One-Shot Facial Recognition using Siamese Neural Networks

Idris Nemsia | Computer Science | Advised by Dr. Sofia Visa

Project Overview

This project focuses on training and evaluating a **Siamese Neural Network (SNN)** for **one-shot facial recognition**. The model is tested on data from unseen classes, evaluating its generalization capabilities in classifying from just a single example.

Problem: Making predictions using limited data during training

- CNNs are powerful for image classification due to their ability to extract hierarchical features through layered transformations.
- **Weakness**: they require **many labeled examples per class** to perform well.
- Collecting such data is often impractical in dynamic or low-resource settings.
- **One-shot learning** addresses this: classifies using only one **example per class**.
- **Compares similarities** between inputs rather than learning a set of fixed classes.

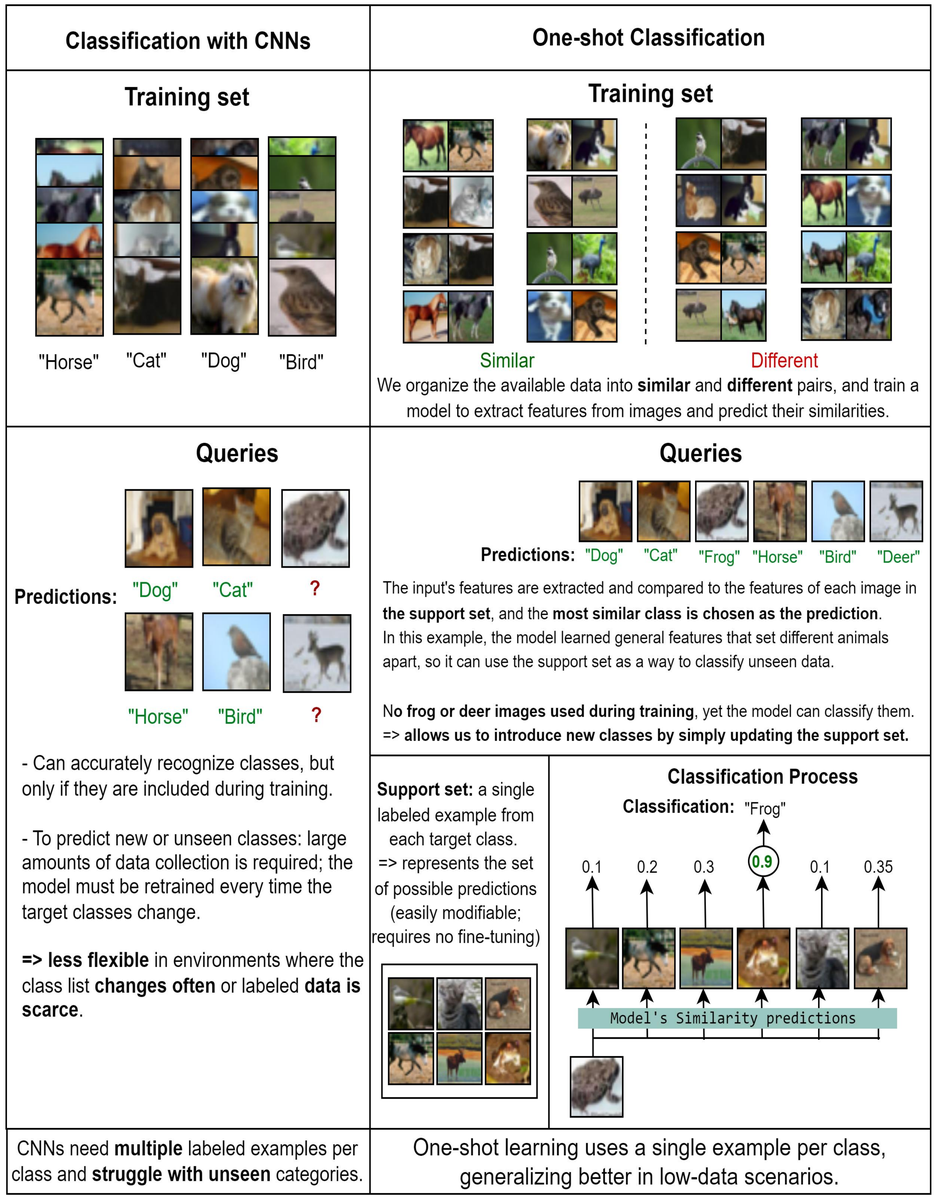

Figure 1: Making predictions with supervised learning vs one-shot learning

Siamese Neural Networks (SNN)

- consists of **two identical feature extraction subnets** (convolutional NN) that share the same weights. Each subnet processes one pair-mage and outputs feature embeddings.
- The **absolute distance** between the two embedding vectors is the **output (1 or 0 for same or different images; Fig.3)**.

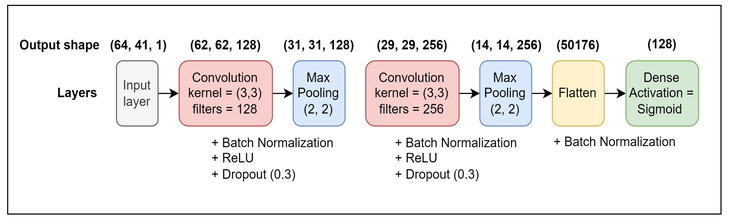

Figure 2. Summary of feature extraction layers

Training

AT&T Lab. Cambridge Olivetti Faces Dataset; images of **40 individuals**, with **10 images** per person.
Training: created 8**0,000 pairs from 30 individuals (**evenly split **similar/different)**.

