Diyala Journal of Engineering Sciences

Journal homepage: https://djes.info/index.php/djes

ISSN: 1999-8716 (Print); 2616-6909 (Online)

A Privacy-Preserving E-Voting System using Federated Learning and CNNs for Secure Fingerprint and Biometric Verification

Wisam Ali Mahmood[1,2,*], Jumana Waleed[1] and Ayad R. Abbas[2]

[1]Department of Computer Science, College of Science, University of Diyala, Iraq
[2]Department of Computer Science, University of Technology- Iraq, Baghdad, Iraq

ARTICLE INFO

Article history:
Received August 14, 2024
Revised December 20, 2024
Accepted December 23, 2024
Available online September 01, 2025

Keywords:
Federated Learning
E-voting; Biometric Verification
Convolutional Neural Networks
Privacy-Preserving
Fingerprint Recognition
Gender Recognition
ID Verification

ABSTRACT

The increasing reliance on electronic voting systems introduces challenges in securing voter data and ensuring privacy. Traditional e-voting systems are prone to cyber threats, compromising the integrity of elections, including data leakage, tampering, and spoofing attacks, as well as the problem of data centralization that increases the risk of hacking and loss of electoral integrity. To address this, we propose a secure e-voting system that uses federated learning to enhance security and privacy. The system employs multiple Convolutional Neural Networks (CNNs) including VGG16, VGG19, ResNet18, MobileNetV2, a custom CNN, and a ResNet-VGG hybrid across six distributed sites. In order to verify that only eligible and registered voters may cast ballots, the proposed system uses a three-step biometric process that includes user ID identification, fingerprint matching for high security using unique biometric data to prevent identity fraud, and gender identification as an extra layer of verification to lower the risk of identity theft. Furthermore, the system trains models locally and only transfer trained weights to avoid centralized sensitive data. When compared to the Sokoto-Coventry fingerprint dataset, it achieved high rates of 99.75% for identity recognition, 99.90% for fingerprint recognition, and 99.97% for gender recognition. These results highlight the effectiveness of the proposed system in providing a secure, privacy-preserving, and scalable solution for e-voting.

1. Introduction

Electronic voting, popularly known as e-voting, is the newest form of voting. It provides several benefits relating to accessibility, speed in vote count, and reducing human errors. However, there exist a number of serious challenges and threats for an e-voting system that ought to be addressed so as to ensure it is effective and secure. Security is one major issue in e-voting. Nothing is more paramount in an election than the integrity and confidentiality of votes; a crime that directly influences even one vote undermines public confidence in the electoral process. Traditional e-voting systems face several cyber threats, including hacking attacks, malware infections, and denial-of-service attacks. These may result in the tampering of votes, unauthorized access to information about voters, or disruption of the voting process [1]. Maintaining the privacy of the voters is another major challenge. A voter should believe that his choices are confidential and identity not revealed or misused. While the e-voting systems can resolve various problems associated with traditional polling, they add some new issues, the most significant being the centralization of voter data makes it a prime

[*] Corresponding author.
E-mail address: wisam.a.alnadeem@uotechnology.edu.iq

DOI: 10.24237/djes.2025.18312

target for attackers [2]. While, in theory, e-voting would give more access to the process, it should be designed with all voters in mind and not forget people with disabilities or from non-dominant cultures who do not have easy access to technology. User-friendly electronic voting systems that work for a variety of diverse population groups can help in improving inclusive participation [3].

In e-voting systems, autonomy would mean that the system works all by itself securely and is not dependent on a central authority. This independence is vital in several ways. It increases the security of the voting process by diminishing these centralized points, which are the favorite targets of cyber attackers [4]. In this regard, a decentralized system is highly resistant to any entity seeking to compromise the data and control of the system. It will not compromise voters' privacy since there is no single repository, as all data remains stored in different nodes. This, therefore, can enhance voters' trust in the system because they feel they are dealing with a system that is controlled by none [5]. Further, autonomous e-voting systems can make the processes involved in voting more resilient and versed toward any kind of situation so that every section of society becomes vocal without a feeling of backwardness [6].

Although traditional e-voting systems mainly deal with the challenges associated with securing data transmission and preventing vote tampering, they often neglect other issues related to privacy, independence, and dealing with heterogeneous data. For example, the new challenge of centralization of voter data emerges, as storing this data in centralized repositories increases the risk of being targeted by attackers. Moreover, traditional systems suffer from weaknesses in dealing with the large diversity of biometric voter data, which affects the model’s ability to generalize. Their reliance on central authorities also makes them vulnerable to breaches that may lead to a loss of public confidence in the system. Hence, this research focuses on addressing these new challenges by introducing a decentralized system based on federated learning and CNNs, which enhances privacy and independence and ensures accurate and comprehensive handling of heterogeneous data.

Federated learning is among state-of-the-art techniques that can be applied to solve autonomy challenges in e-voting systems. FL allows training machine learning models across decentralized devices/servers with local data samples that can collaborate without actually sharing the original datasets. In other words, sensitive biometric data like fingerprints used in voter authentication never leave the local devices, hence ensuring privacy and security. By distributing the training process, FL decreases the risks associated with centralized storage and processing of data, hence making it hard for cyber attackers to attack. It can also enhance voting mechanisms by increasing the diversity of data sources in a manner that provides greater generalizability in the model. This learning paradigm aligns with autonomy, as each device and node is capable of contributing its experience independently to the training process without any central coordinating authority [7,8].

CNNs are very important in the e-voting system, particularly when dealing with biometric data for voter verification. Actually, research has shown that they are very efficient in image recognition and, hence, suitable for processing fingerprint images [9-12]. Federated learning provides a framework by which CNNs can be trained on fingerprint data from different locations and thus learn important features without putting all the fingerprints into the same location. Such a decentralized approach increases the privacy of biometric data and makes use of variations in datasets to improve generalization for a discriminative model. The incorporation of CNNs within the federated learning framework without affecting autonomy or security will provide high accuracy in voter impersonation, an e-voting system. It is stated that a combination of both is incorporated as great achievement in the development of safe, effective, and reliable e-voting software [13].

The following contributions illustrate the improvements introduced by our proposed system:

1. Enhanced Voter Identity Verification, which uses a three-step biometric method,

assures safe voter authentication and allows only eligible voters to participate.

2. Utilizing federated learning, decentralized model training improves data security and privacy by training models locally, storing sensitive biometric data on local devices, and only exchanging trained weights.
3. Several Convolutional Neural Network (CNN) models, such as VGG16 VGG19 ResNet18 MobileNetV2, a custom CNN, and a ResNet-VGG hybrid spread over six dispersed sites, are used in the Diverse Model Application to improve the system's recognition accuracy and resilience.
4. Establishing security, privacy, and integrity for the electronic voting process through the promotion of trust and transparency helps to increase voter confidence.

The paper is organized as follows: Section 2 discusses related research in e-voting systems, federated learning, and biometric verification. Section 3 outlines the technique, which includes distributed system and task allocation, federated learning implementation, CNN model use, model aggregation, and e-voting authentication. Section 4 summarizes the experimental findings and analyses the system's performance. Finally, Section 5 summarizes the findings and suggests areas for future investigation.

