# Fingerprint Identification and Commonality Analysis Using Deep Learning

Ashish Rameshbhai Deriya*, Urvashiben Thakkar*, Yesha Sureshbhai Patel*, Uha Ratna Sudha Achanti*, Prem Kumar Jami*, CJ Chung**
(*) MSCS Candidate (**) PhD, Professor
College of Arts & Science, Lawrence Technological University

## ABSTRACT

Fingerprint biometrics are integral to digital authentication and forensic science. Fingerprint identification serves as a fundamental aspect of biometric security, yet delving into the individuality and potential commonalities of human fingerprints poses interesting questions. This study employs deep learning methodologies to address two primary research questions. Firstly, we explore the feasibility of training deep learning models to identify single person fingerprints. Through Binary Classification tasks, utilizing datasets consisting of fingerprints from two individuals, each containing 400 fingerprint images, we evaluate the efficacy of different CNN and pretrained models (such as VGG16, ResNet50, and EfficientNetB3). Our analysis focuses on evaluating the test accuracy of these models in identifying fingerprints, providing insights into the potential of deep learning for personalized fingerprint recognition. Secondly, we investigate the existence of commonalities or common features between fingerprints of different 17 individuals. Through a series of Binary Classification tasks and Multi-Class Classification tasks, utilizing datasets totaling 3360 fingerprint images, we explore various aspects of commonality. Specifically, in Multi-Class Classification tasks 2 and 3, we examine the impact of excluding specific fingerprint samples from around 17 people, such as excluding the left index finger, on model accuracy. Additionally, for Multi-Class Classification task 3, utilizing the pretrained EfficientNetB3 model, we achieved an F1 score of 0.859, indicating the model's effectiveness in identifying commonalities among fingerprints. Our findings highlight the potential of deep learning in personalized fingerprint recognition and shed light on shared traits among fingerprints, offering implications for enhanced biometric security and forensic analysis. These insights contribute to the development of more reliable authentication systems and aid in the administration of justice. Future research directions include expanding dataset sizes for improved model performance, exploring alternative hidden fingers during training, automating testing procedures, interpreting neural networks for deeper understanding of fingerprint characteristics, and investigating gender prediction and life expectancy estimation based on fingerprints. Moreover, techniques to determine if fingerprints from different crime scenes belong to the same individual could enhance forensic analysis capabilities.

**Keywords:** CNN, EfficientNetB3, VGG16, ResNet50, Fingerprints, Biometric Security, Forensic Analysis, Deep Learning.

## DATA COLLECTION

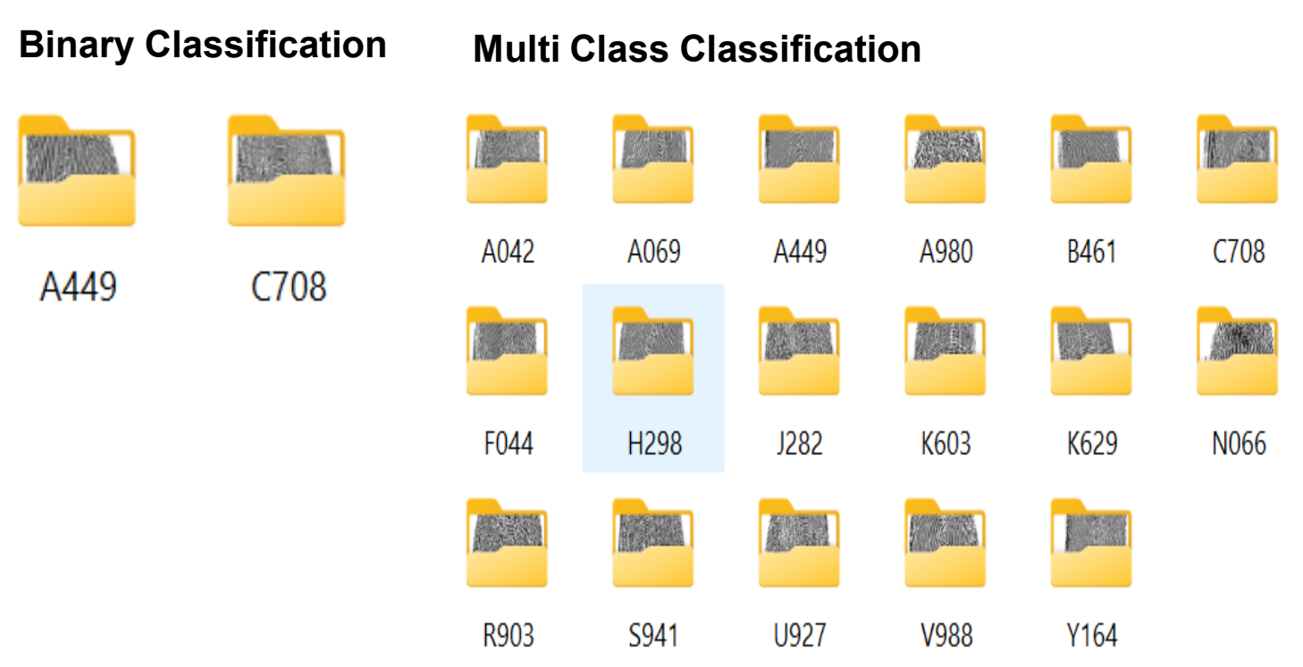
**Data Collection Equipment:**
The data collection process involved using MyFingerprint application.

