alpha_m^2) \tag{6}$$

3.4 Loss Function

For a person authentication task utilizing neural networks, a common choice for the loss function is categorical cross-entropy, especially when dealing with multiple classes. Consequently, a substantial amount of research has been conducted on the development of appropriate loss functions for tasks involving the segmentation of EEG signals. The mathematical expression for categorical cross-entropy loss is as follows:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log\left(p_{ij}\right) \tag{7}$$

In the above Equation (7), N represents the total number of samples and C represents the total number of classes. p_{ij} represents the sample predicted i belongs to class j probability and y_{ij} is a binary indicator. Models are penalized by categorical cross-entropy loss when there is a discrepancy between the ground truth labels and the projected probability distribution. It improves the model's ability to classify samples accurately, giving a high probability of the correct categories and a low probability of the incorrect classes.

4 Experimental Results and Discussion

4.1 Training Setup

The Experiment was conducted in a soundproof, air-conditioned room and a quiet environment. Participants were seated in comfortable chairs and made aware of the experimental procedure by explaining all the steps. The user signed a written consent form before placing the EEG Emotiv Insight 2.0 headset. We used two computers, one for recording the brain signals through Emotiv Pro software while performing the given activities and storing the EDF and CSV files generated by the software on the Emotiv cloud. Another computer connected by Wacom Tablet is dedicated to keeping the various signature characteristics (XY Coordinate, Pressure, Azimuth, Altitude, etc.) by using the Wacom Signature Scope Software. Every individual participated in one recording comprising four activities. These activities include sitting in the rest/relax position with closed eyes, visualizing the signature image with closed eyes, visualizing the signature trajectory with closed eyes, and performing the signature on a Wacom tablet with a Wacom pen with open eyes. A relaxation activity is performed at the beginning and in between two other activities to relax participants. Participants are advised to follow the audio instructions. For the first seven seconds (t=0 to t=7 s), an audio instruction is given to individuals to sit in the rest position with closed eyes. For the next eleven seconds (t=7 to t=18 s), the individual performs the activities provided in the audio instruction (sit in the rest position with closed eyes). From t=18 to t=22 s, another audio instruction is given to visualize the signature image with closed eyes. The participants perform this activity for ten seconds (t=22 to t=32 s). From t=32 to t=36 s, individuals are instructed to sit relaxed with closed eyes. After listening to the instruction, participants relaxed for the next ten seconds (t=36 to t=46 s). Afterward, individuals are instructed to visualize the signature trajectory with closed eyes from t=46 to t=51 s. Participants perform the activity to visualize the signature trajectory from t=51 to t=6 s. For the next four seconds (t=61 to t=65 s), Individuals are given the audio instruction to sit in the rest position with closed eyes. Participants perform this activity for eleven seconds (t=65 to t=76 s). From t=76 to t=84 s, an audio instruction is given to open eyes and do the signature on the Wacom tablet with a Wacom pen. After t=84 s, participants opened their eyes and made their signature on the Wacom tablet. Separate sets for training, testing, and validation have been created from the dataset. Make sure the model learns a variety of characteristics by allocating a substantial percentage for training.

4.2 Performance Evaluation

An assessment of the DeevV-Net is suggested based on metrics such as sensitivity, specificity, accuracy, and precision with respect to both unimodal (EEG Signals) and multimodal (EEG and Hand-Written Wacom Signature) data. We evaluate the suggested deep learning model against state-of-the-art EEG authentication techniques including OCMF, MMIF, mSSN, M-LSTM, and NeuroWave-Net. We compare the effects of several deep learning architectures on the reliability of authentication. Because of this, the training process and test the model on the test dataset using specified classification metrics.

True Positive (T_{Pos}): The number of effectively positive EEG signals and Wacom data is referred to as T_{Pos}.

True Negative (*T_{Neg}*): T_{Neg} is the total number of correctly detected negative EEG signals and negative Wacom data points.
False Positive (*F_{Pos}*): The false positive rate (F_{Pos}) is the total number of false positives in EEG and Wacom data.
False Negative (*F_{Neg}*): F_{Neg} is the number of positive EEG and Wacom signals misclassified as negative.