2. Related works

In this paper, some relevant works are discussed emphasizing privacy, security, and verifiability challenges. In the context of CNN, authors of [14] concentrated on CNN for fingerprint recognition and mentioned the top performance and accuracy. These findings underscore the need to incorporate high-contrast, low-noise fingerprints as valid criteria for recognition. This research uses the Sokoto Coventry Fingerprint Dataset to challenge CNN in its capability of classifying and recognizing fingerprints with an accuracy that is close to 99% when dealing with a clean dataset. CNN is a hierarchical model that has been described as a funnelled structure where fully connected layers with output are offset to achieve fine recognition results. On the other hand, the gap identified by their paper does not involve reliance on alternative authentication methods.

Similarly, A study by [15] proposes a public verifiable e-voting system, which can serve as a reference to prevent privacy and verify the vote. Paper ballots, which are most commonly used in elections, have been restricted by the following time spatial and transparency issues. The new system includes commitment schemes, zero-knowledge succinct non-interactive arguments of knowledge (zk-SNARKs), and machine learning in an attempt to make voting impossible for others, even with face recognition-based voter authentication. It can support sophisticated voting systems and is a private, secret, on-the-fly vote-counting system that scales for a broad range of elections. As a result, the gap identified in this paper is dependence on facial authentication, which may fail to recognize the user with major changes from surgery or medical treatment, create a problem of accessibility, and even restrict the system's inclusiveness.

Similarly, in [16], the authors conclude that a secure electronic voting system was presented that used fingerprint biometrics to authenticate a voter, thereby limiting each individual from coming in with someone else's ID. It was a web-based system that allowed voting online in a secure manner, making the process more accessible and efficient. Fingerprint recognition allows its users to strengthen its proof of security, making it almost impossible for unauthorized users to access or influence the election system. It updated the votes quickly, which makes election results more transparent and trustworthy. It is also built to be hack-proof, which will provide more reliability and security for the election process. The system was geared towards bringing a new face to the electoral processes, requiring little manpower and removing issues that come with traditional ballots, like vote rigging or duplication. The authors promote the superiority of including indelible biometric data in justifying that only a class (of genuinely competent) voters can take part in elections. However, the challenges highlighted include increased complexity and authentication difficulty, especially for older people.

Building on prior work, the authors in [17] leverages CNNs specifically for gender classification from fingerprint biometrics, achieving 96.47% accuracy on the SOCOFing dataset. The algorithm was built using Python 3.6, proving the efficiency of using a convolutional neural network for fingerprint-based gender classification in biometric systems. However, the limitation of this work is that it does not address the issue of data decentralization.

Furthermore, the research depicted in [18] proposed a deep Convolutional Neural Network (ConvNet), which predicts both gender and hand for prints. This is very important in forensic anthropology in that it relates to differentiation between the general criminal population and those fitting to a smaller degree with much-needed specificity. It gives very good validation accuracy for gender classification at 99.40% and a pretty well-done identification rate of 99.17% for hand identification. The model was tested on the publicly available dataset SOCOFing, which served as a benchmark previously in attempts to establish state-of-the-art performance in categorization techniques. Nevertheless, the gap included in this Paper did not address the concept of a data-distributed and decentralized implementation method.

The authors in [19] Concentrate on the crucial concern surrounding privacy protection in Blockchain-based Federated Learning (referred to below FL) systems, and as we have discussed this for being progressively applied to decentralized environments leading data security requirement. They address the amalgamation of blockchain technology to FL and its benefits in preserving data privacy while also retaining decentralized learning. It specifies some important privacy threats like data leakage and inference attacks, provides possible countermeasures-mainly by differential privacy-homomorphic encryption-secure multiparty competition. Their method reveals that blockchain-assisted FL systems work suitably when it comes to maintaining privacy, which is especially useful in critical areas like healthcare and Industry 5.0. These results highlight the value of layering strong privacy guarantees in a changing landscape for distributed machine learning. However, this paper's limitation is that it did not address the use of alternative methods of authentication.

Based on this approach, the authors in [20] showed an interdisciplinary smart Electronic Voting Machine (EVM) using IoT based on fingerprint biometric authentication. The system uses the R307 fingerprint module to authorize voters before they vote so that only authorized persons can enter their votes, ensuring secure and precise voting. This system enables real-time data communication using IoT technology and assures that the voting process takes place securely. The centralized architecture, including a gateway for secure storage of voters' fingerprints and deployment of Arduino-based microcontrollers, allows effective processing and storage of votes. The theory behind this method is to enhance the clarity, credibility, and security of e-voting machines. As a result, challenges in this paper include difficult accessibility interface and risk Data centralization storage.

3. Methodology

This work demonstrates the design of a secure and privacy-preserving e-voting online system using asynchronous six geographically distinct locations by state-of-the-art technologies as shown in Figure 1. Federated Learning distributed model training among all the sites, in a way that each site is enabled to train its models locally without sharing biometric information with a central server-only trained weights. It mitigates the risks associated with centralized storage of data while ensuring a high degree of accuracy with respect to authentication of voters. The e-voting system used three biometric ID verification methods: fingerprint ID identification, gender identification, and finger recognition. Fingerprints act as the primary and trusted biometric, while gender identification adds an additional layer of verification to detect impersonation attempts, and finger recognition enhances the accuracy of fingerprint matching by identifying the finger used. These collectively ensured that the voter established his or her identity as legitimate and valid before casting their vote. The dataset consisted of 600 users and was divided equally among the six geographically separated sites, with each site receiving data on 100 users.

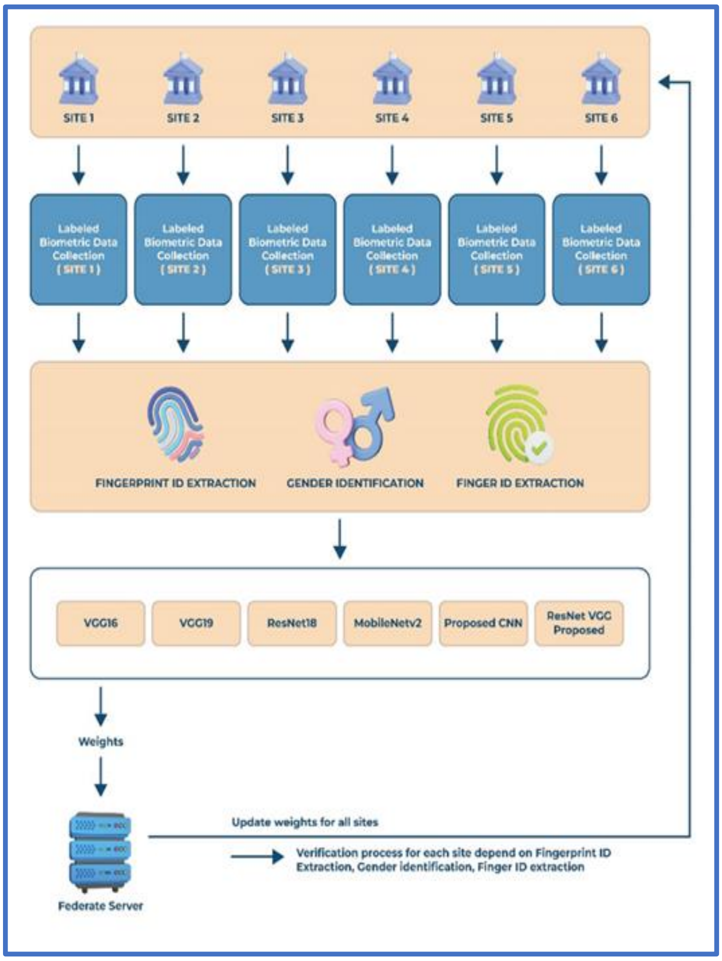
SITE 1
SITE 2
SITE 3
SITE 4
SITE 5
SITE 6
Labeled Biometric Data Collection (SITE 1)
Labeled Biometric Data Collection (SITE 2)
Labeled Biometric Data Collection (SITE 3)
Labeled Biometric Data Collection (SITE 4)
Labeled Biometric Data Collection (SITE 5)
Labeled Biometric Data Collection (SITE 6)
FINGERPRINT ID EXTRACTION
GENDER IDENTIFICATION
FINGER ID EXTRACTION
VGG16
VGG19
ResNet18
MobileNetv2
Proposed CNN
ResNet VGG Proposed
Weights
Federate Server
Update weights for all sites
Verification process for each site depend on Fingerprint ID Extraction, Gender identification, Finger ID extraction