**Data Classification and Class Settings:**
Our dataset encompasses binary and multi-class classification tasks for fingerprint identification. Tasks include differentiating person 0 and person1 fingerprints, as well as identifying individuals from a pool of 17 people's fingerprints, with variations like excluding specific fingerprints for training, validation and testing. In this data setting, data is splitted in 80%, 10% and 10% ratio for training, validation and testing respectively.

| Dataset Info | Classes | Total Model Training Images | Total Model Validation Images | Total Model Testing Images |
|---|---|---|---|---|
| Binary Classification 1 | Person 0<br>Person 1 | **320** | **40** | **40** |
| Binary Classification 2 | Person 0<br>Person 1 | **322**<br>Model trained without person 0's left index fingerprint | **38**<br>Model validate without person 0's left index fingerprint | Tested on person 0's left index fingerprint<br>**20 for test 1**<br>**40 for test 2** |
| Multi-Class Classification 1 | 17 Individuals | **2680** | **340** | **340** |
| Multi-Class Classification 2 | 17 Individuals | **2772**<br>Model trained without person 0's left index fingerprint | **338**<br>Model validate without person 0's left index fingerprint | Tested on person 1's left index fingerprint<br>**20 for test 1**<br>**340 for test 2** |
| Multi-Class Classification 3 | 17 Individuals | **2715**<br>Models trained without left index fingerprints | **306**<br>Model Validate without left index fingerprints | Tested on left index fingerprints of all individuals<br>**339** |