The accuracy measure, which compares the anticipated labels to the actual labels, is described in Equation (8). A minimum of 0 and a maximum of 1 are the possible values for the accuracy score.

$$Accuracy = \frac{T_{Pos}+F_{Neg}}{T_{Pos}+F_{Pos}+T_{Neg}+F_{Neg}} \tag{8}$$

Where F_{Pos} stands for count false positive. F_{Neg} stands for quantity of false negative and T_{Pos} refers to count of true positive. The use of precision allows for the evaluation of classifiers. The precision measure in Equation (9) is determined by dividing the total number of true positives and false negatives by the number of true positives.

$$Precision = \frac{T_{Pos}}{T_{Pos}+F_{Pos}} \tag{9}$$

Sensitivity, as shown in Equation (10), refers to as true positive rate, measures how many positive examples are correctly categorized, for example. In Equation (11) we can see that the specificity measure is the number of true negatives divided by the sum of the false positive and true negative counts.

$$Sensitivity = \frac{T_{Pos}}{T_{Pos}+F_{Neg}} \tag{10}$$

$$Specificity = \frac{T_{Neg}}{T_{Neg}+F_{Pos}} \tag{11}$$

Using several pre-trained models including OCMC, MMIF, mSNN, M-LSTM, and NeuroWave-Net in conjunction with particular EEG datasets acquired by DeepV-Net is the main goal of this investigation. A training set with 70% of the data and a test set with 30% of the data have been created for each kind of data, whether it's unimodal or multimodal. Six pre-trained models have been developed for EEG signals, one for hand-written Wacom signatures, and one for multimodal data sets that combine the two types of inputs. Each of the pre-trained models achieved a specificity score higher than 0.989, demonstrating remarkable performance according to the multimodal data. Results from the first epoch show that DeepV-Net had the highest specificity score of 0.984, 0.961, and 0.999 for EEG signals, hand-written Wacom signatures, and multimodal data, respectively. However, in both the unimodal and multimodal datasets, the OCMC model achieved a specificity score lower than 0.99. Although there are models that perform better than the others, there isn't much of a difference between them. Also, the reported performance metrics have a low standard deviation, which proves that the proposed technique for user authentication is consistent and reliable. Several pre-trained models' accuracy, specificity, and sensitivity are shown in Figure

6, Figure 7, and Figure 8. Table 2, Table 3 and Table 4 shows the precision, sensitivity and specificity values for unimodal and multimodal data in training phase.

Models	Precision		
	Unimodal (EEG)	Unimodal (Wacom)	Multimodal (EEG+Wacom)
DeepV-Net	0.893 ± 0.010	0.798 ± 0.013	0.996 ± 0.003
NeuroWave-Net	0.815 ± 0.012	0.777 ± 0.014	0.887 ± 0.008
M-LSTM	0.778 ± 0.015	0.763 ± 0.015	0.883 ± 0.007
mSNN	0.764 ± 0.016	0.751 ± 0.016	0.863 ± 0.011
MMIF	0.785 ± 0.014	0.76 ± 0.015	0.872 ± 0.009
OCMC	0.814 ± 0.013	0.735 ± 0.017	0.858 ± 0.010

Table 2: Comparing two training model datasets for precision with 95% of Confidence interval

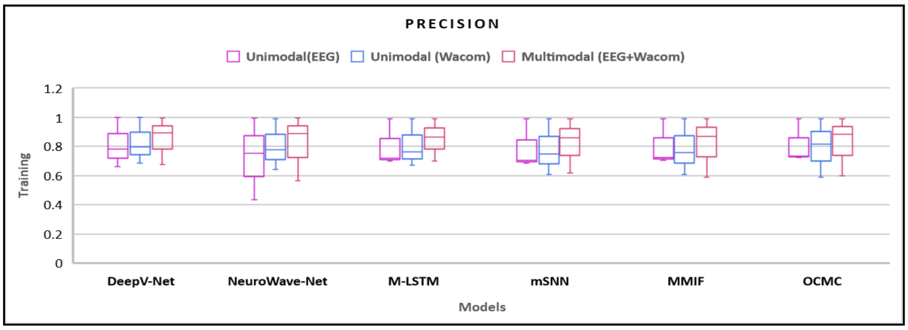

Figure 6: Comparison analysis of training models

Models	Sensitivity		
	Unimodal (EEG)	Unimodal (Wacom)	Multimodal (EEG+Wacom)
DeepV-Net	0.899	0.835	0.998
NeuroWave-Net	0.865	0.803	0.992
M-LSTM	0.852	0.771	0.992
mSNN	0.851	0.753	0.984
MMIF	0.842	0.739	0.967
OCMC	0.821	0.733	0.943

Table 3: Two datasets from different training models were used for sensitivity analysis

