International Journal on Advanced Science Engineering Information Technology

Vol.15 (2025) No. 1
ISSN: 2088-5334

Implementing Computer Vision and Biometrics into User Authentication

Danial Muhammad Firdaus Anson [a], Nur Erlida Ruslan [a,*], Su-Cheng Haw [a]

[a] *Faculty of Computing and Informatics, Multimedia University, Persiaran Multimedia, Cyberjaya, Malaysia*
Corresponding author: [*]nurerlida@mmu.edu.my

***Abstract*—The highly digital world that is present today requires robust protection of personal information, primarily when it is handled digitally. As cyber threats continually expand, the demand for reliable authentication systems becomes more critical. Many security systems are more secure than currently implemented since some still use basic and multi-factor authentication features. Unfortunately, these systems are tedious to navigate and would compromise user experience so that it can be more secure. This paper goes through research in the field, focusing on facial and speech recognition. Additionally, implementing current authentication systems will be evaluated and used further as benchmarking. The review intends to gather an understanding of the current state of research and real-world implementation so that a method of implementing computer vision in biometric authentication can be proposed. This paper comprehensively overviews the current state-of-the-art facial and brief speech recognition state. Training a model and evaluating its accuracy can be viable for biometric authentication. This paper first demonstrates facial recognition as Labelled Faces in the Wild (LFW) taken from Kaggle.com. The proposed result was focused on the accuracy metric. This paper shall be continued by using libraries such as Keras Tuner and Optuna to assist in selecting the optimal set of hyperparameters.**

***Keywords*—Computer vision; image processing; biometrics; user authentication; facial recognition.**

Manuscript received 11 Sep. 2024; revised 24 Nov. 2024; accepted 4 Jan. 2025. Date of publication 28 Feb. 2025.

I. INTRODUCTION

Nowadays, information is key, and securing personal information is crucial. It has become a challenge to secure information as malicious parties try to steal information to exploit it to gain a benefit. The same can be said for digital information since, nowadays, cybersecurity is as critical as traditional security because information is mainly stored in servers and digital databases. Cybersecurity and Infrastructure Security Agency (CISA) defines cybersecurity as the art of protecting networks, devices, and data from unauthorized access or criminal use and the practice of ensuring confidentiality, integrity, and availability of information. Cybersecurity is critical because proper permissions should be set so that sensitive information does not fall into the wrong hands. The Internet Crime Complaint Center (IC3) received over 80,000 complaints of data breaches and identity theft in 2022. As such, authentication is vital in ensuring that information theft does not occur. Traditionally, authentication is handled by identification cards usually issued by an authority (e.g., Identification Card (IC), driving license, passport, etc.).

In the digital world, user authentication is done through the use of usernames and passwords, and most online services require the creation of an account with a username and password. According to a report published by Okta Inc., passwords are still the most commonplace authentication method worldwide. This becomes more important as more businesses start to provide digital services.

On the other hand, as computing advances, there is an interest in enabling computers to gain an understanding of the world around us. The field of computer vision aims to do this by implementing algorithms and techniques so that a computer can replicate human sight by processing information, namely images and videos. The implementation of computer vision technology can be seen in biometric identification. For example, facial recognition and fingerprint recognition use computer vision technology so that a computer could identify if the biometric information given matches the information in its database [1], [2]. The implementation of biometric identification exceeds that of human capabilities because it is far more accurate and quicker. In recent years, research such as [3], [4], [5], [6], [7] has been done to improve biometric authentication using

newer technologies such as blockchain technology, artificial intelligence (AI), and deep learning.

For a long time, the username and password combination proved to be a sufficient authentication method in combating identity theft. Unfortunately, [8] shows that malicious actors have evolved, and a simple username and password system is not secure enough for the modern age, making its use a bad practice. As such, most services would implement a multi-factor authentication (MFA) system such as the Transaction Authorization Code (TAC) shown in Fig. 1.

Fig. 1 TAC sent through Short Message Service (SMS)

The study in [9] found that employing MFAs significantly reduces the risk of an account being compromised. However, some services prefer a different approach to heighten security by implementing biometric authentication. Biometrics serves as a good way to increase security without compromising on user-friendliness. According to [10], [11], fingerprint and facial recognition are used in most smartphones' authentication systems as they are secure enough for use. Here, computer vision technology is being implemented in user authentication through biometrics.

This paper aims to review recent research papers advancing in the field. Additionally, current authentication systems are studied to see how user authentication is handled. In doing so, a good understanding of computer vision and biometric technology in user authentication can be gathered. Biometrics can be easily associated with a user's identity, which it pertains to how it is unique to each person [12]. Furthermore, this study demonstrates the plausibility of using facial recognition in biometric authentication.

