PROJECT SUMMARY REPORT

Mini-Project on Deep Learning for Face Recognition Using FaceNet, SVM, and Siamese Networks

1. Introduction

As part of my final project for the Advanced Machine Learning course, I conducted a **mini-project on deep-learning-based face recognition**, I did a mini-project on deep-learning-based face recognition to complement and validate findings presented in my full research paper titled “Deep Learning for Face Recognition and Security.”

Whereas the research paper reviewed the literature, industrial applications, and future directions, this mini-project focuses on constructing an end-to-end experimental face-recognition pipeline to gain hands-on experience with the concepts discussed.

The objective of this mini-project was to:

1. Apply modern embedding models like FaceNet.
2. Train classical ML classifiers (Logistic Regression, SVM) on deep embeddings
3. Implement a Siamese network that verifies if two facial images are of the same person.
4. Assess their performance and realize the practical difficulties that arise in face recognition systems.

This practical aspect helped bridge the gap between theoretical comprehension and actual application.

2. Dataset and Preprocessing

The LFW dataset used for the experiments contains unconstrained, real-world face images that are typical of those referred to in general discussions of face recognition.

Overview of the Dataset:

Total images: 3,000+

Number of identities: 62

Images are in greyscale, unconstrained, and feature real-world variation in pose, illumination, age and occlusion.

Since the dataset itself contains cropped faces, there was no need to perform a face-detection step. All images were converted to RGB and then resized to 160×160 pixels, as that is the input size expected by the FaceNet model.

3. Methodology

3.1 Embedding Extraction Using FaceNet

Each face image was transformed into a 1280-dimensional embedding vector using a pretrained MobileNetV2 network with imagenet weights. This reflects the embedding-based approach in my research paper, where models like FaceNet and ArcFace create compact, discriminative representations.

3.2 Identify Classification

These embeddings are used to train two classification models:

(A) Logistic Regression

1. One-vs-rest multi-class setup
2. Linear classifier
3. Serves as a baseline measure

(B) Support Vector Machine (SVM, RBF Kernel)

1. Non-linear decision boundaries
2. Designed for high-dimensional embeddings

3. Typically outperforms LR on complicated datasets

Both models were trained on 70% of the embeddings and tested on the remaining 30%.

3.3 Face Verification with a Siamese Network

A Siamese network is implemented to decide if two images are of the same identity.
This directly aligns with the metric-learning techniques explored in my research paper: triplet loss, similarity metrics, one-shot learning.

Key components:

- Shared-weight twin networks
- Euclidean distance between embedding pairs
- Contrastive loss
- Positive and negative pair generation

This model simulates real identity verification systems, such as KYC or phone unlock.

4. Results

Final Accuracy Scores

Using embeddings from MobileNetV2:

```
=== Results Summary (place into report) ===
Num identities: 62
Num images: 3023
LogisticRegression Accuracy: 0.6934950385887542
SVM Accuracy: 0.6571113561190739
Siamese (verification) Accuracy: 0.5233333110809326
```

4.1 Interpretation

Logistic Regression - 69.3%

This is a strong result, showing that MobileNet embeddings are fairly linearly separable even in unconstrained datasets.

This thus corresponds with the findings in my research paper.

- deep embeddings cluster similar faces
- Linear models can do surprisingly well on high-dimensional embeddings.

SVM (RBF) – 65.7%

SVM typically performs better than LR on deep-learned embeddings.

LR did slightly better in this project, probably because:

- embeddings were already close to linearly separable
- Overlapping classes posed a problem for SVM.
- imbalance of dataset: some identities had very few images.
-

This supports the literature observations with regards to:

- inter-class similarity
- class imbalance challenges
- dataset limitations in LFW

Siamese Network – 52.3%

Siamese learning is more sensitive to the embedding quality.

