

White Paper

Person Face Classification Using AuraFace Model and SVM

An AI-powered face recognition system that learns facial characteristics from training images to accurately recognize individuals.

Date: 22/07/2026

Abdul Basith P M

AI / ML

N-KORR Technologies Pvt Ltd.

1 Executive Summary

The customer requires a solution capable of identifying and recognising individuals from images using a custom dataset maintained within the organization. The solution is intended to support scenarios where an individual's identity needs to be determined from captured images without relying on manually labelled or pre-defined image metadata.

To address this requirement, a Proof of Concept (POC) was developed to evaluate the feasibility of building a person face recognition solution using deep learning-based face embeddings and machine learning classification techniques. The POC uses AuraFace to generate facial embeddings from images and an SVM classifier to identify individuals based on those embeddings.

The primary objective of the POC is to validate whether the proposed approach can accurately recognise individuals from the organization's custom dataset and provide a foundation for further development of a production-ready solution.

For production deployment, the solution would require appropriate privacy and compliance controls, including consent where required, data retention and deletion policies, secure storage and access controls, and assessment of applicable regulatory and organizational requirements. These considerations are outside the scope of the current POC but must be addressed for production readiness.

2 Problem in Context

Developing a face classification model with the person dataset has some technical challenges. Unlike general image classification, facial recognition requires the model to learn subtle differences between facial structures.

The person dataset available here contains a limited number of images per person, especially when compared to datasets on Kaggle, making direct fine-tuning prone to overfitting. Adding to the challenge, the image quality is often poor, and many photos feature groups with multiple people rather than isolated, clear portraits.

3 Technical Deep Dive

The face recognition system was evaluated using two fundamentally different approaches: transfer learning with ResNet34 and an embedding-based recognition pipeline using AuraFace and SVM. Although both methods achieved high training accuracy, their behavior during real-world testing differed significantly.

Preprocessing

Standard preprocessing techniques, including resizing, normalization, rotation, and data augmentation, were applied before training.

Reference Image



After Preprocessing



The images shown below have been preprocessed using standard transformations, including resizing, random rotation, random cropping, normalization, and other data augmentation techniques.

Fine-Tuned ResNet34

A pretrained ResNet34 model was fine-tuned by replacing the final classification layer to match the number of persons in the dataset.

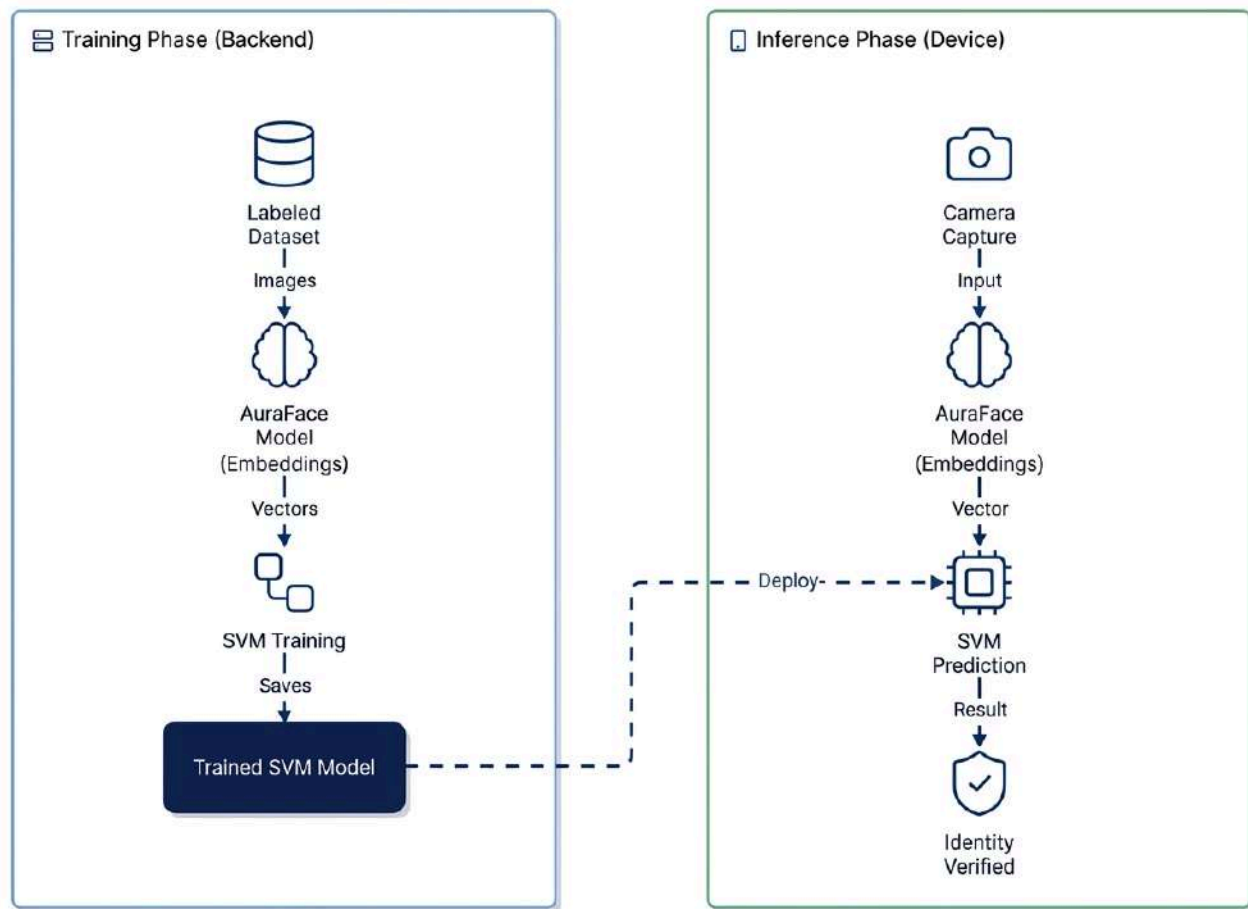
The model achieved over 90% training and validation accuracy. However, testing showed that predictions became unreliable when images contained excessive background information or inconsistent framing. This indicated that the network had learned background context alongside facial features, a common issue when training with relatively small datasets.