In recent years, computer vision and biometric technology have been a focus for researchers because of their potential ability to be applied to our daily lives. A ton of research, [11], [13], [14], [15], [16] has been conducted on the different implementations of computer vision and biometric technology to improve their performance. This review analyzes the implementations of facial and speech recognition.

II. Materials and Method

A. Background

A simple test on implementing facial recognition is subject to show the technology.

1) Gabor Wavelets: Gabor wavelets are a feature-based method that uses mathematical functions to do a feature extraction method before being input to the Back Propagation Neural Network (BPN). Introduced by Dennis Gabor in the 1940s to analyze the frequency content of signals, the method is very popular in computer vision. For facial recognition, Gabor wavelets help extract discriminative features from images, enabling the handling of lighting variations and facial expressions. The Principal Component Analysis (PCA) is also used to extract facial features before dimensionality reduction. The Gabor-PCA method resulted in better performance accuracy. Besides that, another hybrid of Gabor and PCA is proposed whereby the Support Vector Machine (SVM) is used as the classifier whereby the performance of the technique is tested with FRGCv2 and ORL face database Unfortunately, generally, Gabor wavelets have a few downsides, namely being computationally complex, having high parameter sensitivity, limited robustness to noisy images, and having difficulty in handling scale variations of facial features. For these reasons, current research on Gabor wavelets also includes the usage of machine-learning methods, as seen in [17] and [18]. The rise in popularity of machine learning has brought about a focus on implementing machine learning methods in different research domains, including face recognition. Two popular deep learning methods later used for facial recognition are Convolutional Neural Networks (CNN) and Siamese Neural Networks.

2) Convolutional Neural Network (CNN): CNNs are a popular type of deep learning model designed for processing and analyzing visual data or extracting landmark features from the given input image [19]. CNNs are well suited for tasks such as image processing because of their ability to automatically learn hierarchical features, eliminating the need to define local features. CNNs are also favored because of their scalability, having better performance with a big dataset with varied images such as the one used in [11]. Besides that, these landmarks are passed to the K-Nearest Neighbor (KNN) algorithm, a popular machine learning technique because of its simple implementation and accurate results, which identify a person by comparing the facial key points extracted from the person's input image with the facial keypoint values that are captured and stored in the database during the registration of that particular person. The accuracy of CNNs is also impacted by the loss function chosen. Loss functions are used to find the difference between the predicted output and the ground truth as a measure of model evaluation, giving a lower value when the performance is high. One example of a high-performance loss function is SoftMax-Loss, as shown in equation 1, commonly used in the final layer of a neural network for multi-class classification tasks. Although researchers in [20] and [21] have supplemented it to gain more discriminative features.

$$\mathcal{E}(y,z) = -\log\left(\frac{e^{z}y}{\sum_{j=1}^{m} e^{z}j}\right) = log\left(\sum_{j=1}^{m} e^{z}j\right) - z_y \quad (1)$$

CNNs for facial recognition provide an implementation that is high in accuracy without being computationally complex and having a shorter execution time, leading to the possibility of real-time applications of facial recognition, as shown in [20] and [22]. Researchers in [20] have studied some popular examples of face recognition library packages using CNNs such as VGG-Face, OpenFace, DeepFace, DeepID, and Dlib. Besides, FaceNet also employs a deep convolutional network that is optimized directly for generating embeddings, avoiding the intermediate bottleneck layers used in earlier deep learning methods [23].

3) Siamese Neural Network: Conversely, there is a Siamese Neural Network, specifically designed to compare inputs in a pair. They are different compared to CNNs in that

they do not classify images into labels, instead comparing the distances i.e., the similarities between the two input images, explained in [21]Siamese Neural Networks are made up of a pair of identical neural networks, usually CNNs, sometimes called sister networks. The network uses the Inception Resnet version 1 with pre-trained weights for encoding the image and a Multi-Task Cascaded Convolutional Network (MTCNN) for face detection. Fig. 2 shows these sister networks that process the input samples independently, and the outputs are then compared to measure their similarity.

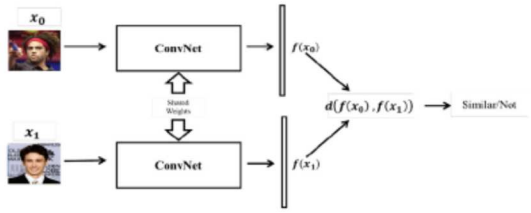

Fig. 2 Siamese Neural Network consisting of two CNNs

Siamese neural networks are used in face recognition because it excels in one-shot learning, a technique that lets models learn using a single sample, as researched in [21]. Because of its structure of having a pair of neural networks, the performance of Siamese neural networks is also impacted by the chosen loss function. Research in [24] This shows that the advantage of using a Siamese neural network over CNNs is that it can perform as well as CNNs with a smaller dataset. However, it also has higher computational complexity and requires careful tuning of hyperparameters.