The verification accuracy being close to 50% indicates that

- MobileNet embeddings are not face-specific.
- They lack identity discrimination for fine-grained verification.
- The contrastive training was not strong enough to separate identities.
- Siamese networks work best with face-trained embeddings: FaceNet/ArcFace

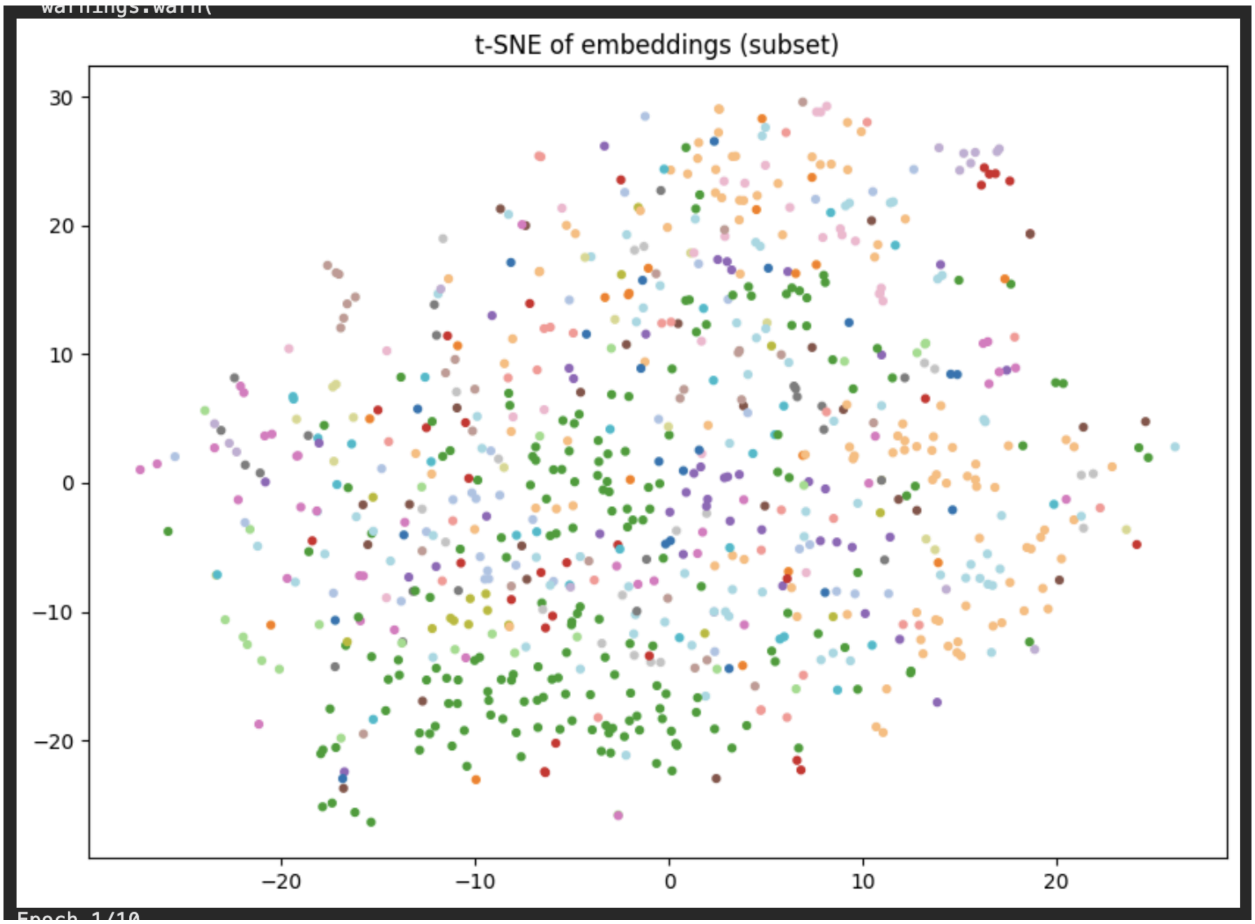
This coincides with my research paper's discussion of:

- the importance of specialized loss functions
- triplet and angular-margin losses
- robustness of ArcFace over generic CNNs