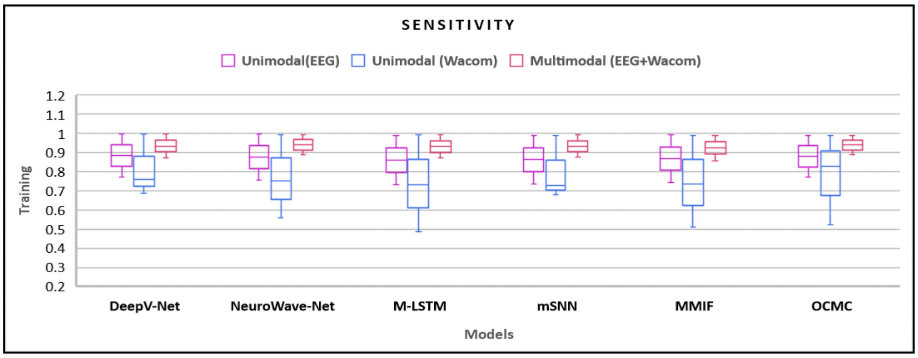

Figure 7: Comparison analysis of training models

Models	**Specificity**		
	Unimodal (EEG)	Unimodal (Wacom)	Multimodal (EEG+Wacom)
DeepV-Net	0.984	0.961	0.999
NeuroWave-Net	0.922	0.891	0.992
M-LSTM	0.916	0.886	0.989
mSNN	0.921	0.889	0.989
MMIF	0.911	0.91	0.985
OCMC	0.899	0.887	0.993

Table 4: Comparison of two training model datasets for specificity

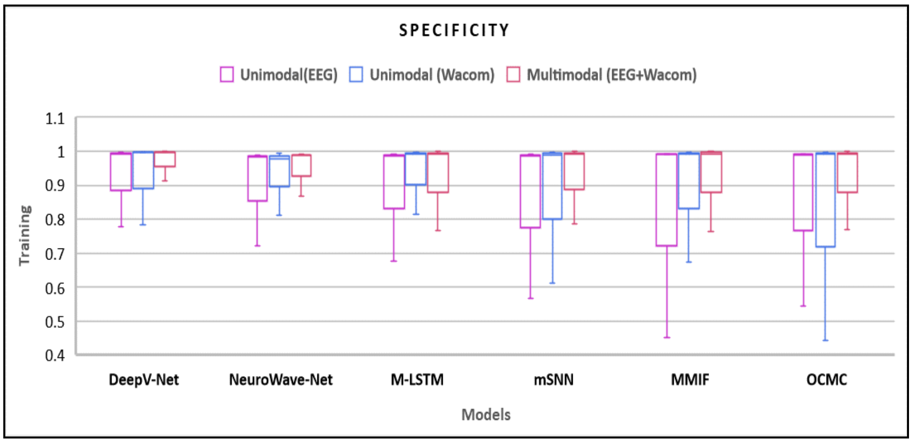

Figure 8: Comparison analysis of training models

During the validation phase, DeepV-Net with pre-trained models like mSNN, M-LSTM, and Neuro wave-net may be used to test the appropriateness using particular EEG datasets. Figures following show that both pre-trained models performed well, with validation specificity scores above 0.95 for each. Epoch 1 results show that on both unimodel and multimodal datasets, NeuroWave-Net and DeepV-Net attained the maximum specificity score of 0.99. In both datasets, the mSNN model achieved a specificity score of 0.982. Although there are models that outperform the others, there isn't much of a difference between the other versions. Figure 9, Figure 10, and Figure 11 show the results of the validation performance analysis for sensitivity, specificity, and accuracy for a number of pre-trained models.