Next, we also provide reviews on speech recognition and authentication systems.

1) Mel-frequency cepstral coefficients (MFCC): MFCC is a feature extraction technique common in audio processing. MFCCs capture the characteristics of speech signals and have been proven in [25] to be effective in machine learning. Although MFCCs can be used for speech recognition by pattern matching, modern researchers combine MFCCs with machine learning methods to enhance their performance.

2) Convolutional Neural Network (CNN): Just like facial recognition, CNNs referred to [26], [27] are popular for speech recognition and are employed in acoustic modeling to capture hierarchical patterns in spectrogram representations of sound signals such as the one in Fig. 3.

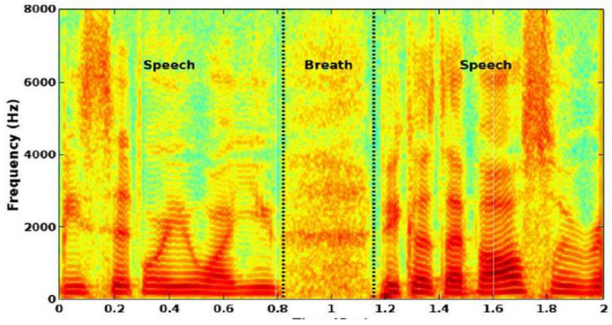

Fig. 3 Spectrogram of a sound signal in [17]

CNNs excel at finding patterns and structures in image-like spatial data. CNNs have translation invariance, giving them the ability to recognize patterns even though their exact position in the input is not known. In speech recognition, translation invariance leads to the model being good at handling variations in temporal misalignments due to a variability in speaking speed, environmental noise, and intonation. CNNs have also been shown by researchers in [28] to be good at handling speech with varying pronunciations and accents but is still behind in performance compared to Recurrent Neural Networks (RNNs).

3) Recurrent Neural Network (RNN): RNNs, especially Long Short-Term Memory (LSTM), are widely used in speech recognition because they can capture long-range dependencies and work out the vanishing gradient problem associated with traditional RNNs. LSTMs perform well in modeling sequential data, making them good at tasks where understanding temporal dependencies, such as language processing, is essential. LSTMs are equipped with memory cells that can store and retrieve information over long sequences, which is critical for determining dependencies in speech signals across multiple time steps. Like CNNs, LSTMs effectively capture temporal misalignments in speech signals.

CNNs and RNNs are like their implementation in speech recognition in that their performance relies on the loss function that is being used, and they also require hyperparameter tuning. Researchers in [28] and [26] combining CNN and RNN in a speech recognition model can yield very high performance when done correctly. Online services in the modern day are used by millions, if not billions, of people worldwide. As such, companies prioritize ensuring their services have enough security, which is where user authentication takes place. Most online services have an account system with user authentication to protect sensitive information. After reviewing a few commonly used services, data has been collected on the authentication systems used and tabulated in Table 1.

TABLE I
AUTHENTICATION SYSTEMS OF SERVICES REVIEWED

Service	Basic Authentication Features	Multi-Factor Authentication (MIFA)	Additional Notes
Maybank	Username and Password, PIN	Fingerprint authentication, Secure2u	Fingerprint authentication on the mobile app
Netflix	Email and Password	-	Email sent when account accesses in different location
WhatsApp	Phone Number	6-digit one time password (OTP)	MFA is required for login
X (formally known as Twitter)	Username and Password	SMS/Mobile app/Security key authentication	SMS authentication is locked behind a subscription
Shopee	Mobile Number/Username /Email and Password	PIN/Fingerprint authentication	PIN authentication and fingerprint authentication for making payments
Steam	Phone Number	PIN/Fingerprint authentication/Face recognition	Face verification to access additional benefits

As shown, Maybank [29] and WhatsApp [30] both successfully implement Multi-Factor Authentication (MFA). The former does it through Secure2u, as shown in Fig. 4, which uses a one-tap approval on its mobile application and MAE or 6-digit Transaction Authorization Code (TAC) sent to the user's mobile number through SMS. Maybank does this

because of the nature of the business, which provides financial services and requires adequate security measures. The latter utilizes a 6-digit code sent to the user's connected phone number through SMS. Secure2u is a safer and more convenient way to authorize Maybank2u web and MAE app transactions using Secure Verification and Secure TAC. Fig. 5 shows the 6-digit code required for a user to log into their account. As an instant messaging service used by millions of users worldwide, WhatsApp needs to take security seriously, as messages sent might contain sensitive information or documents.