Cropping images to include only the face reduced this problem and improved prediction consistency, but it also made the model dependent on consistent preprocessing during inference.

AuraFace + SVM Pipeline

The second approach separates feature extraction from classification. AuraFace generates high-dimensional facial embeddings that capture distinctive facial characteristics while remaining less sensitive to background variations. These embeddings are then used to train an SVM classifier.

Architecture



This architecture offers several advantages:

- Robust feature extraction using a pretrained model
- Lightweight and fast classifier training
- Simple retraining by updating embeddings and the SVM only
- Better generalization on limited datasets

Tradeoffs

Aspect	Fine-Tuned ResNet34	AuraFace + SVM
Training	End-to-end fine-tuning	Fixed embeddings + SVM
Background Sensitivity	Higher	Lower
Retraining	Neural network retraining	SVM retraining only
Dataset Requirement	Larger dataset preferred	Good for smaller datasets
Real-world Robustness	Moderate	High

4 Model Results

Image	Model	Confidence	Correctness
	Fine-Tuned ResNet34	Hari Krishnan B S (99.96 %)	✓
	AuraFace + SVM	Hari Krishnan B S (63 %)	✓
	Fine-Tuned ResNet34	NAJIYA P P K (25.01 %)	✗
	AuraFace + SVM	Hari Krishnan B S (68 %)	✓
	Fine-Tuned ResNet34	Hari Krishnan B S (99.93 %)	✓
	AuraFace + SVM	Hari Krishnan B S (76 %)	✓

Note: The confidence scores represent the model's estimated probability for the predicted class and should not be interpreted as a direct measure of model accuracy. Although the fine-tuned ResNet34 produces near-100% confidence scores, it also generates incorrect predictions with high confidence, indicating overconfident behaviour. AuraFace + SVM produces more conservative confidence scores while achieving more reliable recognition on the evaluated samples.

5 NKORR's Approach

At NKORR, we approach face recognition as a feature representation problem rather than solely a classification task. Instead of retraining deep neural networks for every deployment, we leverage pretrained embedding models to generate robust facial representations and train lightweight classifiers for identity recognition.

Our approach is guided by four principles:

- Leverage pretrained knowledge to reduce training time and improve generalization.
- Separate feature extraction from classification to simplify updates and maintenance.
- Design for scalability, allowing new persons to be added with minimal retraining.
- Validate under real-world conditions, focusing on reliability across varying backgrounds, lighting conditions, and image quality rather than benchmark accuracy alone.

This modular design results in a solution that is easier to maintain, scalable as datasets grow, and better suited for production environments.

6 Recommendations

1. Evaluate models using real-world images rather than relying solely on validation accuracy.
2. Use embedding-based approaches when working with limited datasets or when identity updates occur frequently.
3. Apply consistent face detection, alignment, and cropping during both training and inference.
4. Separate feature extraction from classification to reduce retraining effort and improve maintainability.
5. Use confidence thresholds and periodic model evaluation to maintain reliable recognition performance in production.

7 References

#	Source	Notes
AuraFace Model - https://huggingface.co/fal/AuraFace-v1	<i>HuggingFace</i>	<i>Models in HuggingFace</i>
Practical Deep Learning for Coders - https://course.fast.ai/	<i>Jeremy Howard</i>	<i>Fastai</i>
Introduction to Pytorch - https://docs.pytorch.org/docs/2.13/index.html	<i>Pytorch</i>	<i>Learn Pytorch</i>