Table 1: All task Dataset Information

**Acknowledgment**: We thank all 17 individuals who provided their fingerprint image data for this research.



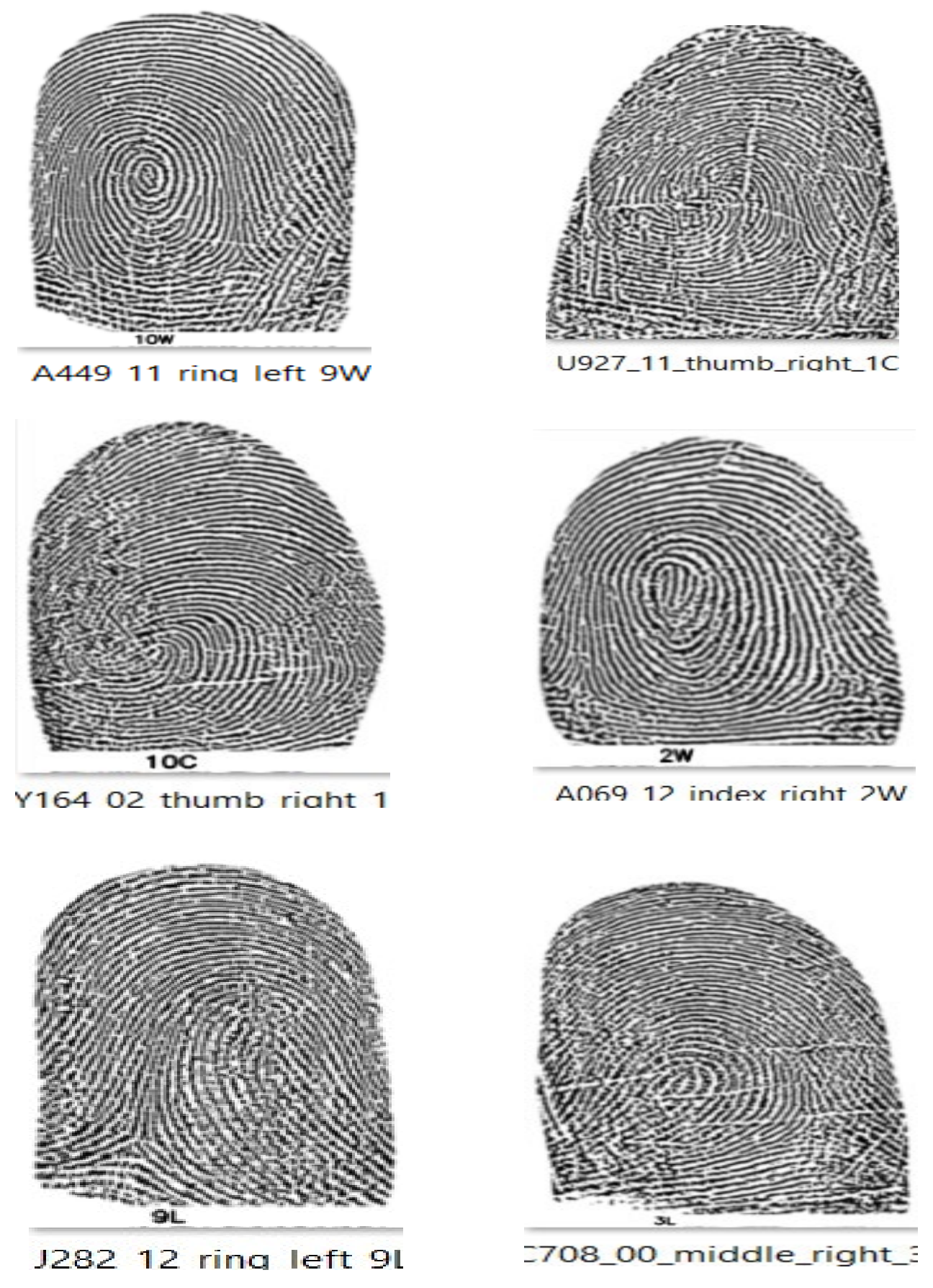


Figure 1: The Fingerprint Examples from some classes

## MODEL EXPERIMENTS

This research employed two distinct models, a Convolutional Neural Network (CNN) and a pre-trained EfficientNetB3, to investigate fingerprint identification.

**CNN Model Experiment:**

**Data Preprocessing:** Data sourced from Google Drive was processed, splitting it into training, validation, and testing sets.
**Data Augmentation:** To enrich the training data and improve model performance, data augmentation techniques were applied.
**Model Architecture:** A CNN architecture was meticulously designed to suit the fingerprint identification task.
**Overfitting Mitigation:** Techniques like early stopping and model checkpointing were implemented to prevent overfitting during training.
**Training Process:** The model underwent training, with validation employed to monitor and optimize training progress.
**Model Evaluation:** Upon completion of training, the model's performance was assessed using the unseen test dataset.