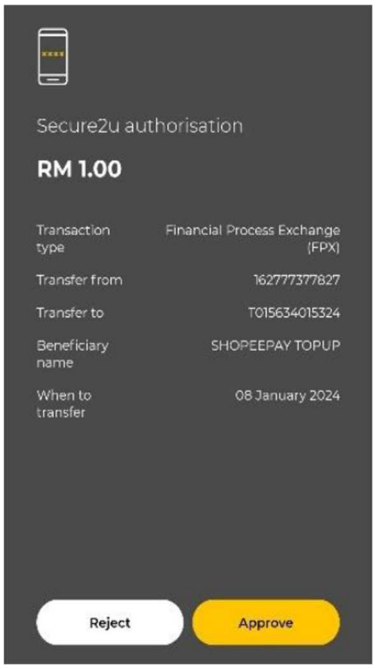

Fig. 4 Secure2u on MAE

Fig. 5 WhatsApp 6-digit code

Netflix, on the other hand, is lacking in security as it does not implement any form of MFA [31]. The only extra security feature is that if an account is accessed in a new location, an email containing the location and time that the account was accessed will be sent to notify the user. However, there is no way of revoking access to the account without changing the password. Because of its weak authentication system, numerous cases have been reported where a user's Netflix account has been accessed by a third party from a different location.

As shown in Fig. 6, X provides their users with three methods of enabling MFA: SMS authentication, mobile application authentication, and security key. Although SMS authentication is locked behind a subscription, mobile application authentication, and security keys can be set up quickly. Mobile application authentication requires a supported authenticator app such as Authy and Google Authenticator, while a security key requires a supported device/web browser. As businesses can use X, account security is handled properly.

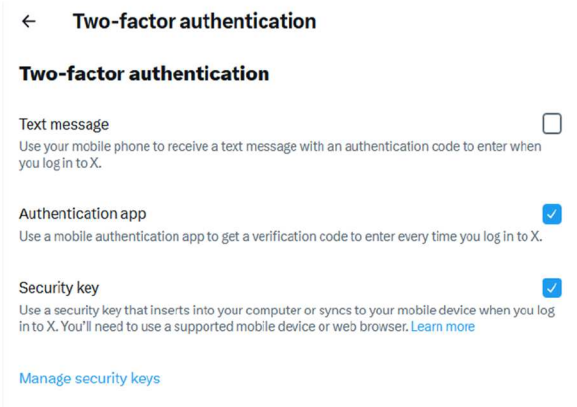

Fig. 6 MFA methods on X

While the services they provide are different, Shopee is an e-commerce website, and Steam is an online game store. They provide users with a wallet system to store money to buy products. Both services also allow users to add their payment cards to their accounts, and Shopee also keeps your home address for shipping purposes. As a result, both Shopee and Steam employ MFA in some way, Shopee uses a PIN/fingerprint authentication for payments, as shown in Fig. 7. Fig. 8 shows Steam's mobile application, which uses its Steam Guard system that allows users to authenticate logins using a code or logging in by scanning a quick-response (QR) code.

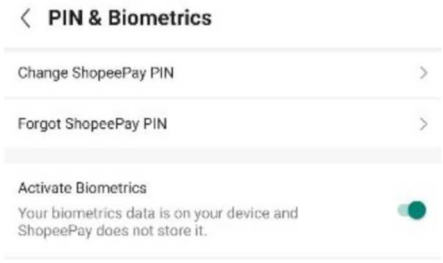

Fig. 7 Shopee PIN/fingerprint authentication

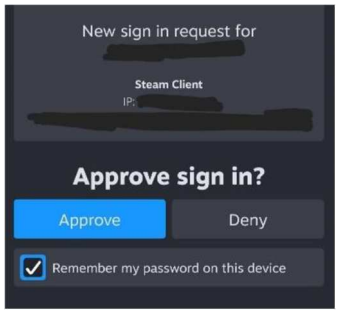

Fig. 8 Steam mobile authentication

Touch ‘n Go eWallet implements MFA during signup by asking users to add a phone number and enter the TAC sent. Additionally, users will need to create a PIN to authenticate payments. Users can verify their account to access additional benefits using their government-issued IC. Facial verification is used to validate the information entered.