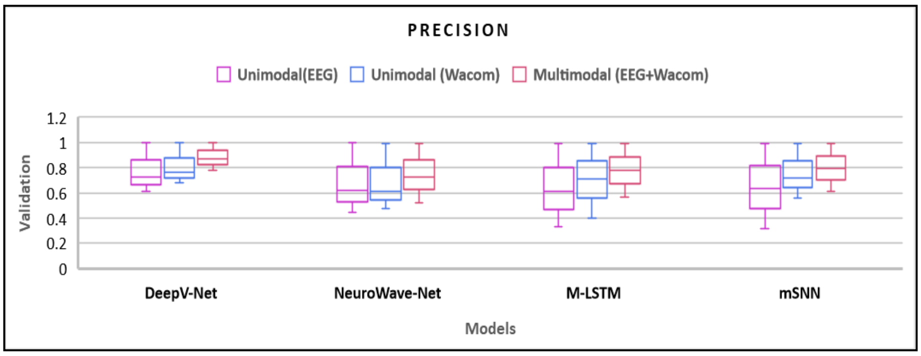

Figure 9: Comparison analysis validation performance of precision score

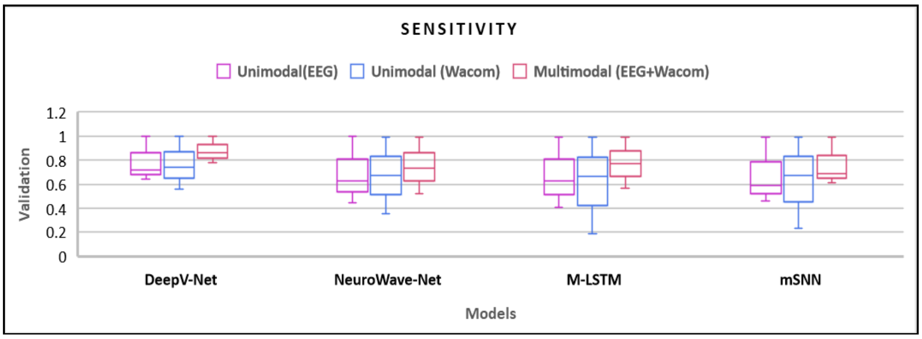

Figure 10: Comparison analysis validation performance of sensitivity score

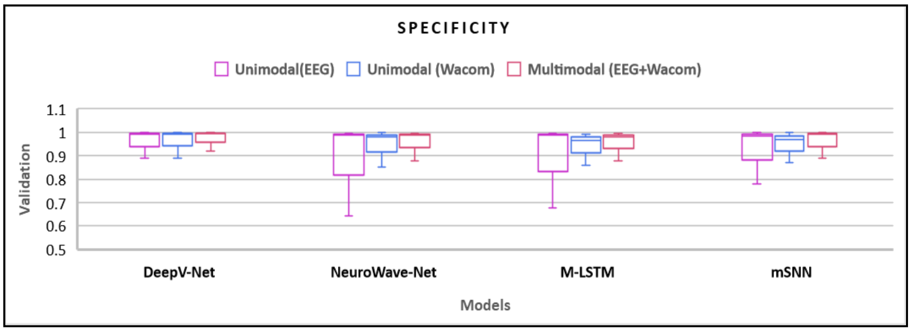

Figure 11: Comparison analysis validation performance of specificity score

Figure 12 shows a comparison of the DeepV-Net model used with the multimodal dataset and the one used with the unimodal dataset. Figure 12 shows the results of the accuracy test. For a 70% training data set, the paper details the testing accuracy attained by many methods, such as OCMC, MMIF, mSNN, M-LSTM, and NeuroWave-Net. While other models such as OCMC, MMIF, mSNN, M-LSTM, and NeuroWave-Net have reported accuracy values of 0.818, 0.796, 0.879, 0.884, and 0.928 for multimodal (EEG + Hand-Written Wacom signature) testing, DeepV-Net's stated accuracy is 0.987. When compared to previous approaches, the DeepV-Net technology outperforms them by 16.92%, 19.13%, 11.85%, 10.36%, and 3.91%, respectively.

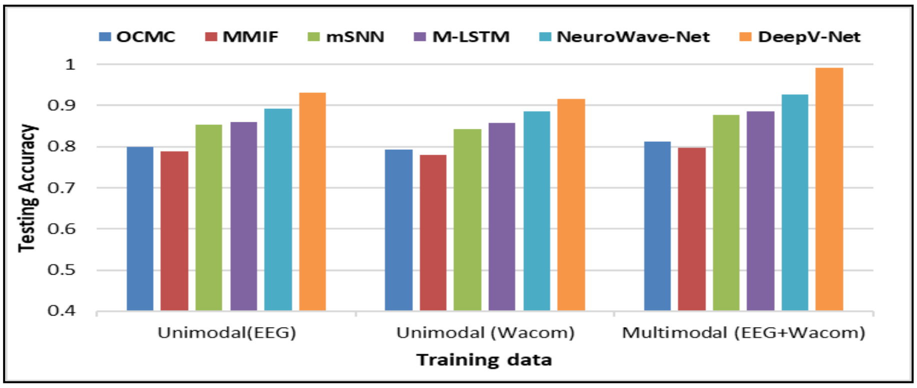

Figure 12: Comparison analysis average testing accuracy

Model	Modality	Precision	Sensitivity	Specificity	Accuracy
DeepV-Net	Multimodal (EEG-Wacom)	0.996	0.998	0.999	0.987
DeepV-Net	EEG	0.893	0.899	0.984	0.924
DeepV-Net	Wacom Signature	0.798	0.835	0.961	0.912

Table 5: Performance comparison of DeepV-Net with unimodal approaches

Table 5 represents the performance comparison of DeepV-Net with unimodal approaches. DeepV-Net not only achieves a high authentication accuracy but also exhibit an impressive computational efficiency compared to other deep learning-based models utilized in multimodal biometric systems. Our architecture makes use of optimized 3D convolutional layers and an end-to-end training pipeline, alleviating the requirement for elaborate hand-engineered features and fusion strategies external to the network. It provides up to 29% Top-1 accuracy gain without adding extra parameters and computational cost. Based on shared feature learning structure on both the EEG and signature modalities, DeepV-Net can further reduce redundancy in parameter updates and converge faster in the training phase. On the other hand, based on our experiments, DeepV-Net was trained within shorter time with lesser epochs to cost optimal performance than M-LSTM, mSNN, and MMIF. This provides a creative way of implementing the same functionality as those approximate methods, making it more feasible for practical applications in cases of real-time or resource-limited settings, where scalability and response time are crucial.