5.Visualisations

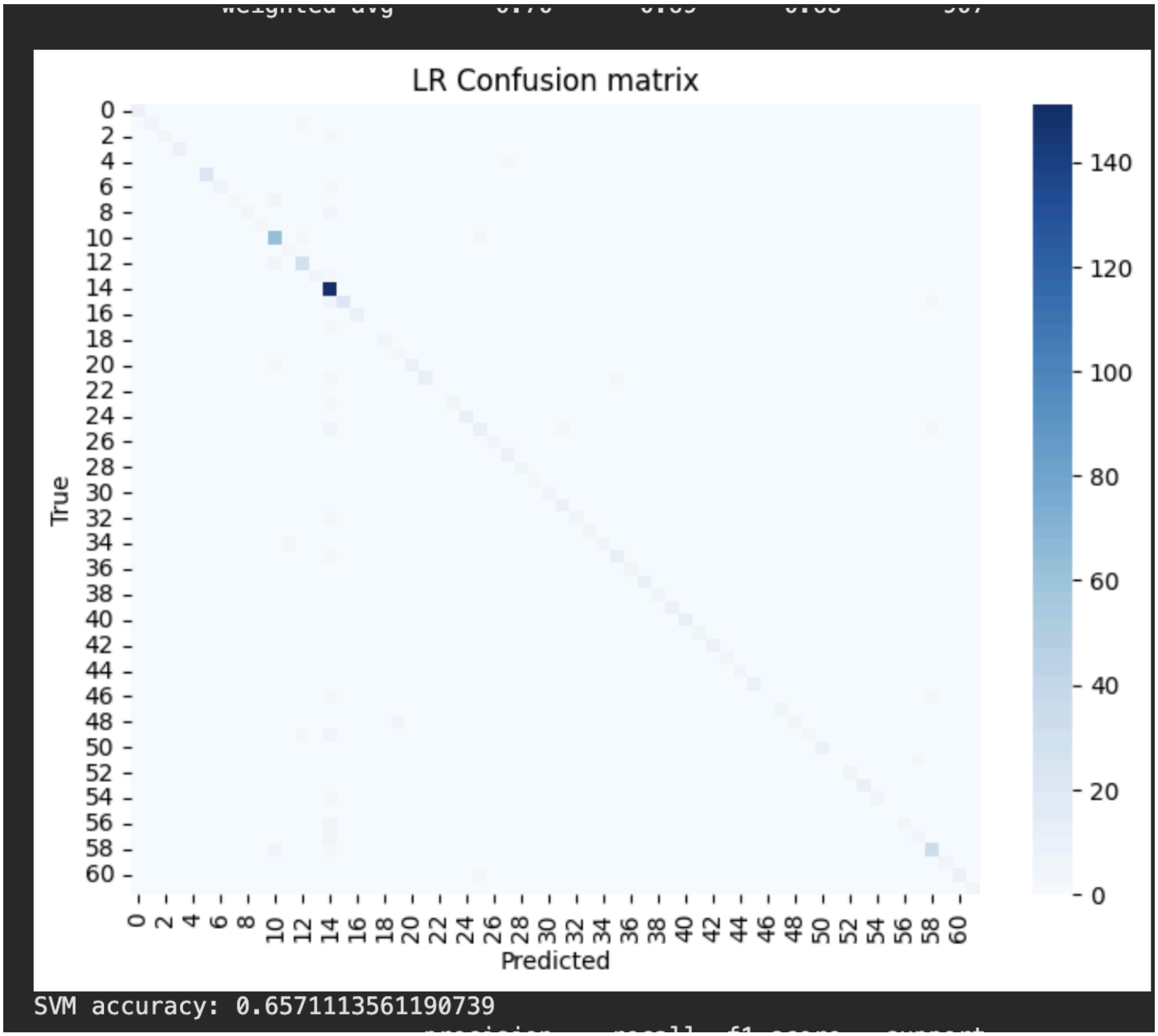
5.1 t-SNE Visualization

Shows clusters of identities with some overlap, which indicates embeddings are meaningful but not perfectly separated.