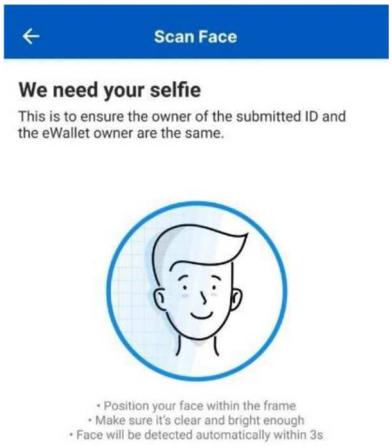

Fig. 9 Touch ‘n Go eWallet face verification

Alternative methods provided to authenticate payments are device biometric, which uses the fingerprint data or face data stored by the users’ device, and face verification, which uses the face data stored by Touch ‘n Go [32].

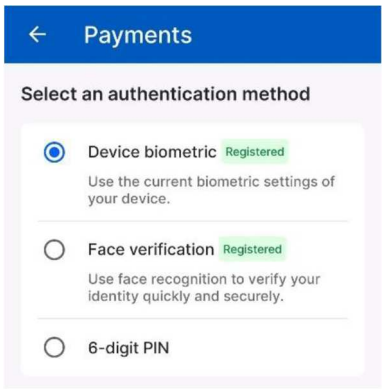

Fig. 10 Touch ‘n Go eWallet payment authentication methods

The review has shown that most companies are taking authentication seriously. Apart from Netflix, the services reviewed implement MFA in one way or another, with Maybank and Shopee opting to use biometrics as fingerprint authentication. Additionally, Touch ‘n Go eWallet users can authenticate payments using face recognition. Most of the services do not use facial or speech recognition as an authentication method, most probably because of fingerprint authentication’s familiarity and accuracy.

B. Proposed Method

The dataset that will be used to prove facial recognition is Labelled Faces in the Wild (LFW), which was taken from Kaggle.com. LFW is made and maintained by researchers at the University of Massachusetts and contains 13,233 face images of 5,749 people. It is widely used to assess neural network model performance because of its large size and diverse facial variations, including different poses, lighting conditions, and expressions. LFW images are pre-processed to be resized and center-aligned using the Viola-Jones algorithm so that features such as nose and eyes are positioned relatively consistently across all photos. Doing this has yielded better results when training a facial recognition model. LFW also contains metadata tags for each image and configuration file that will help split the dataset into training and testing sets. Some examples of images in the LFW dataset are shown in Fig. 11.

Fig. 11 Sample images from LFW dataset showcases variations of race, angle, lighting, and visibility

From the LFW dataset, we used facial data images from 19 people aged 30 to 70, consisting of at least 40 images of each person, to train the model to predict that specific person's identity. The images in LFW have already been pre-processed using the deep funnel method, which rotates and resizes the original images to have consistent eye and nose positions, and the face is centered in the image, making the dataset better for model training. The format of the images is in jpg, and the image number is padded to four characters with leading zeroes to ease the navigation of the images in a better way. Additionally, all images in the dataset have been resized and reformatted into a 250x250 pixel detected and centered using the OpenCV implementation of the Viola-Jones face detector for uniformity and configuration files to split the dataset into training and testing sets. After normalizing the pixel values, the dataset can be used for model training.

Fig. 12 shows that further preprocessing is done, such as resizing the images to 75 percent of their original size and converting them to grayscale to decrease the computational power needed. The edges of the images are also sliced so that the model is not confused by background information. The dataset is split into a training and testing set, with 80 percent for the former and 20 percent for the latter.

Fig. 12 Example of an image before and after further preprocessing

Next, methods such as CNN, RNN, and Siamese Neural Networks have been used to tackle the problem of facial recognition. For this demonstration, CNN will be implemented. CNN was chosen because it is relatively fast while maintaining high accuracy. Fig. 13 shows the structure

of the CNN used. It is a basic CNN consisting of only one convolution layer and one pooling layer.

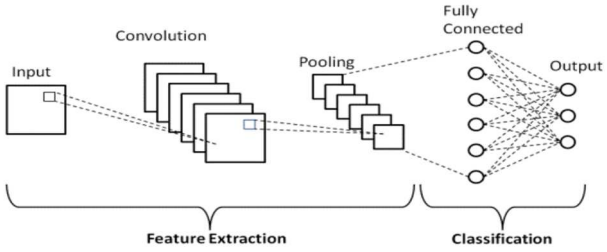

Fig. 13 Structure of CNN used

The neuron in a neural network's output mainly depends on its activation function. An activation function is a mathematical function that is applied to the weighted sum of a neuron's inputs. The convolution layer will utilize Rectified Linear Unit (ReLU) for its activation function. ReLU helps the vanishing gradient problem and reduces training time. Fig. 14 shows that in ReLU, if the input is positive, it outputs the value as is and outputs 0 if it is negative.