Figure 1. Federated Learning Framework for Secure Biometric Authentication

This would ensure balanced and diverse training environments for the federated learning models while preserving user privacy at each site. Six neural networks-VGG16, VGG19, ResNet18, MobileNetV2, a proposed custom CNN, and a proposed ResNet-VGG hybrid model-were trained for all these deployment scenarios. Every neural network optimized using FL and capable of learning collectively from exposure to diverse data environments in the course of verification achieved high-level performance for specific tasks. This rigorous verification pipeline, enforced by the integrated influence of strong neural networks and FL, served not only to safeguard voter security and privacy from cyber threats in the conduct of these elections but also ensured that this E-election system was authentic.

The following steps are summarizing the data flow of the proposed system:

1- Collecting the biometric data (Fingerprint, gender, and finger identification for each voter at the six locations.
2- Each model at each location using biometric data without sharing any raw data.
3- The locations send only the trained weights to the central federal server.
4- The central server aggregates the weights received from all locations to create a global model.
5- The aggregated weights are sent back to the locations to improve the local models.

6- The final model is used locally to verify the voter's identity using fingerprint, gender, and finger identification.
7- This flow protects the privacy of voter data while achieving high performance and accuracy in verification.

3.1 Distributed system and task allocation

The deployed e-voting system is realized in six different distributed sites, where biometric data for each voter is collected independently from any other site. This data may include fingerprint ID, gender, and finger ID. The data is then delivered on-site to neural networks that are locally trained by their deep learning models, which have a single purpose; which is determining the proper identity of voters so they can participate in an election. At this point, each site runs independently and trains the models only on its own data to improve security while maintaining privacy for biometric information.

3.2 Implementation of Federated Learning

The proposed system used Federated Learning for collaborative model training across the six distributed sites to preserve voter data privacy and security. All the biometric data, including fingerprint ID, gender identity, and finger ID, is collected at each site separately, and all models are locally trained on these raw attributes without sharing any of this information with other sites. Instead, the central Federated Server accepts only the model weights (the learned features), as shown in Figure 2, from a user and proceeds to adjust these on its own. The received model weights from all sites are aggregated at the Federated Server in a process called Global Model Aggregation. The information learned from the varied environments and variances in data is then compiled to create a globally trained model. The averaged model weights are then returned to each of the sites, enabling them to benefit from collective learning while keeping all data secret. In this way, the e-voting system is made to be more objective and fault-tolerant without compromising on the privacy aspects since nobody can ever have access (whether at the front end or back end) to biometric information transcending its server.

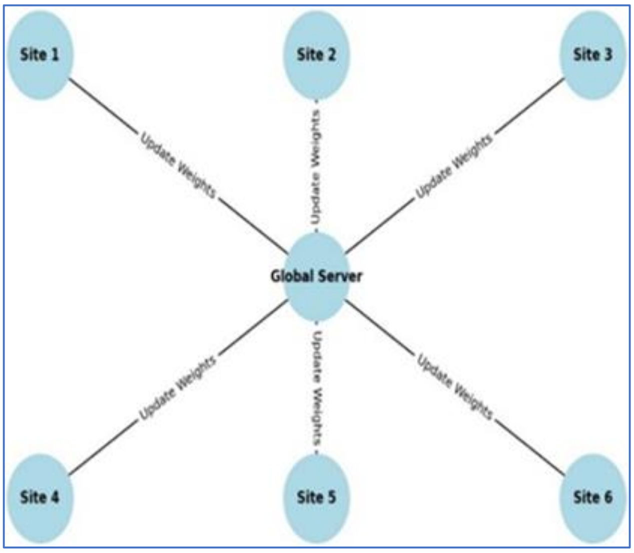

Figure 2. Weight Aggregation Process for Federated Learning

3.3 CNN Models

There are six CNN architectures employed in this system, namely: VGG16, VGG19, ResNet18, MobileNetV2, a proposed CNN, and a proposed ResNet-VGG. These CNNs efficiently process input data through various scenarios for accurate voter verification across distributed locations, using fingerprint biometric images, gender classifier image constructs, and the finger ID parameter during sign-up sessions. These models are trained using Federated Learning to ensure privacy-preserving innovations, enabling collaborative learning without sharing sensitive data.

1. Preprocessing Stage: The first-step preprocessing in the e-voting system is a normalization and uniform step intended to preprocess raw biometrics, including but not limited to fingerprint images, gender, and finger IDs, from their unprocessed form into processed data necessary for accurate model training. As a first step, fingerprint images of different sizes due to various capture devices and conditions are resized into the same dimension (e.g., 100×100 pixels). In this process, all the images will be of uniform size and can further be fitted

into neural networks. Normalizing them shifts their pixel values so that they fall between 0 and 1. This normalization step is beneficial not only to speed up convergence during model training but also enables the neural network to better differentiate fine features found in fingerprint patterns.

Algorithm 1. (Preprocessing Stage)

Input: Raw biometric data (including fingerprint images, gender, and finger ID), desired image size (img _size), and number of classes (num_classes).

Output: Preprocessed image data and corresponding categorical labels.

Begin

Step 1: For every data entry in the dataset, extract the following elements: fingerprint image, voter ID, finger number, gender, and any associated metadata (such as filenames).

Step 2: Verify the size of the fingerprint image. If it does not match the required dimensions, resize the image to ensure consistency across all data.

Step 3: Adjust the pixel values of the fingerprint image to a standardized range, typically by scaling the values to fall between 0 and 1.

End

As shown in the Algorithm 1. At the same time Algorithm 2 illustrate the processing stage.