Configuration	**Precision**	**Sensitivity**	**Specificity**	**Accuracy**
Full DeepV-Net	0.996	0.998	0.999	0.987
Without Squezee and Excitation (SE)	0.961	0.969	0.982	0.954
Without Attention Guide Filter (AG)	0.948	0.955	0.977	0.947
Without Multimodal Fusion	0.893	0.899	0.984	0.925

Table 6: Ablation study on DeepV-Net components

Table 6 summarizes the performance of our network under ablation studies to determine the effect of each of the key components of DeepV-Net, each of which we progressively deactivate whilst monitoring changes in performance measures. The results provided in Figure 13 and Figure 14 clearly demonstrate the comparative benefit of the suggested strategy compared to other modern techniques, even when utilizing the most successful classifier. Testing on the multimodal dataset (EEG + Hand-Written Wacom signature) yielded these results. Using the pre-trained DeepV-Net on a dataset that includes the Inception module is the main reason to use this approach.

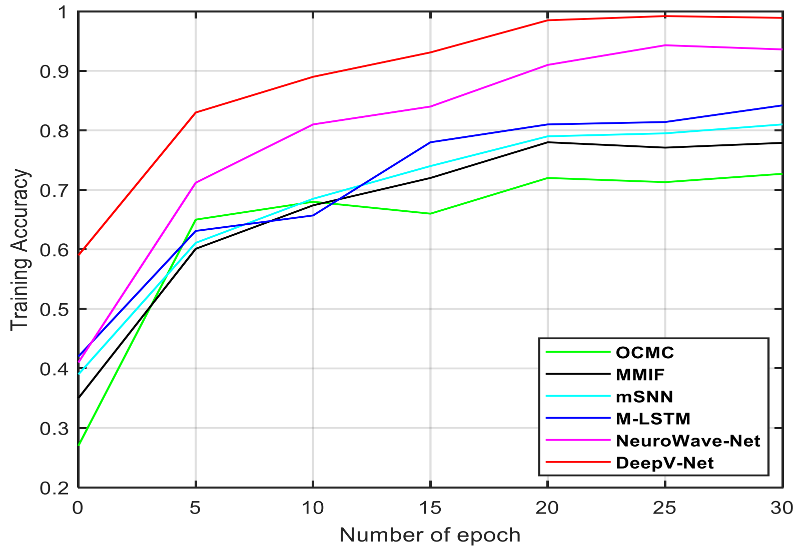

Figure 13: Comparison analysis of multimodal (EEG + Wacom signature) training accuracy

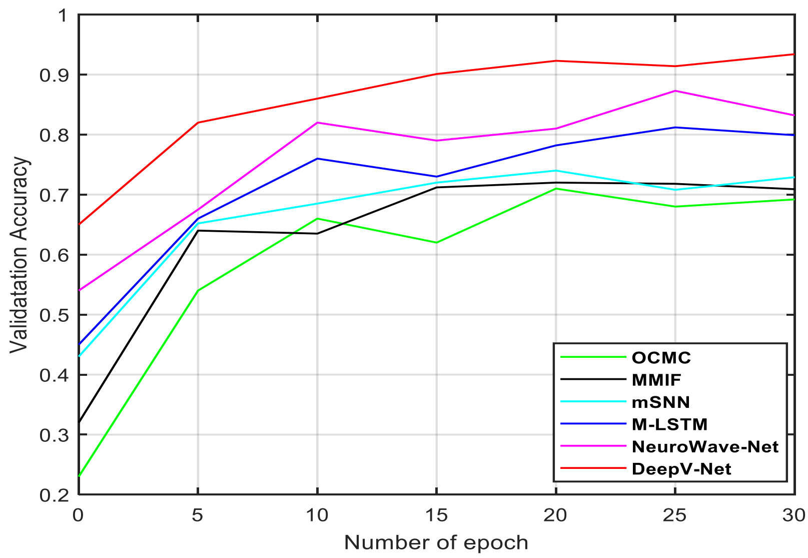

Figure 14: Comparison analysis of multimodal (EEG + Wacom signature) validation accuracy

Consequently, this method makes good use of the benefits offered by both components. An essential method for extracting basic traits is to use a pre-trained module. Concurrently, higher-dimensional feature extraction and categorization are the purview of the remaining layers. Figure 13 and Figure 14 demonstrate that the suggested strategy yields a validation accuracy of 93.3% and a training accuracy of 99.1% after 30 training epochs, respectively. This suggests that getting optimum outcomes for these models is still important, even when using pre-trained weights.