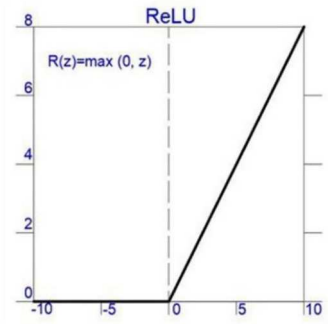

Fig. 14 ReLU activation function

The output layer, on the other hand, can be used to utilize SoftMax as its activation function. SoftMax is commonly used for multi-class classification problems where images are classified into one of many possible classes, in this case, the different people that the model is trained on. The loss function in CNN computes the difference between the expected and actual output, and the goal of training a CNN is to minimize the difference. CNN can use one of several loss functions, depending on the problem it must solve. When compiling the model, sparse categorical cross-entropy is used as a loss function. Because the model applies SoftMax to the output layer, SoftMax-loss is utilized.

For training, a batch size of 32 is used. This means the model will process 32 image samples during training, updating the weights after each batch of 32 samples has been processed. This is done to help manage memory usage and help improve the model's performance in generalizing to unseen data. Additionally, a learning rate of 0.001 is used. Learning rate is a hyperparameter that determines how much to tweak the model in response to the estimated error each time the model weights are updated.

III. Results and Discussion

A. Model Testing and Evaluation

After training and testing the model, evaluation can be done by making a confusion matrix of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). The confusion matrix can derive several significant evaluation metrics, including accuracy, precision, recall, and F-score. Equations 2 to 5 show the formulas used to determine these metrics.

$$Precision = \frac{TP}{TP0+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F-Score} = \frac{2xPrecision\ x\ Recall}{Precision+Recall} \quad (4)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FN} \quad (5)$$

Fig. 15 shows the confusion matrix and evaluation metrics of the model. The model performs well with evaluation values of more than 80 percent.

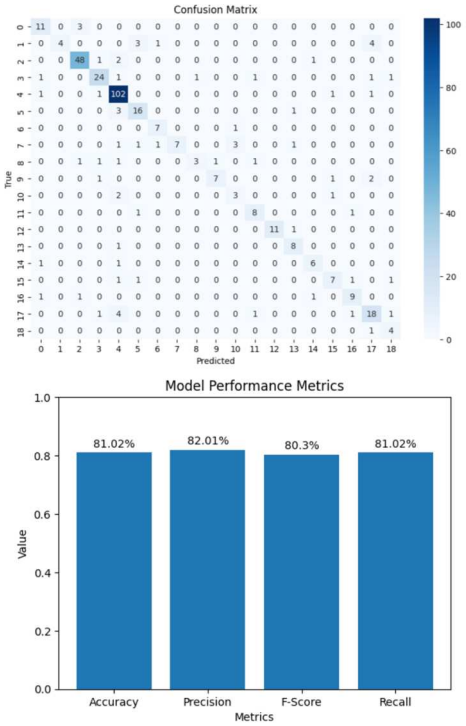

Fig. 15 Confusion matrix and evaluation metrics of model

However, to determine the accuracy metric, it was also necessary to evaluate the specificity and sensitivity. As the computation of the accuracy metric also involves calculating the specificity and sensitivity, we focus on the accuracy metric in this research.

B. Model Refinement

If the initial algorithm modelled is not ideal, it could be because of data overfitting or underfitting. In both cases, the model generalizes badly; in other words, it does not adapt well to unseen data. Overfitting means that the model learns too well and picks up the noise in the training data, and underfitting happens when the model is too simple to learn the complexity in the training data. In both cases, the model needs to be refined.

Refinement can be made to the models by employing hyperparameter tuning, which involves tuning the hyperparameters to improve performance. Hyperparameters manually set, such as learning rate and batch size, can be tuned by giving different values for the tuner to test out.

Tuning should be done to find the best accuracy, and sampling algorithms such as random search, grid search, or Bayesian optimization can be utilized to search for the best possible hyperparameter values. Libraries such as Keras Tuner and Optuna can assist in selecting the optimal set of hyperparameters.

C. Benchmarking

To prove the results, the accuracy performance will be compared with other methods. All the datasets mentioned in A and the methods involved in B will be used and implemented.

IV. Conclusion

The main goal of this paper is to determine how computer vision technology may be integrated into biometric authentication. In the first part of the review, current research in the field was studied to determine the direction in which the efforts are being made. Doing so concluded that machine learning techniques are the most common method of implementing computer vision into biometric authentication. Then, authentication systems currently in use were examined to see real-world implementations. This sheds some light on how fingerprinting is often used as a form of biometric authentication. Meanwhile, the implementation of facial and speech recognition has room to improve.