Algorithm 2. (Processing Stage)

Input: Preprocessed biometric data (fingerprint images, gender, and finger ID), trained models (FingerprintModel, GenderModel, FingerIDModel), and stored records (fingerprint templates, gender, and finger IDs).

Output: Verification results for fingerprint, gender, and finger ID matches, and the final decision for voter authentication.

Begin

Step 1: Extract biometric elements from each preprocessed entry, including the fingerprint image, stored gender label, and stored finger ID, for further processing.

Step 2: Passed the fingerprint image through the trained models to predict the fingerprint and gender and finger ID and compared each with its stored value in the database for declaring the match-fingerprint match, gender match, finger_id_match-marking each one as True if matched or False.

Step 3: Combine the results of fingerprint, gender, and finger ID verifications. Set final decision to True if all matches are successful; otherwise, set it to False.

Step 4: If final decision is True, grant the voter permission to cast their vote. Otherwise, deny voting access and log the failure for security purposes.

End

2. Training Stage*:* At the training stage, all six distributed sites utilize Federated Learning to locally train the three models: VGG16, VGG19, and ResNet18, along with MobileNetV2, the proposed CNN, and the hybrid model (PROPOSED-RESNET-VGG) for their corresponding biometric data (fingerprint, gender, and finger ID). The local models contribute their learned weights to a central server, and the aggregated global model is shared with all sites for further refinement, ensuring secure and robust performance across all locations.

A. *Proposed CNN*

The CNN model that will be proposed is meant to take in biometric data, such as fingerprints, gender, and finger IDs. The model design includes a series of convolutional layers with batch normalization after each layer for stable and efficient training. Three batches of convolutional layers are configured, using 3×3 kernels to undertake detailed feature extractions from input images. A dropout layer is added, which will randomly neglect 25% of the units during training to prevent overfitting and improve generalization performance. Fully connected layers are used at the very end to perform feature map classification, where the feature maps are flattened and then processed by dense neural networks-fully connected ones-with a final output SoftMax layer. Besides, a dropout rate of 0.5 has been used in the fully connected layers to further improve the generalization ability of this model. This involves using the Adam optimizer with a learning rate of 0.0001, whereas categorical cross-entropy itself makes sure that this model is proficient in highly difficult tasks involving biometric classification while ensuring high accuracy. Figure 3 illustrates proposed CNN.

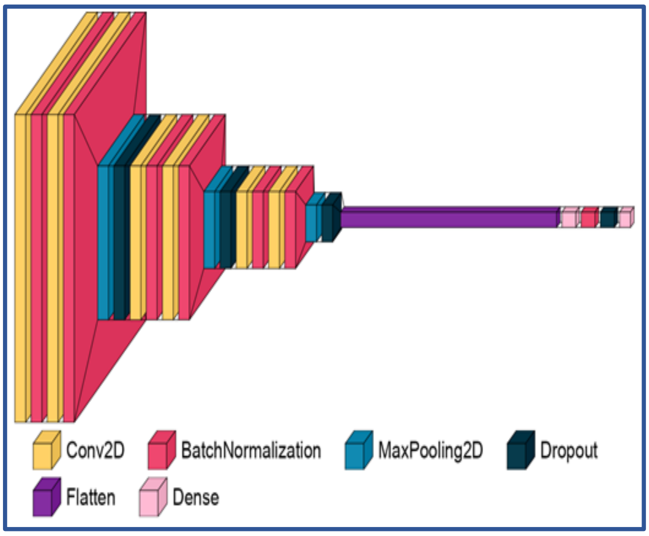

Figure 3. Proposed CNN

B. *VGG16 Model*

The VGG16 [21, 22] model is a deep convolutional neural network that is shown to be stronger than other hand-crafted, feature-based methods (as shown in Figure 4). It is normally shown to be strong, above all, in image recognition tasks. The architecture consists of five blocks of convolution, where each consists of several layers of convolution followed by ReLU activation and Batch Normalization for training stability. The number of filters in the convolutional blocks increases from 64 to 128 and 256, and for the last two blocks, 512. MaxPooling is sequentially added to these blocks subsequently to decrease the spatial size of the representations, thereby maintaining their main features.

After the convolutional blocks, there are two fully connected layers with 4096 units each for accurate interpretation of high-level features extracted by the combined workflow. To avoid overfitting in the dense layers, Dropout layers are added, where roughly half of the neurons' inputs are zeroed out as they send their output to another neuron. In the final layer, a fully connected softmax is used to determine probabilities for each class output: fingerprint ID, gender, or finger ID. The model is compiled with the Adam optimizer, with a learning rate of 0.0001, to balance the trade-off between convergence speed and accuracy. With a categorical cross-entropy loss function, it is well-suited for the multi-class classification tasks associated with biometric verification.

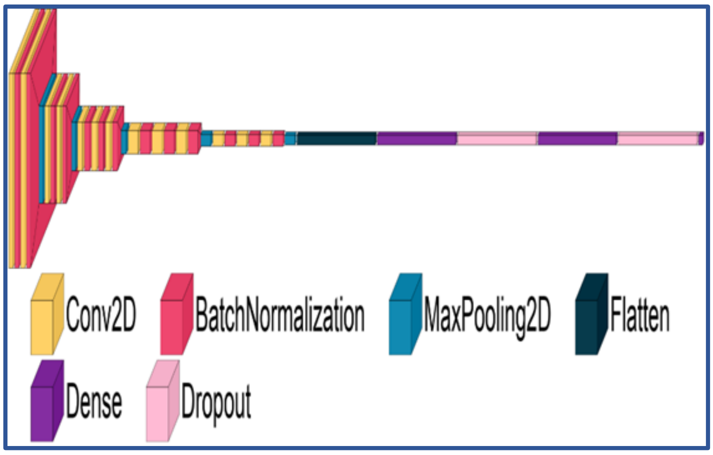

Figure 4. VGG16 Model

C. *VGG19 Model*

VGG19 [21, 23] is an enhanced version of the VGG16 architecture, meaning that it simply adds more convolutional layers to enable this network to maintain further depth for capturing more fine features from the input data. There are five convolution blocks in this model; every block contains a few convolutional layers that are then followed by batch normalization and ReLU activation. The first two blocks had 64 and 128 filters, respectively. Now, the third, fourth, and fifth blocks have 256 and 512 filters. At the end of every block, a MaxPooling layer reduces spatial dimensions to further assist the model in focusing on the most effective features.

After the convolutional blocks are two fully connected 4096-unit layers that decode the complex features extracted by the convolutional layers. Added at the end of these fully connected layers is a softmax layer, which provides class probabilities and makes the architecture fit for fingerprint classification, gender, and finger ID tasks. It is then compiled with the Adam optimizer at a learning rate of 0.0001, along with categorical cross-entropy as the loss function, to ensure effective training on multi-class classification tasks across all distributed sites. Powered by this deeper architecture, VGG19 achieves impressive precision, especially in tasks requiring detailed feature recognition. The Figure 5 illustrates VGG19 model.