**Pre-trained EfficientNetB3 Model Experiment:**

**Model Initialization**: Model Initialization: The pre-trained EfficientNetB3 model was initialized, and its layers were initially frozen to retain the learned features from the pre-training phase.
**Basic Architecture:** A basic model architecture was constructed, incorporating data augmentation techniques.
**Model Compilation:** The model was compiled in preparation for training.
**Training and Fine-tuning:** We initiated the process by training the model to establish a performance baseline. Following this, we unfroze specific layers to fine-tune the model, enhancing its suitability for fingerprint identification tasks. Additional training rounds were then carried out to further enhance the model's overall performance and adaptability.
**Evaluation Metrics:** Model effectiveness was evaluated using metrics such as F1 score and confusion matrix analysis on the test dataset.

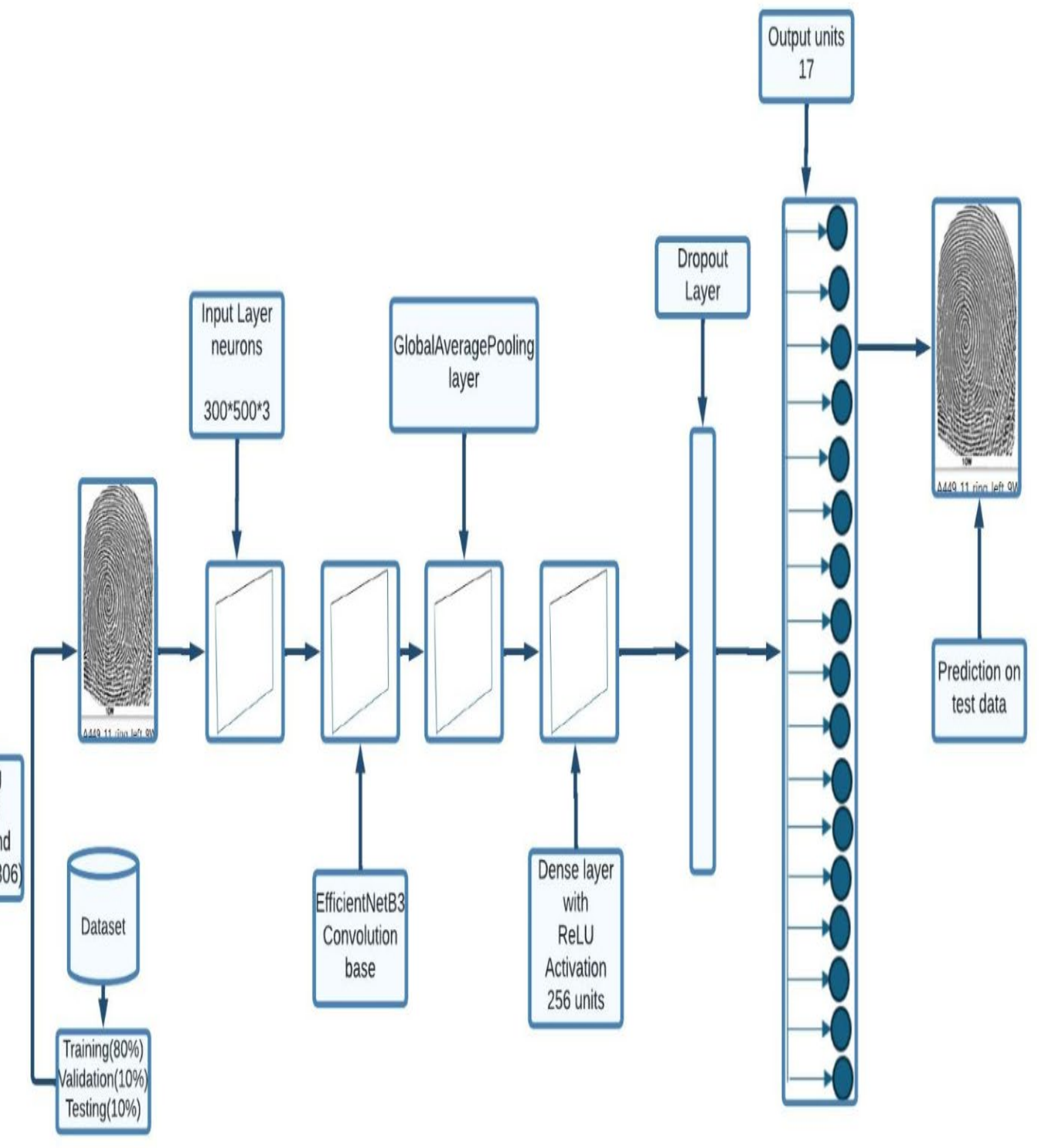


Figure 2: Pre-Trained EfficientNetB3

## RESULTS

The utilization of pre-trained models, specifically EfficientNetB3, ResNet50, and VGG16, showcased superior performance compared to custom CNN models. By leveraging transfer learning and fine-tuning techniques, the pre-trained models demonstrated robustness and adaptability in capturing intricate features of fingerprints. EfficientNetB3, in particular, stood out for its innovative scaling methodology and parameter efficiency, achieving notable accuracy scores across binary and multi-class classification tasks. Therefore, we will discuss the results of Multi-Class Classification tasks Multi-Class Classification 1, Multi-Class Classification 2, and Multi-Class Classification 3 using the EfficientNetB3 pre-trained model, along with performance metrics and analysis. It's important to note that in Binary Classification 2, Multi-Class Classification 2, and Multi-Class Classification 3, unseen left index fingerprints were used to test if our trained models can recognize the person. This approach ensures a rigorous evaluation of our models' ability to generalize to unseen data and accurately identify individuals, further validating the effectiveness and reliability of our approach in personalized fingerprint recognition.