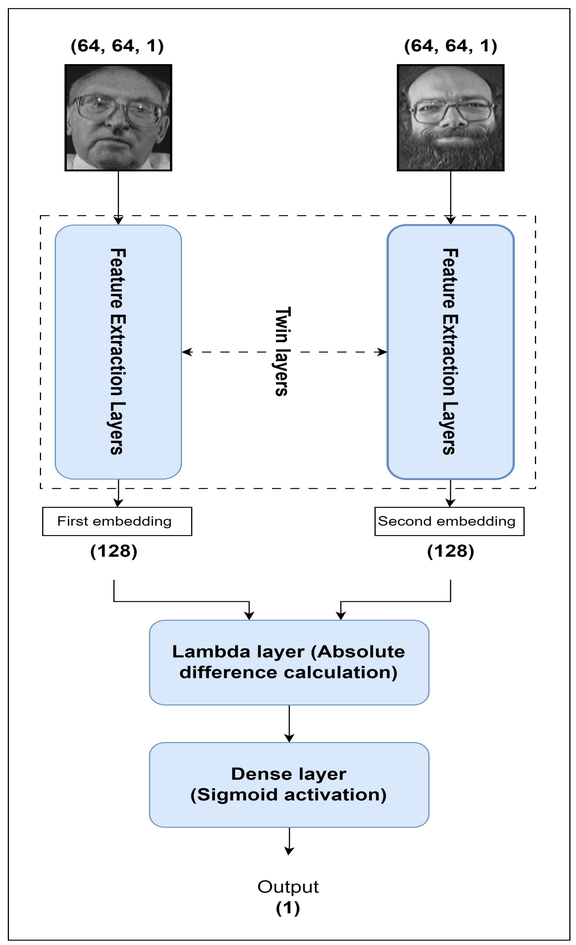

Figure 3. Summary of the Siamese Neural Network structure

Evaluation: One-Shot Trials

One-shot trials are used to evaluate the model's ability to classify unseen classes using a single example per class. In each trial:

1. A **support set** is formed from randomly selected images, each representing a unique class.
2. A **query image** (not in the support set) is classified.
3. Accuracy calculation: **proportion of correct predictions** across trials.

First Evaluation: Unseen Identities from the Original Dataset

- **10,000** 10-way one-shot classification trials are performed using the Olivetti Faces dataset
- The model achieves an accuracy of **68.07%**.

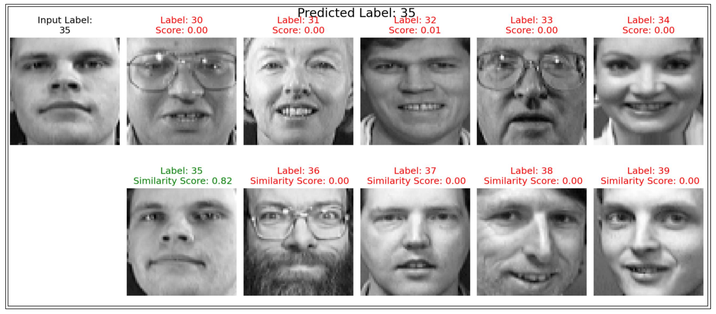

Figure 4: Example one-shot trial on the Olivetti dataset (original dataset)

Second Evaluation: Unseen Identities from an Unseen Dataset

- The model was further tested on the **Yale Face Database** (unseen during training), with **no fine-tuning**.
- **10,000** 10-way one-shot trials using **10 randomly selected individuals**, yielding **44.27% accuracy**. The drop reflects the training data's limitations.

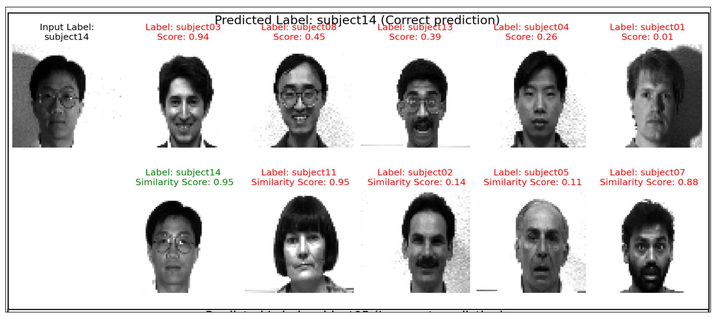

Figure 5: Example one-shot trial on the Yale Face Database (Unseen dataset)

Limitations and Future Work

- The training dataset lacked variation in **lighting, pose, and angle.**
- While uniformity in the training data supported faster learning, it **limited generalization** to new conditions. The model struggled with variations that were not represented during training, like strong lighting in the Yale dataset.
- Future work: use **more diverse training dataset**; improve **data augmentation, loss function** refinement, and **optimization** strategies.

References:
1. Koch G. et al. Siamese Neural Networks for One-shot Image Recognition. Department of Computer Science, University of Toronto. 2015
2. Loy J. "Neural Network Projects with Python". In: Packt Publishing, Feb. 2019.
3. Sewak M. et al. Practical Convolutional Neural Networks. Packt Publishing, Feb. 2018
4. Géron A. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd Edition". In: O'Reilly Media, Inc., Oct. 2022