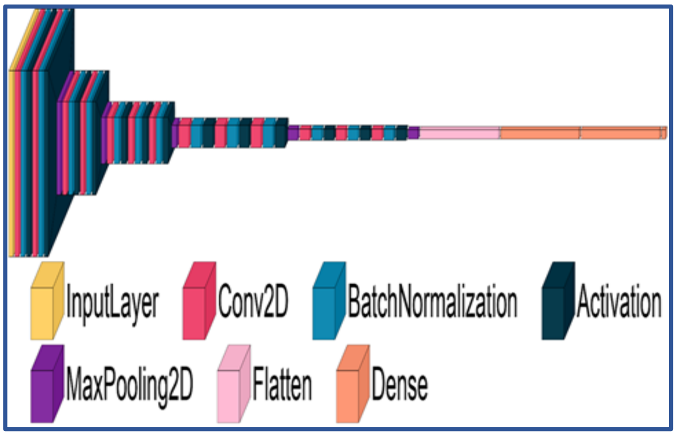

Figure 5. VGG19 Model

D. *ResNet18*

ResNet [24, 25] is a deep convolutional neural network that can train very deep networks; these were the challenges faced by earlier models like VGG16. Specifically, it starts with an initial convolutional layer with 64 filters using a 7x7 kernel and stride but is immediately followed by a MaxPooling layer to reduce the input spatially while it captures important low-level features. ResNet18 has residual blocks—one with fewer layers to ensure better information flow. Included are the identity block, which preserves spatial dimensions, and the convolutional block, which implements 1x1 convolutions in the skip connections to add the input correctly to that coming out of subsequent layers. It is structured into four sets of residual blocks: 64, 128, 256, and 512 filters, respectively (see Figure 6). All these will capture evermore complex features as the network deepens. In the first set, there is one convolutional block followed by two identity blocks. In the next three sets, it starts out with a convolutional block followed by three identity blocks. After these layers, global average pooling is applied, averaging each feature map into a single vector, effectively summarizing the learned features. Then there's a fully connected layer at the top, with the softmax function at the end assigning probabilities to classes; this makes it very suitable for classification tasks. The Adam optimizer used to finally compile it, along with categorical cross-entropy loss, secures strong performance when using it for biometric verification.

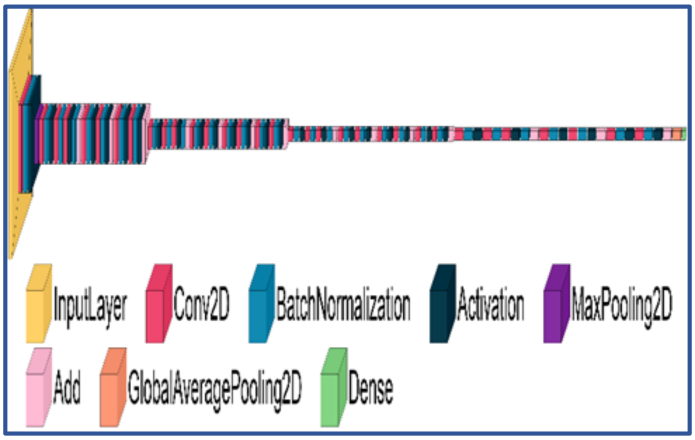

Figure 6. ResNet18 Model

E. *MobileNetv2*

MobileNetV2 [26, 27] model is lightweight and efficient, yet it remains very high performing, hence suitable for use on mobiles. In this model's architecture, depth wise separable convolutions were applied, which reduce the number of parameters and computational cost significantly compared to traditional convolutions. This could be realized through factorization of convolution into a depth wise operation followed by a 1x1 pointwise convolution, where the first layer performs filtering and the subsequent layers combine features. MobileNetV2 has also introduced inverted residual blocks with linear bottlenecks that further optimize the network by first increasing the dimensions of input features before applying depth wise separable convolution, and then projecting the output back into a lower-dimensional space. These blocks also contain shortcut connections preserving input information and facilitating the gradient flow during training.

The model starts with an initial convolutional layer, succeeded by several inverted residual blocks with depth and width, at an increasing rate, but retaining a low computational footprint. In the case of this model, it starts with 32 filters in the first convolutional layer and then scales through inverted residual blocks to 320 filters in the deeper layers. The final layers include a 1x1 convolution to increase the feature dimension to 1280, followed by a global average pooling and a fully connected SoftMax layer for classification. Because of this architecture, MobileNetV2 is very good at fingerprint, gender, and finger classification tasks while remaining computationally efficient-making it desirable for deployment on edge devices with minimal resources. It is then compiled with the Adam optimizer, along with categorical cross-entropy loss for fitting and faster training of the model related to biometric verification tasks.

F. *Proposed ResNet-VGG*

The proposed ResNet-VGG consolidated architecture merges the capabilities for effective feature extraction of both the VGG and ResNet architectures (as shown in Figure 7), hence making it more efficient for biometric classification. There are major four blocks within this model that extract features at different levels of abstractions, all of which help in performance. Blocks 1 and 2 follow the VGG style of design, where each block contains two convolutional layers with 64 filters in Block 1 and 128 filters in Block 2. The layers all have 'same' padding, using 3x3 kernels to maintain spatial dimensions. The convolutions follow Batch Normalization for stable training and are joined by ReLU activations for non-linearity. After the convolution operations are done, a MaxPooling layer is applied with a 2x2 window to drop spatial dimensions and focus on relevant high-level features.

Block 3 and Block 4 trace the evolution into ResNet, where residual blocks begin to be used. In each block, there are three residual blocks with two convolutional layers-256 filters in

Block 3 and 512 filters in Block 4. Similarly, these residual blocks add the input directly to the output by bypassing the convolutional layers due to the inclusion of shortcut connections. This design allows more information to reach deeper layers while reducing the vanishing gradient problem. Again, as has been the case with blocks described earlier, these convolutional layers are followed by Batch Normalization and a ReLU activation, which works wonderfully for effective feature learning and stable training processes. Each block is then concluded by including a MaxPooling layer to further down sample feature maps.

Finally, the model transitions from these convolutional and residual blocks into dense layers. After flattening the feature maps, the model contains two dense layers with 4,096 units each; leverage added with L2 regularization to prevent overfitting and Dropout for generalization purposes, with a dropout rate of 0.5. Finally, in the end, a SoftMax layer is used that estimate class probabilities related to tasks such as fingerprint classification, gender detection, and finger ID detection. The hybrid architecture combines the feature extraction capabilities of VGGS with ResNet's learning capabilities to give an effectual solution for complicated.