### Pre-trained EfficientNetB3 Results

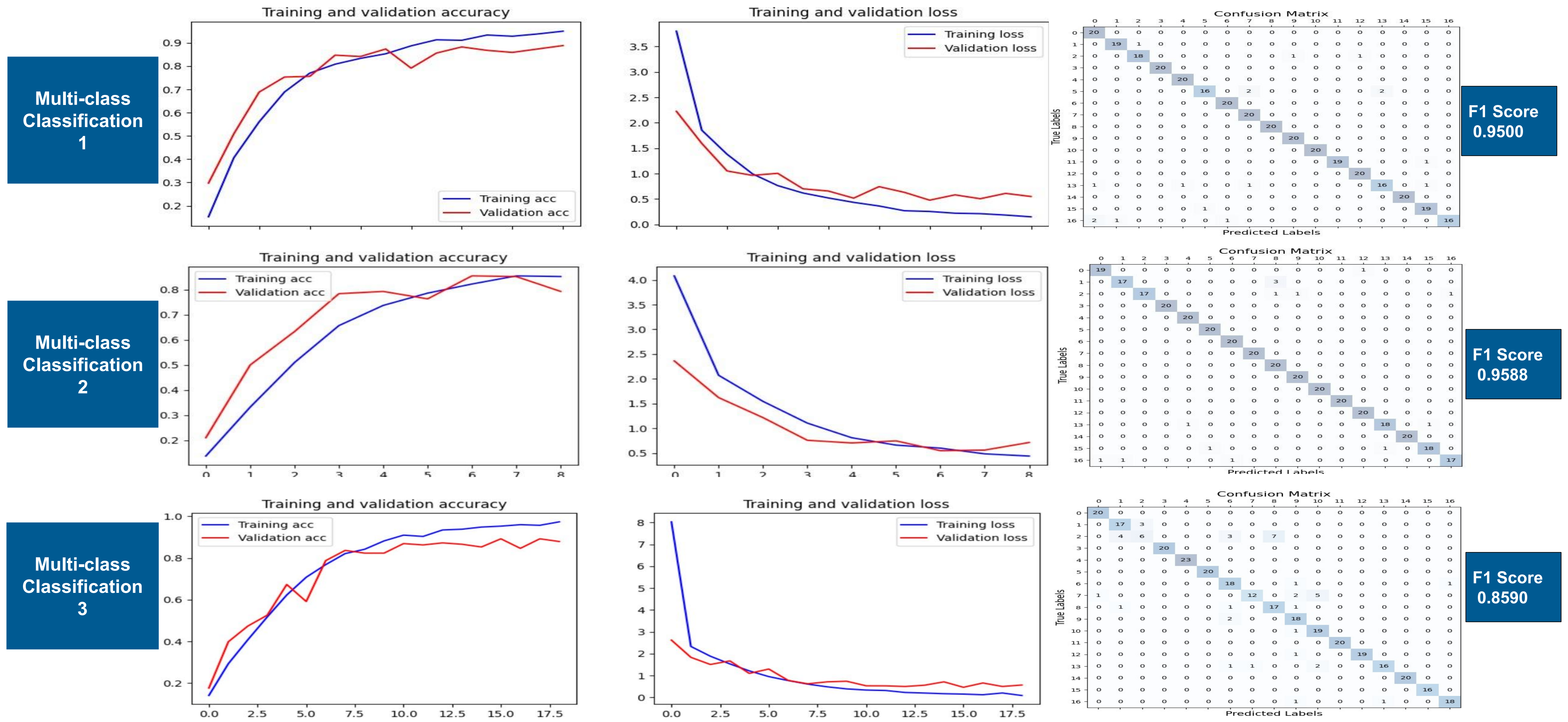


Figure 3: Accuracy, Loss, Confusion Matrix and F1 Score

| Model Information | CNN-1 | CNN-2 | CNN-3 | Pre-trained Efficient Netb3 | Pre-trained ResNet 50 | Pre-trained VGG16 |
|---|---|---|---|---|---|---|
| Data Augmentation | Yes | Yes | No | Yes | Yes | No |
| Input Size | 300/500 | 300/500 | 600/600 | 300/500 | 300/500 | 600/600 |
| Training Epochs | 30 | 30 | 30 | 30 | 30 | 30 |
| Early Stopping | Yes | Yes | No | No | No | No |
| Fine-Tuning | No | No | No | Yes | Yes | No |
| Test Accuracy | | | | | | |
| | | | | | | |
| Binary Classification 1<br>All fps of P0 and P1 | 87.5 | 95 | 95 | 97.50 | 94 | 100 |
| Binary Classification 2<br>Eval 1: Left index fps of P0 only | 100 | 70 | 90 | 100 | 30 | 100 |
| Eval 2: Left index fps of P0 + all fps of P1 | 62.5 | 62.5 | 95 | 97.5 | 94 | 100 |
| Multiclass Classification 1<br>ALL fps of p0 and p(N-1) | 80.58 | 83 | 88.20 | 94.41 | 90 | 91 |
| Multiclass Classification 2<br>Eval 1: Left index fps of P0 | 75 | 80 | 89 | 100 | 55 | 90 |
| Eval 2: Left index fps of P0 + all fps of P1 – P(N-1) | 78.23 | 80 | 86 | 94.99 | 83.52 | 82 |
| Multiclass Classification 3<br>Left Index Fps of everyone, P0-P(N-1) | 79.84 | 66.96 | 73 | 86.89 | 82.89 | 73 |

Table 2: All task Test accuracies

## SUMMARY and CONCLUSION

**CNN Performance:**
All three CNN models were trained for the same number of epochs (30) with early stopping implemented for CNN-1 and CNN-2, potentially preventing overfitting. The input sizes for CNN-1 and CNN-2 were smaller (300x500) compared to CNN-3 (600x600), which may suggest that CNN-3 could capture more detailed features but did not necessarily translate to higher accuracies across all tasks. None of the custom CNN models used fine-tuning or learning rate adjustments, which may have limited their performance when compared to the pre-trained models that utilized these techniques.

**Pretrained Performance :**
The pre-trained models had varying input sizes, with EfficientNetB3 and ResNet50 having a medium input size (300x500 for EfficientNetB3 and 500x500 for ResNet50) and VGG16 having the largest (600x600). This did not straightforwardly correlate with performance, as EfficientNetB3 often outperformed the others despite not having the largest input size.
Both EfficientNetB3 and ResNet50 were fine-tuned, which likely contributed significantly to their strong performance, as they could adjust pre-learned weights to better fit the specific dataset.
Early stopping was not used in any of the pre-trained models, suggesting confidence in the convergence of the models within the set epochs.
The absence of data augmentation for all models implies that the results are based on the models' abilities to learn from the given datasets without additional variance from augmented data.

