FacePass: A Facial Identification Powered Logbook for Libraries Using Internet of Things

A Case Study proposal presented to the

College of Computer Studies of

Pamantasan ng Lungsod ng Pasig

In partial fulfillment of the requirements for the course

Programming Languages

Prepared by:

Bernardino, Joshua Vincent M.

Villanueva, Francis Niño E.

Yap, Louise Ann E.

Engr. Greta M. Rosario

Adviser

May 2025

1.0 Introduction

The development of deep learning has particularly enhanced the effectiveness of facial recognition technologies. This case study will focus on the advancement of three pioneering works in facial recognition convolutional neural networks (CNN): FaceNet, VGG, and Inception-ResNet V2.

FaceNet maps face images into a high-dimensional concept space where the distance of vectors indicates similarity, and was used for comparing facial similarity. This is done by training a neural network on millions of face images, and then using the model to predict features for a full face to build a representation (Schroff, F. et al., 2015). The layered structure of VGG resulted in deep specialization, thereby improving accuracy in facial recognition and identification, which now includes gender classification and emotion detection, enhancing the scope of its application in multimodal biometric systems (Rabea, M. et al., 2024). In the field of face recognition, Inception-ResNet-V2 is known for creating highly accurate and robust facial embeddings. This high-performance convolutional neural network employs residual connections alongside Inception modules for enhanced accuracy in detail classification such as delineating frames of features on the face.

The case study aims to analyze these models' efficiency in facial recognition considering their accuracy, computational effectiveness, and usefulness. The researchers want to identify the most effective method for integrating deep learning-based facial recognition into practical systems by comparing their strengths and limitations.

2.0 Purpose of the Study

The purpose of this study is to conduct a comparative analysis of three widely recognized convolutional neural network (CNN) models—*FaceNet, VGG, and Inception-ResNet V2*—to evaluate their efficiency, accuracy, and computational effectiveness in facial recognition tasks. These models have demonstrated significant advancements in deep learning-based facial recognition, but each has distinct strengths and limitations.

This study aims to:

- **Evaluate Accuracy:** Measure how well each model can identify and distinguish facial identities, including recognition accuracy and error rates.
- **Computational Efficiency:** You may need to measure the amount of processing time (as well as memory requirement) used by each model to process the data, and also the possibility of deploying it in real-time settings.
- **Determine practical application:** Look into which model is applicable in practice such as security, authentication, biometrics systems.

By comparing FaceNet's high-dimensional feature mapping, VGG's structured deep learning approach, and Inception-ResNet V2's hybrid residual-inception architecture, the researchers try to find the best way to integrate deep learning based facial recognition model with a practical system.

With the use of the public dataset of facial images, the researchers present a comparative study between the three models under a consistent setup to clarify clear understanding of their strengths and limitations. The results are expected to provide developers, researchers, and system integrators with suggestions on which type(s) of CNN-based model will be more suitable for efficient and reliable face recognition across different applications.

3.0 Comparison of Algorithms and Discussion

The study uses three different models: FaceNet, Inception-ResNet v2, and VGG-19. Three different CNN models with differing algorithms. FaceNet uses an algorithm that generates high-quality face mapping with machine-learning architectures, Inception-ResNet v2 is almost the same as Inception using the two 3x3 convolutions and incorporates the use of residual connections instead of the concatenation part usually seen in the Inception architecture, and VGG-19 that uses with 19 weight layers comprised with 16 convolutional layers and 3 fully connected layers.

Using the Confusion Matrix, the FaceNet model achieved the highest performance with accuracy, macro average, and weighted average of 99% in precision, recall, and f1-score. VGG-19 reached an accuracy, macro average, and weighted average of 28%, 27%, and 27% respectively on all metrics. The Inception-ResNet v2 reached the lowest accuracy, macro average, and weighted average of 25% on all classification evaluation metrics all with 2206 support and using the Pins Face Recognition dataset. It is possible that the two models had low accuracy as a result of the small image resolution of the images inside the dataset forcing them to upscale the images. FaceNet accepts 160x160 image resolution compared to the two models, Inception-ResNet v2 with 224x224 and VGG-19 with 299x299.

4.0 Conclusion

This research examined and compared the performance of three convolutional neural network (CNN) models: FaceNet, VGG-19, and Inception-ResNet V2 in the task of face recognition. The models were chosen based on their architectural strengths and performance in prior studies. Inception-ResNet V2 was acknowledged for its computational efficiency through the combination of residual and inception modules in its design, FaceNet for its face mapping function, and VGG-19 for its feature extraction capabilities utilized by its layered structure.

FaceNet was the best model, according to the analysis, with an accuracy of 99.05%, an inference time of 327.7 seconds, and a feedback delay of 0.1467 seconds. Compared to VGG-19 with 27.33% and Inception-ResNet V2 with 24.93% accuracy, which was less accurate and required more processing, these results clearly show better efficiency. This is mainly because of the dataset's resolution restrictions.
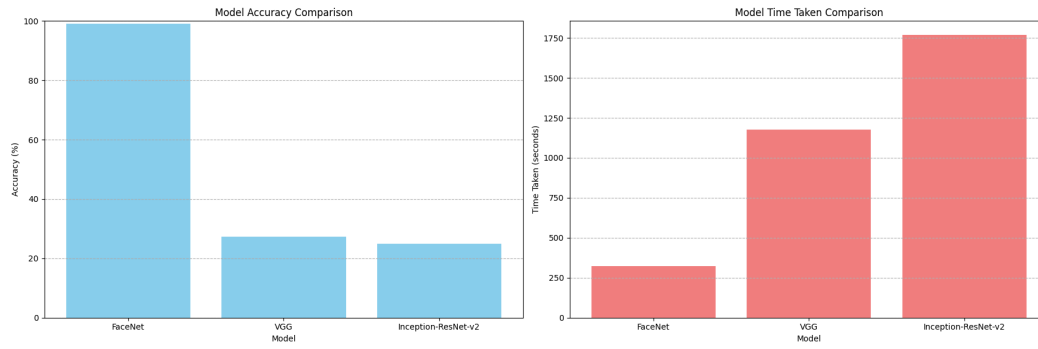

Figure 1. Comparison of Models by Its Accuracy and Time Taken

In conclusion, FaceNet is the most suitable for full and high-speed facial recognition systems particularly when performance and speed are significant factors. Its effectiveness makes it eligible for real-time application in services such as verification, security, and biometrics. The outcome of this research can be used as a valuable source of data for researchers, developers, and system designers planning to apply deep learning-based facial recognition in real-world systems.

References:

Pins face recognition. (2020, March 7). Kaggle. *https://www.kaggle.com/datasets/hereisburak/pins-face-recognition?fbclid=IwY2xjawKj4 v1leHRuA2FlbQIxMABicmlkETFrQ1g2Q2w3eHVKNnhLWUMzAR4xurxPVNqxaadx9y8I c1TZjJD_1mBHwR1tDYaUxRfPWFlaivi75N24Z0ZkOA_aem_cSBH3fo58IEUCr4WAZTA vw*

Rabea, M., Ahmed, H., Mahmoud, S., & Sayed, N. (2024, January 2). *IdentiFace : a VGG based multimodal facial biometric system.* arXiv.org. https://arxiv.org/abs/2401.01227

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 815-823).

Zhang, H. (2024). Real-time face recognition method based on MTCNN-Inception-ResNet-v2-SVM model. *Applied and Computational Engineering*, *45*(1), 179–189. *https://doi.org/10.54254/2755-2721/45/20241677*