Additionally, a demonstration was made to show how facial recognition can be applicably used for biometric authentication. The CNN model trained showed a promising accuracy of over 80 percent. For future work, speech recognition should be tested to see its viability. In addition, hyperparameter tuning of both facial and speech recognition models can improve accuracy even further. Libraries such as Keras Tuner and Optuna can assist in selecting the optimal set of hyperparameters.

References

[1] V. W. S. Tan, W. X. Ooi, Y. F. Chan, C. Tee, and M. K. O. Goh, "Vision-based gait analysis for neurodegenerative disorders detection," *J. Inform. Web Eng.*, vol. 3, no. 1, pp. 136–154, 2024, doi: 10.33093/jiwe.2024.3.1.9.

[2] J. Kim, T.-S. Ng, and A. B. J. Teoh, "Conditional deployable biometrics: Matching periocular and face in various settings," *J. Inform. Web Eng.*, vol. 3, no. 3, pp. 302–313, Oct. 2024, doi: 10.33093/jiwe.2024.3.3.19.

[3] F. Ahamed, F. Farid, B. Suleiman, Z. Jan, L. A. Wahsheh, and S. Shahrestani, "An intelligent multimodal biometric authentication model for personalised healthcare services," *Future Internet*, vol. 14, no. 8, pp. 1–28, 2022, doi: 10.3390/fi14080222.

[4] A. Jaya Prakash, K. K. Patro, M. Hammad, R. Tadeusiewicz, and P. Pławiak, "BAED: A secured biometric authentication system using ECG signal based on deep learning techniques," *Biocybern. Biomed. Eng.*, vol. 42, no. 4, pp. 1081–1093, 2022, doi: 10.1016/j.bbe.2022.08.004.

[5] N. D. Sarier, "Privacy preserving biometric authentication on the blockchain for smart healthcare," *Pervasive Mob. Comput.*, vol. 86, p. 101683, 2022, doi: 10.1016/j.pmcj.2022.101683.

[6] Y. Liang, S. Samtani, B. Guo, and Z. Yu, "Behavioral biometrics for continuous authentication in the Internet-of-Things era: An artificial intelligence perspective," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 9128–9143, 2020, doi: 10.1109/jiot.2020.3004077.

[7] P. Bothra, R. Karmakar, S. Bhattacharya, and S. De, "How can applications of blockchain and artificial intelligence improve performance of Internet of Things? – A survey," *Comput. Netw.*, vol. 224, p. 109634, 2023, doi: 10.1016/j.comnet.2023.109634.

[8] S. Furnell, "Assessing website password practices – Unchanged after fifteen years?," *Comput. Secur.*, vol. 120, p. 102790, 2022, doi: 10.1016/j.cose.2022.102790.

[9] L. A. Meyer, S. Romero, G. Bertoli, T. Burt, A. Weinert, and J. L. Ferres, "How effective is multifactor authentication at deterring cyberattacks?," *arXiv*, 2023, doi: 10.48550/arxiv.2305.00945.

[10] B. Jeon et al., "A facial recognition mobile app for patient safety and biometric identification: Design, development, and validation," *JMIR Mhealth Uhealth*, vol. 7, no. 4, p. e11472, 2019, doi: 10.2196/11472.

[11] R. R. Datta, D. D. Suman, C. Nabanita, and K. D. Ranjan, "Biometric-based computer vision for boundless possibilities: Process, techniques, and challenges," Oct. 31, 2024. doi: 10.1049/pbpc064e_ch3.

[12] J. Kim, T. Ng, and A. J. Teoh, "Conditional deployable biometrics: Matching periocular and face in various settings," *J. Inform. Web Eng.*, vol. 3, no. 3, 2024, doi: 10.33093/jiwe.2024.3.3.19.

[13] V. Upadhyaya, "Advancements in computer vision for biometrics enhancing security and identification," in *Leveraging Computer Vision to Biometric Applications*, Chapman and Hall/CRC, 2025, pp. 260–292. doi: 10.1201/9781032614663-14.

[14] L. V. Chernenkaya, E. N. Desyatirikova, and A. V. Rechinskii, "Realization of computer vision system for biometric identification of personality," in *2021 Int. Russ. Autom. Conf. (RusAutoCon)*, 2021, pp. 409–414. doi: 10.1109/rusautocon52004.2021.9537374.

[15] F. Hashmi, K. Ashish, S. Katiyar, and A. Keskar, "Computer vision in contactless biometric systems," *Int. Arab J. Inf. Technol.*, vol. 18, no. 3, pp. 484–492, 2021, doi: 10.34028/iajit/18/3a/12.

