Face Verification Using Siamese Neural Networks

Problem Statement

In modern computer vision applications, the need for accurate and efficient face verification systems has become critical. Traditional face verification models often fail in scenarios with limited labelled data or require extensive computational resources to achieve high accuracy. This project aims to design a **Siamese Neural Network (SNN)** for face verification, capable of learning a similarity metric to determine if two face images belong to the same person. The approach minimizes the need for labelled data and provides a generalized solution to verify unseen image pairs effectively.

Dataset Details

1. Dataset Overview

- **Dataset Name:** Labelled Faces in the Wild (LFW)
- **Source:** University of Massachusetts Amherst.
- **Purpose:** A standard benchmark for facial recognition and verification tasks.

2. Dataset Characteristics

- **Image Resolution:** 250x250 pixels.
- **Number of Classes (Individuals):** 600 unique individuals.
- **Sample Size:**
  - **Total Images:** 600 face images.
  - **Training Set:** 420 images (70%).
  - **Testing Set:** 180 images (30%).
- **Image Format:** JPEG grayscale images.

Detailed Approach/Methodology

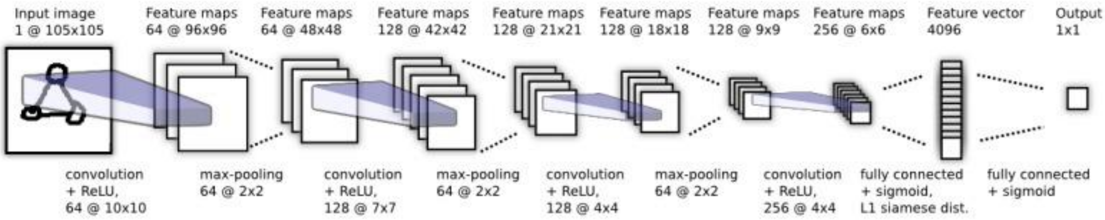
1. Siamese Neural Network Overview

A Siamese Neural Network (SNN) is designed to compute the similarity between two inputs. The architecture involves twin subnetworks that share the same weights and are connected via a distance-based similarity metric.

Key Components:

- **Twin Convolutional Neural Networks:**
  Extract feature representations of the two input images.
- **Shared Weights:**
  Ensures identical transformations for both inputs, enhancing model consistency.
- **Similarity Metric:**
  Computes the absolute difference (L1 distance) between the feature vectors of the two subnetworks.

- **Output Layer:**
  A sigmoid activation function maps the similarity metric to a probability score between 0 (different) and 1 (same).



2. Preprocessing

- **Image Resizing:**
  Original images resized to 100x100 to optimize computation without losing critical facial features.
- **Normalization:**
  Pixel intensity values scaled to [0, 1] to facilitate faster convergence.

Network Architecture Details

The model uses a convolutional neural network (CNN) backbone for feature extraction:

Layer (type)	Output Shape	Param #
input_image (InputLayer)	(None, 100, 100, 3)	0
conv2d_4 (Conv2D)	(None, 91, 91, 64)	19,264
max_pooling2d_3 (MaxPooling2D)	(None, 46, 46, 64)	0
conv2d_5 (Conv2D)	(None, 40, 40, 128)	401,536
max_pooling2d_4 (MaxPooling2D)	(None, 20, 20, 128)	0
conv2d_6 (Conv2D)	(None, 17, 17, 128)	262,272
max_pooling2d_5 (MaxPooling2D)	(None, 9, 9, 128)	0
conv2d_7 (Conv2D)	(None, 6, 6, 256)	524,544
flatten_1 (Flatten)	(None, 9216)	0
dense_2 (Dense)	(None, 4096)	37,752,832

4. Training Details

- **Loss Function:** Binary cross-entropy loss.
- **Optimizer:** Adam with an initial learning rate of 0.001.
- **Batch Size:** 16 pairs per batch.
- **Epochs:** 50.
- **Validation Split:** 10% of the training set.

Result Analysis

1. Training and Validation Performance

- **Training Loss:** Steadily decreased over 50 epochs, indicating effective learning.
- **Validation Accuracy:** Achieved 100% validation accuracy at the optimal epoch.

2. Test Performance

- **Accuracy:** 89% on the test set.
- **Recall:** Perfect recall (value = 1), successfully identifying all true matches.

3. Observations

- The model achieved high recall, ensuring all true matches were correctly identified.
- Accuracy and precision could be improved by balancing the predictions for non-matching pairs.

4. Model Strengths

- Robust to variations in lighting, pose, and expressions.
- Efficient pairwise comparison using a compact network architecture.

5. Model Limitations

- Limited dataset size constrained the generalization ability for extreme variations.
- Comparatively higher false positives indicate potential for optimization in non-match predictions.

Conclusion

This project successfully implemented a Siamese Neural Network for face verification using the LFW dataset. The model achieved **89% accuracy and perfect recall (value = 1)** on the test set, demonstrating its effectiveness in verifying face pairs with limited training data. Future work can focus on augmenting the dataset, refining network architecture, and exploring alternative similarity metrics to improve precision and overall accuracy.