**EfficientNetB3:**
EfficientNetB3, a member of the EfficientNet model family, stands out for its compound scaling methodology, which balances network depth, width, and resolution to optimize accuracy and efficiency. This innovative scaling approach, combined with its design for parameter efficiency, enables EfficientNetB3 to deliver superior performance with fewer parameters than earlier architectures like VGG16 and ResNet50, making it not only faster but also more resource-effective for both training and inference. Furthermore, its strong performance is consistent across various image recognition tasks, and it particularly excels when fine-tuned for specific datasets, showcasing an adaptability that may surpass that of its predecessors. The model's versatility and reduced computational demands make it an optimal choice for deployment in resource-constrained environments, thus marking it as a leading option for modern convolutional neural network applications.

**Conclusion:**
Based on our research findings, it's evident that deep learning models, particularly pre-trained ones such as EfficientNetB3, VGG16, and ResNet50, demonstrate remarkable efficacy in fingerprint recognition tasks. The implications of our study extend to enhancing biometric security measures and bolstering forensic analysis capabilities. By leveraging deep learning methodologies, we pave the way for the development of more reliable authentication systems, which in turn contributes to the administration of justice. Moving forward, expanding dataset sizes and exploring alternative training methodologies could further enhance model performance and deepen our understanding of fingerprint characteristics, ultimately advancing the field of biometric security and forensic science.

## FUTURE WORK

Building upon the insights gained from this study, several avenues for future research can be explored to advance the field of fingerprint identification and deepen our understanding of biometric security:

**Diverse Data Collection:** Collecting a more diverse range of fingerprint datasets from various demographics, ethnicities, and environmental conditions will enhance the robustness and generalization capabilities of the models. This will enable better adaptation to real-world scenarios and improve overall performance.

**Advanced Pre-processing Techniques:** Investigating advanced pre-processing methods and exploring deep learning architectures tailored to handle various fingerprint characteristics, including noise, low contrast, and complex background patterns, will be crucial for improving model accuracy and reliability.

**Model Optimization:** Experimenting with different CNN architectures and conducting thorough hyperparameter tuning will contribute to optimizing model performance, particularly in scenarios with larger datasets and complex fingerprint variations. Balancing noise reduction with detail preservation will also be essential for refining the models.

**Expanding Research Scope:** Broadening the scope of research to include additional biometric modalities, such as palm prints or iris scans, and exploring interdisciplinary approaches with computer vision and pattern recognition will enrich the understanding of biometric security and identification techniques.

By pursuing these future research directions, the field of fingerprint identification and biometric security can progress towards more robust, accurate, and reliable systems, contributing to enhanced authentication methods and bolstering security measures in various domains.

## REFERENCES

[1] G. Guo, A. K. Ray, M. Izydorczak, J. Goldfeder, H. Lipson, and W. Xu, "Unveiling intra-person fingerprint similarity via deep contrastive learning," *Science Advances*, vol. 10, no. 2, Jan. 2024, doi: https://doi.org/10.1126/sciadv.adi0329.

[2] AI Technology & Systems, "Fingerprint pattern classification using Deep Learning," *AITS Journal*, Aug. 27, 2021. https://medium.com/ai-techsystems/fingerprint-pattern-classification-using-deep-learning-9eb93757df11

[3] "MyFingerprint - Apps on Google Play," *play.google.com*. https://play.google.com/store/apps/details?id=com.issart.fingerprints.

[4] "Convolutional Neural Network (CNN) Architectures," *GeeksforGeeks*, Mar. 21, 2023. (accessed Apr. 16, 2024). https://www.geeksforgeeks.org/convolutional-neural-network-cnn-architectures/?ref=lbp

[5] K. Team, "Keras documentation: Keras Applications," *keras.io*. https://keras.io/api/applications/

[6] "AI Discovers That Not Every Fingerprint Is Unique," *Columbia Engineering*, Jan. 10, 2024. https://www.engineering.columbia.edu/news/ai-discovers-not-every-fingerprint-unique