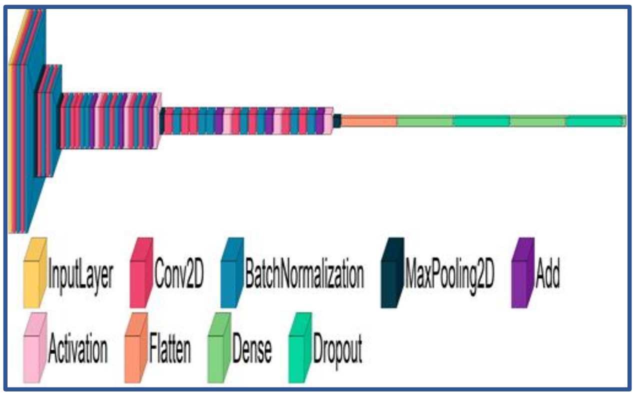

Figure 7. Proposed ResNet-VGG

3.4 Federated Learning and Model Aggregation

Federated learning is important for the e-voting system described herein because it allows collaborative training across many sites to be performed without giving away data privacy. This technique involves training a separate model at every site using its own biometric data, such as fingerprint images, gender, and finger ID, guaranteeing that no sensitive information will leave a location. A central server updates the model weights sent by each site in every epoch of training. Now, all shared models from each participating site are aggregated through an averaged weight by a central server to be the global model containing transferred knowledge from diverse datasets. The line-by-line explanation of each operation is therefore: first, flatten the weights; second, calculate the mean across sites; and third, reshape the result to original dimensions. After this, the model is aggregated and sent back to all sites for further training. This process is repeated over several epochs where the global model keeps improving as it learns from collective data.

During training, at each site, track losses and accuracies to obtain key performance metrics. These metrics are then averaged across all sites to yield the complete status of performance. Save model weights - the ones that obtain high validation accuracy-after every end of an epoch so that the best version of the model is retained. Such a federated learning scheme to let the system learn from multiparty data sources will improve generalization with strong privacy controls; hence, it can be quite suitable for distributed and sensitive biometric verification applications, including e-voting.

In Federated Learning, the update of model weights across multiple sites can be represented by the following equation:

$$w_{t+1} = \frac{1}{N}\sum_{i=1}^{n} w_t^i \tag{1}$$

Where: w_{t+1} is the updated global model weight after aggregation, N is the number of participating sites, and w_t^i Represents the model weights from site i at the current iteration t.

This equation is derived from the process where each site sends its weights to a central server, which averages them to create a generalizable model. These aggregated weights Wt+1 is used to update the global model and are then sent back to each site for the next round of local training. This iterative process continues until the model converges, usually over several epochs.

3.5 E-Voting Authentication

The authentication process in the e-voting system ensures that only authenticated voters can cast their vote (see Figure 8). Step 1 is Fingerprint Verification of the voter, where a set of deep learning models trained specifically for this purpose compares the fingerprint to a database of authorized voters. Once the fingerprint is confirmed, the system moves to the gender verification step. This step uses complementary deep learning models to check the gender associated with the fingerprint against the voter's recorded data. If gender verification is successful, the system performs a final check using Finger ID, ensuring that the correct finger is used for voting. This enhances security and reduces errors in authentication. The voter is allowed to cast a vote only when all three criteria—Fingerprint ID, gender, and Finger ID—are verified. This multi-layered authentication process mitigates unauthorized voting and maintains the integrity of the e-voting system.

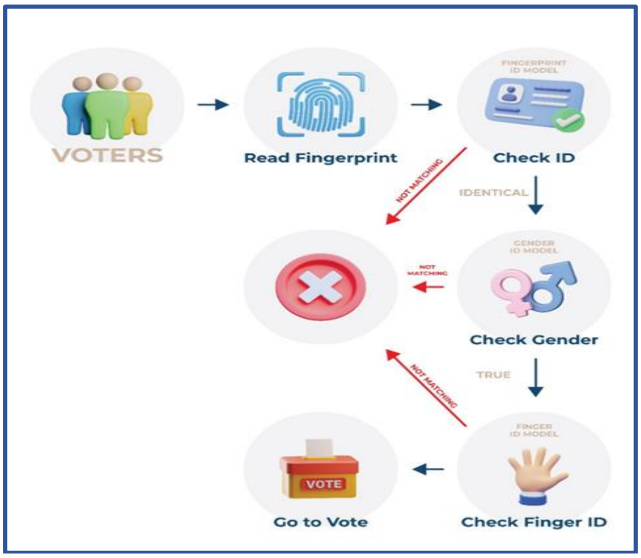

Figure 8. E-Voting Authentication

3.6 Dataset

The Sokoto Coventry Fingerprint Dataset (SOCOFing) is a biometric fingerprint database specifically created for academic research endeavors. SOCOFing comprises 6,000 fingerprint photos obtained from 600 African individuals. It includes distinctive characteristics such as gender labels, hand and finger names, and synthetically modified versions with three varying levels of alteration for obliteration, central rotation, and z-cut [28].

4. Experimental Results

4.1 Evaluation

This sub-section depicts the performance of six neural network models-VGG16, VGG19, ResNet18, MobileNetV2, Proposed CNN, and the Proposed ResNet-VGG - across three key biometric recognition tasks: ID, Finger, and Gender recognition (see Table 1). The models were evaluated on data from six geographically distinct sites. The performance metrics considered include recognition accuracy, precision, recall, and F1 score [29].

Table 1: The evaluation results of various models for biometric recognition tasks

Task	**Model**	Accuracy	Precision	Recall	F1 Score
ID Recognition	VGG16	96.75%	0.9705	0.9675	0.9668
	VGG19	94.22%	0.9481	0.9422	0.9410
	ResNet18	99.70%	0.9975	0.9970	0.9969
	MobileNetV2	99.72%	0.9975	0.9970	0.9969
	Proposed CNN	99.75%	0.9978	0.9973	0.9972
	Proposed ResNet-VGG	99.70%	0.9973	0.9970	0.9969
Finger Recognition	VGG16	99.88%	0.9989	0.9988	0.9988
	VGG19	99.85%	0.9985	0.9985	0.9985
	ResNet18	99.90%	0.9990	0.9990	0.9990
	MobileNetV2	99.90%	0.9990	0.9990	0.9990
	Proposed CNN	99.87%	0.9985	0.9985	0.9985
	Proposed ResNet-VGG	99.90%	0.9990	0.9990	0.9990
Gender Recognition	VGG16	99.85%	0.9983	0.9975	0.9979
	VGG19	99.97%	0.9995	0.9994	0.9994
	ResNet18	99.95%	0.9994	0.9993	0.9993
	MobileNetV2	99.90%	0.9993	0.9976	0.9985
	Proposed CNN	99.35%	0.9838	0.9954	0.9892
	Proposed ResNet-VGG	99.97%	0.9991	0.9998	0.9994

The high validation rates across all tested deep learning algorithms can be attributed to several factors, including the quality of the SOCOFing dataset with minimal noise and high-resolution biometric samples, as well as the choice of these architectures was guided by their diverse characteristics and widespread success in related computer vision tasks. For instance, VGG16 is recognized for its simplicity and effectiveness in image recognition, ResNet18 mitigates vanishing gradient issues with residual connections, and MobileNetV2 is optimized for efficiency in resource-constrained environments. The custom model was designed to explore task-specific optimizations.

These models provide a comprehensive evaluation of potential applicability to e-voting tasks. Improved training techniques such as data augmentation also contributed to the improved generalization ability of the models. Consequently, biometric tasks (such as identity verification, fingerprint recognition, and gender classification) are a natural fit for deep learning’s feature extraction capabilities. Table 2 summarizes the hyperparameter of biometric recognition for various models.