[16] E. Chen, "Machine learning in cybersecurity: Computer vision for biometric authentication systems," *Afr. J. Artif. Intell. Sustain. Dev.*, vol. 3, no. 2, pp. 367–373, 2023.

[17] G. Zou, G. Fu, M. Gao, J. Pan, and Z. Liu, "A new approach for small sample face recognition with pose variation by fusing Gabor encoding features and deep features," *Multimed. Tools Appl.*, vol. 79, no. 31, pp. 23571–23598, 2020, doi: 10.1007/s11042-020-09076-1.

[18] J. Y. Choi and B. Lee, "Ensemble of deep convolutional neural networks with Gabor face representations for face recognition," *IEEE Trans. Image Process.*, vol. 29, pp. 3270–3281, 2020, doi: 10.1109/tip.2019.2958404.

[19] M. T. Masud, M. Keshk, N. Moustafa, I. Linkov, and D. K. Emge, "Explainable artificial intelligence for resilient security applications in the Internet of Things," *IEEE Open J. Commun. Soc.*, vol. PP, p. 1, 2024, doi: 10.1109/ojcoms.2024.3413790.

[20] S. I. Serengil and A. Ozpinar, "LightFace: A hybrid deep face recognition framework," in *2020 Innov. Intell. Syst. Appl. Conf. (ASYU)*, 2020, pp. 1–5. doi: 10.1109/asyu50717.2020.9259802.

[21] W. Cui, W. Zhan, J. Yu, C. Sun, and Y. Zhang, "Face recognition via convolutional neural networks and Siamese neural networks," in *2019 Int. Conf. Intell. Comput. Autom. Syst. (ICICAS)*, 2019, pp. 746–750. doi: 10.1109/icicas48597.2019.00161.

[22] M. Zulfiqar, F. Syed, M. J. Khan, and K. Khurshid, "Deep face recognition for biometric authentication," in *2019 Int. Conf. Electr., Commun., Comput. Eng. (ICECCE)*, 2019, pp. 1–6. doi: 10.1109/icecce47252.2019.8940725.

[23] R. Goel, M. Alamgir, H. Wahab, M. Alamgir, H. Ugail, and I. Mehmood, "Sibling discrimination using linear fusion on deep learning face recognition models," *J. Inform. Web Eng.*, vol. 3, no. 3, 2024, doi: 10.33093/jiwe.2024.3.3.14.

[24] M. Heidari and K. Fouladi-Ghaleh, "Using Siamese networks with transfer learning for face recognition on small-samples datasets," in *2020 Int. Conf. Mach. Vis. Image Process. (MVIP)*, 2020, pp. 1–4. doi: 10.1109/mvip49855.2020.9116915.

[25] N. H. Tandel, H. B. Prajapati, and V. K. Dabhi, "Voice recognition and voice comparison using machine learning techniques: A survey," in *2020 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, 2020, pp. 459–465. doi: 10.1109/icaccs48705.2020.9074184.

[26] V. Passricha and R. K. Aggarwal, "A hybrid of deep CNN and bidirectional LSTM for automatic speech recognition," *J. Intell. Syst.*, vol. 29, no. 1, pp. 1261–1274, 2020, doi: 10.1515/jisys-2018-0372.

[27] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019, doi: 10.1109/access.2019.2896880.

[28] W. Han et al., "ContextNet: Improving convolutional neural networks for automatic speech recognition with global context," in *Proc. Annu. Conf. Int. Speech Commun. Assoc. (INTERSPEECH)*, vol. 2020-Octob, no. 1, pp. 3610–3614, 2020, doi: 10.21437/interspeech.2020-2059.

[29] S. F. Tan and G. C. Chung, "An evaluation study of user authentication in the Malaysian FinTech industry with uAuth security analytics framework," *J. Cases Inf. Technol.*, vol. 25, no. 1, pp. 1–27, 2023, doi: 10.4018/jcit.318703.

[30] E. Mostafa, M. M. Hassan, and W. Said, "An interactive multi-factor user authentication framework in cloud computing," *Int. J. Comput. Sci. Netw. Secur.*, vol. 23, no. 8, p. 63, 2023, doi:10.22937/ijcsns.2023.23.8.8.

[31] E. Cadet, O. S. Osundare, H. O. Ekpobimi, Z. Samira, and Y. Wondaferew, "Cloud migration and microservices optimization framework for large-scale enterprises," *Open Access Res. J. Eng. Technol.*, no. October, 2024, doi: 10.53022/oarjet.2024.7.2.0059.

[32] F. Yang, C. Ma, J. Zhang, J. Zhu, W. Yuan, and A. Owens, "Touch and go: Learning from human-collected vision and touch," *arXiv*, 2022, doi: 10.48550/arxiv.2211.12498.