4.3 Discussion

The reason for the better outcomes of DeepV-Net lies in its rigorous end-to-end multimodal fusion scheme, which combines the complementary features of different modalities such as EEG signals and dynamic handwritten signatures. Utilizing spatial and temporal information within EEG data and complementing that with behavioral signature dynamics, this model offers a richer representation of the user identity with higher precision, sensitivity, and specificity. Additionally, components such as the Squeeze-and-Excitation (SE) module are employed to further tailor the model's focus on informative features, thereby enhancing discriminative capacity. While DeepV-Net has many strengths, it may not be ideal for every situation. Its performance may degrade when one of the its input modalities is noisy or missing, for example due to EEG data corrupted by motion artifacts, or poor contact quality with the scalp sensors. Moreover, the model suffers from the limitation of being trained under controlled laboratory conditions, thus experiencing reduced generalizability in real-world environments with varying environmental light, stress conditions, and distraction from the user itself. In addition, the high computational complexity of DeepV-Net might prevent deployment on low-resource or edge devices unless much more model compression or optimization is applied. These considerations underscore the potential and the practical limits of DeepV-Nets for biometric authentication.

As EEG-based biometric authentication has a stronger protection against spoofing attacks, it also brings important ethical and privacy implications that need to be addressed. EEG is a direct measure of brain activity, and as such, it may unintentionally expose sensitive information about a person's cognitive state, emotional responses, or neurological conditions that are separate from their identity. This poses the risk of misuse or unauthorized profiling without proper protocols in place. Furthermore, the aggregation and storage of such sensitive information require rigorous data privacy protocols, such as informed consent, secure encryption, and a transparent usage policy. Second, the ability to re-identify someone, or match them with other datasets, raises further autonomy and consent challenges.

DeepV-Net showcases impressive performance in controlled experimental environments; its application for biometric authentication in practical scenarios can be undermined by various concerns that require attention. The need to collect EEG signals and dynamic signature data at the same time requires specialized hardware (i.e., EEG headset and Wacom tablet), which incurs a high cost that may hinder its widespread adoption in the community and portability and user convenience issues. Moreover, EEG signals are vulnerable to environmental influence, like electrical disruption, user motion, and non-exact electrode positioning, impairing signal quality in uncontrolled environments. A major challenge is still to ensure uniform data collection across heterogeneous users and usage scenarios.

5 Conclusion

In this article, a novel approach for person authentication leveraging DeepV-Net, a fully convolution neural architecture, on EEG and Wacom signature data have been presented. The proposed method focused on integrating information from both unimodal and multimodal sources to enhance authentication accuracy. The DeepV-Net approach is demonstrated between genuine authentication attempts and baseline activities effectively. The results indicate that DeepV-Net, equipped with multimodal fusion capabilities, outperforms unimodal approaches regarding authentication accuracy and validation. By effectively combining spatial and temporal information from EEG signals with dynamic signature patterns captured by Wacom, the DeepV-Net model approach achieves superior performance across diverse authentication scenarios. The reported multimodal (EEG + Hand-Written Wacom signature) testing accuracy of DeepV-Net is 0.987, and other OCMC, MMIF, mSNN, M-LSTM, and NeuroWave-Net models' accuracy is 0.818, 0.796, 0.879, 0.884, and 0.928, respectively. The DeepV-Net method outperforms the state-of-the-art by 16.92%, 19.13%, 11.85%, 10.36%, and 3.91%, respectively. By the end of the 30 training epochs, the suggested approach has achieved an accuracy of 93.3% during validation and 99.1% during training.

Furthermore, the conclusion now addresses directions such as the inclusion of more biometric modalities, the deployment of lightweight models to edge devices, and the importance of federated learning for privacy-preserving authentication. In future research, our method will covered practical applications of the technology such as secure access control systems, online identity authentication, and continuous authentication in sensitive environments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

The data that support the findings of this study are available on request from the corresponding author, Ashish Ranjan Mishra, upon reasonable request.

References

[A. El_Rahman et al., 24] A. El_Rahman, S., & Alluhaidan, A. S. (2024). Enhanced multimodal biometric recognition systems based on deep learning and traditional methods in smart environments. Plos one, 19(2), e0291084.