Table 2: Hyperparameters of the biometric recognition

Hyperparameters	**Values**
Batch size	64
Epoch	50
Learning Rate	0.0001

4.2 Comparative Analysis

According to Table 1, we conduct a comparative analysis of the performance of the six neural network models across the three biometric recognition tasks—ID, Finger, and Gender recognition. This analysis aims to identify the strengths and weaknesses of each model in terms of accuracy, precision, recall, and F1 score, providing insights into which models are best suited for specific recognition tasks within the e-voting system.

For ID Recognition, the Proposed CNN model stood out with the highest accuracy of 99.75%, closely followed by MobileNetV2 at 99.72% and ResNet18 at 99.70%. These models also demonstrated excellent precision, recall, and F1 scores, highlighting their robustness in handling ID recognition tasks. Although VGG16 and VGG19 exhibited slightly lower accuracy rates, they still maintained reasonable performance, suggesting their potential viability in certain scenarios where the highest accuracy is not the sole priority.

In the case of Finger Recognition, ResNet18, MobileNetV2, and the Proposed ResNet-VGG model showed near-perfect performance, with accuracy and other metrics consistently around 99.90%. This consistency across models indicates that Finger recognition may be a less challenging task for these neural networks, making any of these models suitable for systems where Finger recognition is a critical function. When analyzing Gender Recognition, VGG19 and the Proposed ResNet-VGG model emerged as the best performers, both achieving the highest accuracy of 99.97%, along with near-perfect precision, recall, and F1 scores. In contrast, while the Proposed CNN model performed well overall, it showed slightly lower accuracy and F1 scores in Gender recognition compared to other models. This suggests that, while the Proposed CNN model is highly effective for ID recognition, it may not be the optimal choice for Gender recognition tasks.

4.3 Impact of the Number of Sites

In this sub-section, the performance of the Proposed CNN model is compared when evaluated across 6 sites versus 12 sites. This comparison aims to understand how the model scales with an increased number of sites and whether the added complexity of a larger dataset impacts its accuracy, precision, recall, and F1 score across ID, Finger, and Gender recognition tasks.

In ID Recognition, when the model was evaluated with 6 sites, an average accuracy of 99.75% was achieved, with precision at 99.78%, recall at 99.73%, and an F1 score of 99.72%. However, when the number of sites was increased to 12, a noticeable drop in performance was observed. The average accuracy decreased to 96.07%, precision to 96.59%, recall to 96.07%, and the F1 score to 96.02%. This decline indicates that maintaining high performance in ID recognition becomes more challenging as the number of sites increases, possibly due to greater data diversity and complexity introduced by the additional sites. The model's performance in Finger recognition remained consistently high regardless of the number of sites. With 6 sites, an average accuracy of 99.87% was recorded, along with precision at 99.85%, recall at 99.85%, and an F1 score of 99.85%. When the sites were increased to 12, the performance metrics remained nearly unchanged, with an average accuracy of 99.75%, precision at 99.76%, recall at 99.75%, and an F1 score of 99.75%. This consistency suggests that Finger recognition tasks are less impacted by the increase in the number of sites, with the model continuing to perform exceptionally well even as the dataset grows more complex.

For Gender recognition, a slight decrease in performance was observed compared to Finger recognition, though the model still performed admirably. With 6 sites, an average accuracy of 99.35% was achieved, with precision at 98.38%, recall at 99.54%, and an F1 score of 98.92%. When the number of sites was increased to 12, the average accuracy was recorded at 98.87%, precision decreased to 97.56%, recall slightly increased to 99.00%, and the F1 score was 98.24%. The slight decrease in precision suggests that the model may be slightly more prone to false positives when handling a larger number of sites, though overall, strong performance was maintained.

As the number of sites in a decentralized e-voting system increases, challenges arise in terms of increased communication load and model convergence delay, increased local data heterogeneity between sites, and resource constraints at some nodes with limited capabilities. The risk of cyber-attacks such as poisoning and Byzantine attacks also increases, which requires addressing these challenges to ensure the scalability and reliability of the system in large-scale applications.

4.4 Comparison Result with State of Art

The comparison of the proposed e-voting system with various state-of-the-art methods are depicted in Table 3. Each approach targets different tasks, such as privacy protection, voter authentication, and biometric recognition, across various datasets. Notably, our system, which uses federated learning with multiple CNN models, achieves high accuracy in ID, Finger, and Gender recognition tasks, with rates of 99.75%, 99.90%, and 99.97%, respectively, when evaluated on the SOCOFing dataset across six sites. This performance is competitive with or superior to existing methods while also offering enhanced privacy protections through federated learning.

Table 3: Comparison Result with state of art

Study	Approach	Task	Dataset	Accuracy	Strengths
[14]	CNN for Fingerprint Recognition	Fingerprint Recognition	SOCOFing	~99%	High accuracy on clean datasets
[15]	zk-SNARKs, Machine Learning	Face Recognition	Custom Dataset	High (not specified)	Verifiable e-voting, privacy in vote-counting
[16]	Fingerprint Biometrics in E-Voting	Voter Authentication	Custom Dataset	High (not specified)	Secure, efficient online voting
[17]	CNN (Fig-net) for Gender Classification	Gender Classification	SOCOFing	96.47%	Efficient gender classification using CNN
[18]	CNN for Gender and Hand Identification	Gender & Hand ID	SOCOFing	99.40% (Gender)	High validation accuracy in forensic applications
[19]	Blockchain-based Federated Learning	Privacy Protection	Healthcare, Industry 5.0	Not Specified	Strong privacy guarantees, decentralized learning
[20]	IoT-based EVM with Fingerprint Biometrics	Real-time Voting	Custom Dataset	High (not specified)	Real-time data communication, secure storage
Proposed System	Federated Learning with Multiple CNN Models	ID, Finger, Gender	SOCOFing (across 6 sites)	99.75% (ID), 99.90% (Finger), 99.97% (Gender)	High accuracy across tasks, enhanced privacy with FL

5. Conclusions

This study proposed the design and performance of a secure, privacy-preserving e-voting system using federated learning across multiple sites, testing six neural network models—VGG16, VGG19, ResNet18, MobileNetV2, Proposed CNN, and Proposed ResNet-VGG—on key biometric recognition tasks: ID, Finger, and Gender recognition. The Proposed CNN model excelled in ID recognition with the highest accuracy, precision, recall, and F1 scores across six sites, though its performance slightly declined when the number of sites increased to twelve, with accuracy dropping to 96.07%, indicating that further optimization may be needed to manage the increased complexity and data diversity. Finger recognition tasks showed remarkable stability, maintaining high accuracy of around 99.75% across all sites, suggesting that this task is less sensitive to increased data complexity. For

Gender recognition, the model experienced a slight decrease in performance with more sites, reflecting a minor rise in false positives, yet it still maintained strong overall performance. These findings highlight the effectiveness of federated learning in achieving secure, distributed biometric verification, but they also underscore the need for task-specific model selection and continuous refinement to ensure consistent performance as the system scales.

Future work could explore scaling the system to more sites and integrating additional biometric modalities to further enhance security and accuracy. Additionally, optimizing the federated learning process for real-time applications in large-scale elections will be a key focus, as will testing the model using more comprehensive age, gender, and facial recognition datasets to ensure fairness and effectiveness.