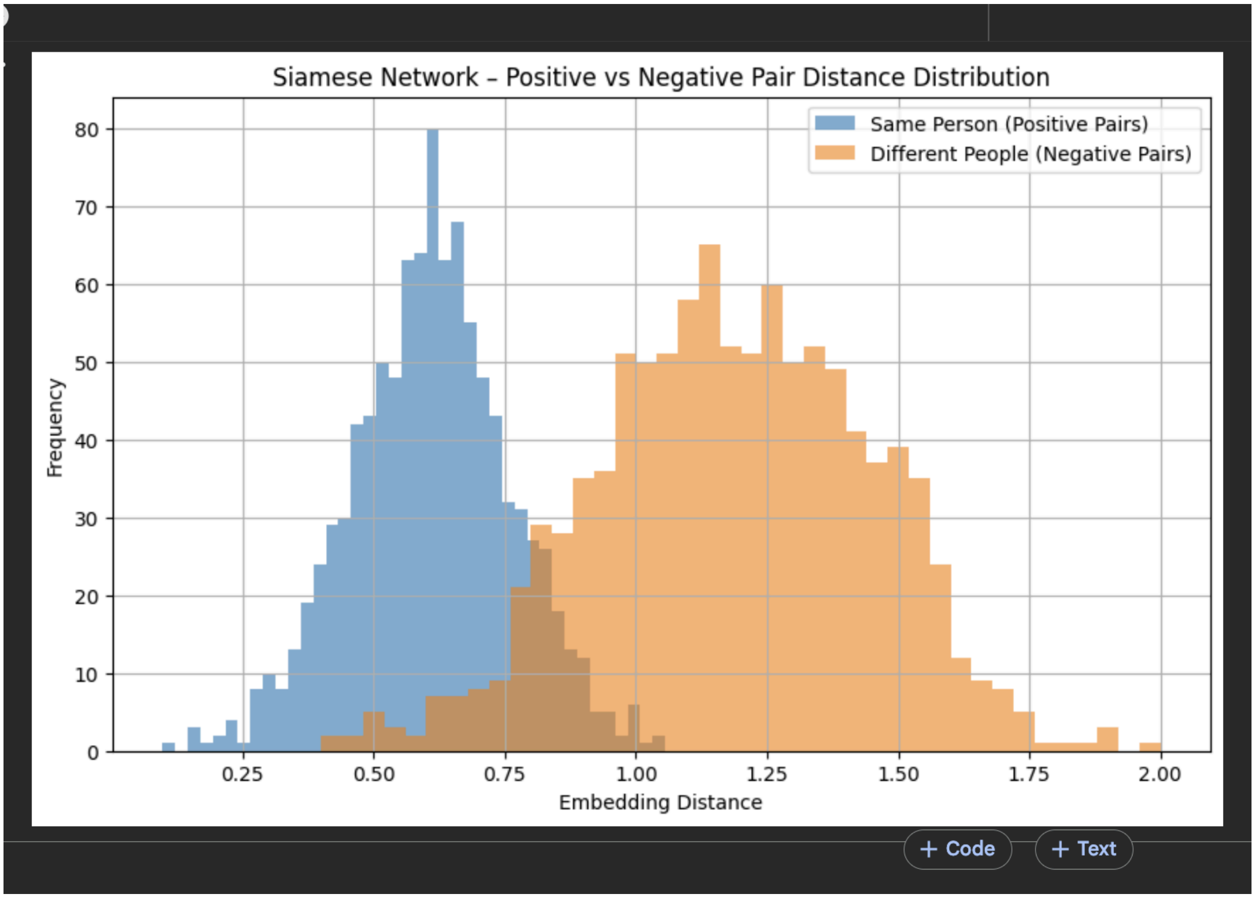
5.2 Confusion Matrix

Confusion matrix for SVM classifier trained on deep embeddings. Darker diagonal blocks represent correct predictions, while off-diagonal regions show misclassifications.



5.3 Siamese Pair Distribution

Positive and negative pairs have considerable overlap, which explains the modest verification accuracy.



These visualizations reinforce the strengths and weaknesses described in both the experimental results and research literature.

6. Key Insights & Reflection

- Recognition based on embeddings is powerful even without training a CNN
- Class imbalance has a significant effect on the results.
- Attention tasks are harder compared to classification.
- Pre-trained generic CNNs cannot compete with models specially designed for face recognition.
- Practical performance reflects real-world challenges, just as discussed in the research paper.

7. Limitations

- Embeddings not specifically trained for faces
- Siamese model trained on embeddings rather than full images.
- No adversarial robustness or fairness evaluation
- LFW has limited samples for some identities.

These limitations are consistent with the broader challenges in face recognition that I documented in my research paper.

8. Conclusion

This mini-project successfully demonstrates the concepts explored in my research paper by implementing a full face-recognition pipeline: embedding extraction, classification, and verification.

Results confirm that

- Deep learning embeddings provide strong baseline performance.
- Classical ML models are effective on high-dimensional deep features.
- Verification accuracy is highly dependent on the quality of embeddings.
- Many limitations to accuracy correspond with the challenges described in current research.

This mini-project provided practical experience in deep-learning-based face recognition and helped to understand real-world complexities while building secure, reliable biometric systems.