[Bertazzoli et al., 21] Bertazzoli, G., Esposito, R., Mutanen, T. P., Ferrari, C., Ilmoniemi, R. J., Miniussi, C., &Bortoletto, M. (2021). The impact of artifact removal approaches on TMS–EEG signal. *NeuroImage*, *239*, 118272.

[Byeon et al., 24] Byeon, H., Raina, V., Sandhu, M., Shabaz, M., Keshta, I., Soni, M., Matrouk, K., Parkash Singh, P. & Lakshmi, T. V. (2024). Artificial intelligence-enabled deep learning model for multimodal biometric fusion. Multimedia Tools and Applications, 83(33), 80105-80128.

[Chaladar et al., 21] Chakladar, D. D., Kumar, P., Roy, P. P., Dogra, D. P., Scheme, E., & Chang, V. (2021). A multimodal-Siamese Neural Network (mSNN) for person verification using signatures and EEG. *Information Fusion*, *71*, 17-27.

[Das et al., 21] Das, K., &Pachori, R. B. (2021). Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals. *Biomedical Signal Processing and Control*, *67*, 102525.

[Debie et al., 21] Debie, E., Moustafa, N., &Vasilakos, A. (2021). Session invariant EEG signatures using elicitation protocol fusion and convolutional neural network. *IEEE Transactions on Dependable and Secure Computing, 19(4), 2488-2500.*

[Fidas et al., 23] Fidas, C. A., &Lyras, D. (2023). A Review of EEG-Based User Authentication: Trends and Future Research Directions. *IEEE Access*.

[Gao et al., 20] Gao, Z., Li, Y., Yang, Y., Wang, X., Dong, N., & Chiang, H. D. (2020). A GPSO-optimized convolutional neural networks for EEG-based emotion recognition. *Neurocomputing*, *380*, 225-235.

[Gorur et al., 23] Gorur, K., Olmez, E., Ozer, Z., & Cetin, O. (2023). EEG-Driven Biometric Authentication for Investigation of Fourier Synchrosqueezed Transform-ICA Robust Framework. *Arabian Journal for Science and Engineering, 1-23*.

[Guan et al., 22] Guan, X., Yang, G., Ye, J., Yang, W., Xu, X., Jiang, W., & Lai, X. (2022). 3D AGSE-VNet: an automatic brain tumor MRI data segmentation framework. *BMC medical imaging*, *22*, 1-18.

[Hussein et al., 19] Hussein, R., Palangi, H., Ward, R. K., & Wang, Z. J. (2019). Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals. *Clinical Neurophysiology*, *130*(1), 25-37.

[Iyer et al., 23] Iyer, A., Das, S. S., Teotia, R., Maheshwari, S., & Sharma, R. R. (2023). CNN and LSTM based ensemble learning for human emotion recognition using EEG recordings. *Multimedia Tools and Applications*, *82*(4), 4883-4896.

[Jayarathne et al., 16] Jayarathne, I., Cohen, M., &Amarakeerthi, S. (2016, October). BrainID: Development of an EEG-based biometric authentication system. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* (pp. 1-6). IEEE.

[Khademi et al., 22] Khademi, Z., Ebrahimi, F., &Kordy, H. M. (2022). A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals. *Computers in biology and medicine*, *143*, 105288.

[Kingsy et al., 22] Kingsy Grace, R., Geetha Devasena, M. S., &Manimegalai, R. (2022). BABW: Biometric-Based Authentication Using DWT and FFNN. Tele-Healthcare: *Applications of Artificial Intelligence and Soft Computing Techniques, 201-219.*

[Kumar et al., 17] Kumar, P., Saini, R., Roy, P. P., & Dogra, D. P. (2017). A bio-signal based framework to secure mobile devices. *Journal of Network and Computer Applications*, *89*, 62-71.

[Landau et al., 20] Landau, O., Puzis, R., & Nissim, N. (2020). Mind your mind: EEG-based brain-computer interfaces and their security in cyber space. *ACM Computing Surveys (CSUR), 53(1), 1-38.*

[Li et al., 22] Li, Q., Luo, Z., & Zheng, J. (2022). A new deep anomaly detection-based method for user authentication using multichannel surface EMG signals of hand gestures. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1-11.

[Li et al., 23] Li, F., Tan, C., Dong, F., & Jia, J. (2020). V-net deep imaging method for electrical resistance tomography. *IEEE Sensors Journal, 20(12), 6460-6469.*

[Mishra et al., 24] Mishra, A.R., Kumar, R., Gupta, V. et al. SignEEG v1.0: Multimodal Dataset with Electroencephalography and Hand-written Signature for Biometric Systems. *Sci Data 11, 718* (2024). https://doi.org/10.1038/s41597-024-03546-z.