References

[1] S. S. Chaeikar, A. Jolfaei, N. Mohammad, and P. Ostovari, “Security principles and challenges in electronic voting,” in 2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW), 2021, https://doi.org/10.1109/EDOCW52865.2021.00030.

[2] A. D. Rubin, “Security considerations for remote electronic voting,” Commun. ACM, vol. 45, no. 12, pp. 39–44, 2002, https://doi.org/10.1145/585597.585599.

[3] L. Carter and F. Bélanger, “The utilization of e-government services: citizen trust, innovation and acceptance factors,” Inf. Syst. J., vol. 15, no. 1, pp. 5–25, 2005, https:// doi.org/10.1111/j.1365-2575.2005.00183.

[4] W. Ali Mahmood, J. Waleed, A. R. Abbas, H. Alaskar, M. Altulyan and A. Jaafar Hussain, "Intelligent Gesture-Enhanced Blockchain Voting: A New Era of Secure and Accessible E-Voting," in IEEE Access, vol. 12, pp. 144055-144068, 2024. doi: 10.1109/ACCESS.2024.3468338.

[5] Shamim and I. Quazi Jahan, “E-Voting Security Protocol: Analysis & Solution,” International Journal of Engineering Research and Applications, vol. 2, pp. 2938–2943, 2012.

[6] M. Sepehri, S. Cimato, and E. Damiani, “Privacy-preserving query processing by multi-party computation,” Comput. J., vol. 58, no. 10, pp. 2195–2212, 2015, https:// doi.org/10.1093/comjnl/bxu093.

[7] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, “Communication-efficient learning of deep networks from decentralized data,” arXiv [cs.LG], 2016, https://doi.org/10.48550/arXiv.1602.05629.

[8] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated Machine Learning: Concept and Applications,” arXiv [cs.AI], 2019, https://doi.org/10.48550/arXiv.1902.04885.

[9] K. Ishida, A. Ercan, T. Nagasato, M. Kiyama, M. Amagasaki, "Use of one-dimensional CNN for input data size reduction in LSTM for improved computational efficiency and accuracy in hourly rainfall-runoff modeling", Journal of Environmental Management, vol. 359, 2024. https://doi.org/10.1016/j.jenvman.2024.120931.

[10] B. S. Mahdi, M. J. Hadi, and A. R. Abbas, “Intelligent security model for password generation and estimation using hand gesture features,” Big Data Cogn. Comput., vol. 6, no. 4, p. 116, 2022, https://doi.org/10.3390/bdcc6040116.

[11] V. Safavi, A. M. Vaniar, N. Bazmohammadi, J. C. Vasquez, O. Keysan, J. M. Guerrero, "Early prediction of battery remaining useful life using CNN-XGBoost model and Coati optimization algorithm", Journal of Energy Storage, vol. 98, 2024. https://doi.org/10.1016/j.est.2024.113176.

[12] J. Waleed, et al. "An Effective Deep Learning Model to Discriminate Coronavirus Disease From Typical Pneumonia", International Journal of Service Science, Management, Engineering, and Technology (IJSSMET), vol.13, no.1, pp.1-16, 2022. http://doi.org/10.4018/IJSSMET.313175.

[13] M. J. Sheller et al., “Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data,” Sci. Rep., vol. 10, no. 1, p. 12598, 2020, https://doi.org/10.1038/s41598-020-69250-1.

[14] M. H. Or Rashid, S. Rahman, J. Satu, and A. S. Tariq, “Convolutional neural network approach for precise fingerprint recognition,” in 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering, 2021, https://doi.org/10.1109/IC4ME253898.2021.9768468.

[15] J. Liu, T. Han, M. Tan, B. Tang, W. Hu, and Y. Yu, “A publicly verifiable E-voting system based on biometrics,” Cryptography, vol. 7, no. 4, p. 62, 2023, https://doi.org/10.3390/cryptography7040062.

[16] H. Hamran, M. Abdullah, M. E. Naveed, and A. R. Afzal, “Design and implementation of secure electronic voting system using fingerprint biometrics,” Journal of Artificial Intelligence and Computing, vol. 1, no. 1, pp. 1–5, 2023, doi: 10.57041/jaic.v1i1.887.

[17] A. Narayanan and Q. M. Hameed Uddin, “Gender detection and classification from fingerprints using convolutional neural network,” in 2023 4th International Conference on Signal Processing and Communication (ICSPC), 2023, https://doi.org/ 10.1109/ICSPC57692.2023.10125703.

[18] D. Maiti and D. Das, “Gender and hand identification based on dactyloscopy using deep convolutional neural network,” in Lecture Notes in Networks and Systems, Singapore: Springer Nature

Singapore, 2023, pp. 145–155, https://doi.org/10.1007/978-981-99-3734-9_13.

[19] Sameera et al., “Privacy-preserving in Blockchain-based Federated Learning systems,” Comput. Commun., vol. 222, pp. 38–67,2024, https://doi.org/10.1016/j.comcom.2024.04.024.

[20] C. H. Srilatha et al., “Fingerprint-based biometric smart electronic voting machine using IoT and advanced interdisciplinary approaches,” E3S Web Conference, vol. 507, p. 01037, 2024, https://doi.org/10.1051/e3sconf/202450701037

[21] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,”2014, https://doi.org/10.48550/arXiv.1409.1556.

[22] O. İ. Çelik, C. Gazioğlu, "Leveraging deep learning for coastal monitoring: A VGG16-based approach to spectral and textural classification of coastal areas with sentinel-2A data", Applied Ocean Research, vol. 151, 2024. https://doi.org/10.1016/j.apor.2024.104163.

[23] S. Mascarenhas and M. Agarwal, “A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification,” in 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON),2021,http://doi.org/10.1109/CENTCON52345.2021.968794.

[24] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015, https://doi.org/10.48550/arXiv.1512.03385.

[25] A. Ullah, H. Elahi, Z. Sun, A. Khatoon, and I. Ahmad, “Comparative analysis of AlexNet, ResNet18 and SqueezeNet with diverse modification and arduous implementation,” Arab,” J. Sci. Eng, vol. 47, no. 2, pp. 2397–2417, 2022 https://doi.org/10.1007/s13369-021-06182-6.

[26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, https://doi.org/10.1109/CVPR.2018.00474.

[27] R. J. Kolaib and J. Waleed, "Crime Activity Detection in Surveillance Videos Based on Developed Deep Learning Approach," Diyala Journal of Engineering Sciences, vol. 17, no. 3, pp. 98-114, 2024. doi: 10.24237/djes.2024.17307.

[28] Y. I. Shehu, A. Ruiz-Garcia, V. Palade, and A. James, “Sokoto Coventry Fingerprint Dataset,” 2018,https://doi.org/10.48550/arXiv.1807.10609.

[29] A. Al-Saegh, A. Daood, and M. H. Ismail, “Dual Optimization of Deep CNN for Motor Imagery EEG Tasks Classification”, Diyala Journal of Engineering Sciences, vol. 17, no. 4, pp. 75-91, 2024. doi: 10.24237/djes.2024.17405.