[Mu et al., 16] Mu, Z., Hu, J., & Min, J. (2016). EEG-based person authentication using a fuzzy entropy-related approach with two electrodes. *Entropy*, *18*(12), 432.

[Nandy et al., 22] Nandy, S., Adhikari, M., Chakraborty, S., Alkhayyat, A., & Kumar, N. (2022). IBoNN: Intelligent Agent-based Internet of Medical Things framework for detecting brain response from Electroencephalography signal using Bag-of-Neural Network. *Future Generation Computer Systems*, *130*, 241-252.

[Pham et al., 14] Pham, T., Ma, W., Tran, D., Nguyen, P., & Phung, D. (2014, July). Multi-factor EEG-based user authentication. In *2014 International Joint Conference on Neural Networks (IJCNN)* (pp. 4029-4034). IEEE.

[Phung et al., 14] Phung, D. Q., Tran, D., Ma, W., Nguyen, P., & Pham, T. (2014, April). Using Shannon Entropy as EEG Signal Feature for Fast Person Identification. In *ESANN* (Vol. 4, No. 1, pp. 413-418).

[Sahoo et al., 23] Sahoo, R. C., Pradhan, S. K., Sahoo, B. M., &Balabantaray, B. K. (2023). Pattern recalling analysis of an auto-associative memory network using FFT and DWT. *Multimedia Tools and Applications, 82(6), 9113-9135.*

[Saichand et al., 21] Saichand, N. V. (2021). Epileptic seizure detection using novel Multilayer LSTM Discriminant Network and dynamic mode Koopman decomposition. *Biomedical Signal Processing and Control*, *68*, 102723.

[Salturk et al., 24] Salturk, S., & Kahraman, N. (2024). Deep learning-powered multimodal biometric authentication: integrating dynamic signatures and facial data for enhanced online security. Neural Computing and Applications, 36(19), 11311-11322.

[Sumalatha et al., 25] Sumalatha, U., Prabhu, S., & Nayak, V. C. (2025). Multimodal biometric authentication: a novel deep learning framework integrating ECG, fingerprint, and finger knuckle print for high-security applications. Engineering Research Express, 7(1), 015207.

[TajDini et al., 23] TajDini, M., Sokolov, V., Kuzminykh, I., & Ghita, B. (2023). Brainwave-based authentication using features fusion. *Computers & Security, 129, 103198.*

[Tatar et al., 23] Tatar, A. B. (2023). Biometric identification system using EEG signals. Neural *Computing and Applications*, 35(1), 1009-1023.

[Triggiani et al., 17] Triggiani, AI., Bevilacqua, V., Brunetti, A., Lizio, R., Tattoli, G., Cassano, F., Soricelli, A., Ferri, R., Nobili, F., Gesualdo, L., Barulli, MR., Tortelli, R., Cardinali, V., Giannini, A., Spagnolo, P., Armenise, S., Stocchi, F., Buenza, G., Scianatico, G., Logroscino, G., Lacidogna. G., Orzi, F., Buttinelli, C., Giubilei, F., Del Percio, C., Frisoni, GB., Babiloni, C. (2017). Classification of Healthy Subjects and Alzheimer's Disease Patients with Dementia from Cortical Sources of Resting State EEG Rhythms: A Study Using Artificial Neural Networks. Front Neurosci.10:604. doi: 10.3389/fnins.2016.00604.

[Vatchala et al., 25] Vatchala, S., Yogesh, C., Govindarajan, Y., Raja, M. K., Ganesan, V. P. A., Vinod, A. A., & Ramesh, D. (2025). Multi-modal biometric authentication: Leveraging shared layer architectures for enhanced security. IEEE Access.

[Yadav et al., 22]Yadav, Y., Chand, S., Sahoo, R. C., Sahoo, B. M., & Kumar, S. (2022). Comparative analysis of detection and classification of diabetic retinopathy by using transfer learning of CNN based models. *Journal of Intelligent & Fuzzy Systems, 43(1), 985-999.*

[Yang et al., 22] Yang, Z., & Huang, Y. (2022). Algorithm for speech emotion recognition classification based on mel-frequency cepstral coefficients and broad learning system. *Evolutionary Intelligence*, *15*(4), 2485-2494.

[Yap et al., 23] Yap, H. Y., Choo, Y. H., Mohd Yusoh, Z. I., &Khoh, W. H. (2023). An evaluation of transfer learning models in EEG-based authentication. *Brain informatics*, *10(1), 19.*

[Zhang et al., 18] Zhang, Y., Zhang, S., & Ji, X. (2018). EEG-based classification of emotions using empirical mode decomposition and autoregressive model. *Multimedia Tools and Applications*, *77*, 26697-26710.