Siamese Networks for Image Similarity

Learning to Compare, Not Just Classify

Team Members

Harsha Bharadwaj	1BG23CS047
Guru Raj Patil	1BG23CS044
Chinmaya V Kavali	1BG23CS032

The Problem with Standard Classification

Standard classifiers answer: "What is this image?"

They struggle with too many classes or limited data.

New, unseen classes require complete retraining.

We need a different approach for verification tasks.

The question becomes: "Are these two images the same?"

The Solution: A Siamese Network

A Siamese Network uses two identical "twin" neural networks.

They share the exact same weights for consistent processing.

1

Two Inputs

Takes a pair of images (Image A, Image B).

2

Shared Network

Both images are fed through the same base network.

3

Embedding Space

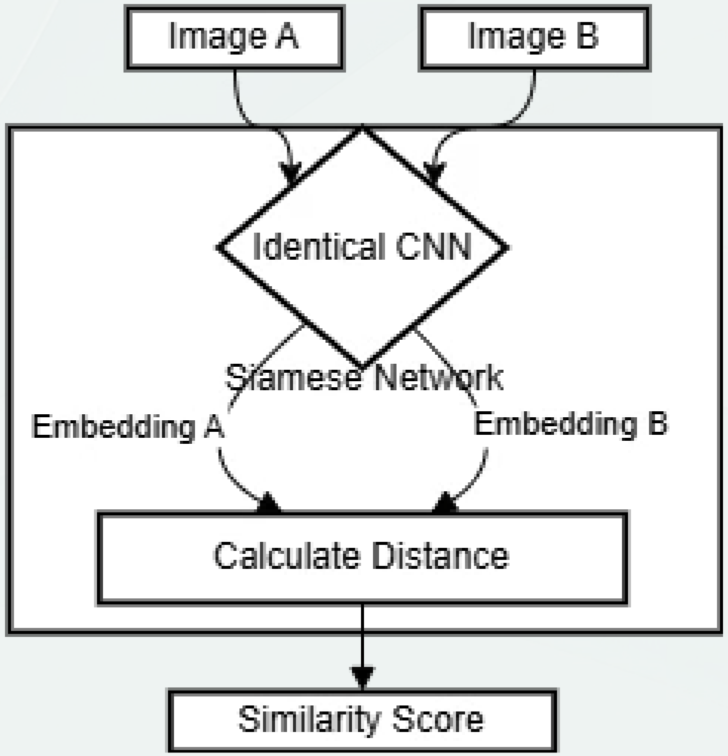
Generates a feature vector for each image.

4

Compare

Calculates the distance between the embeddings.

Goal: Learn an embedding space where similar images are close together.



Siamese Network Flowchart

Both images are processed by the same convolutional network (CNN) to create embeddings. The distance between these embeddings gives us a similarity score, indicating how alike the images are.

Siamese vs. Traditional Networks

Feature	Siamese Network	Traditional Classifier (CNN)
Primary Goal	Measures Similarity / Verification	Assigns a Class Label
Output	A distance or similarity score	Probabilities for each class
Training Data	Pairs of data (similar/dissimilar)	Single, labeled data points
Handling New Classes	Excellent. No retraining needed.	Poor. Requires full retraining.
Best Use Case	Face recognition, signature verification	General object classification
Core Question	"Are these the same?"	"What is this?"

How It Learns: Contrastive Loss

The network learns by being shown pairs and getting feedback on the distance between their embeddings.

Positive Pair (Same)

Embeddings are "pulled" closer together.

Low loss when distance is minimal.

Negative Pair (Different)

Embeddings are "pushed" further apart.

Low loss when distance exceeds a set margin.

This "push-and-pull" technique is guided by the Contrastive Loss function, optimizing the embedding space.

Our Project: Results

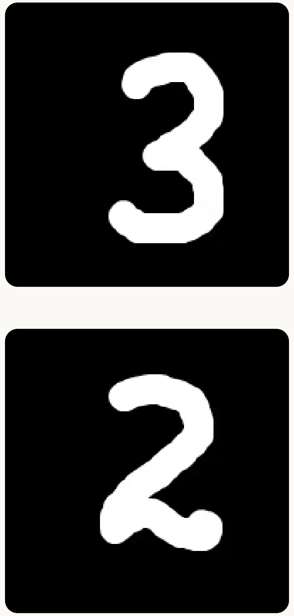
We applied this concept to the MNIST dataset of handwritten digits.

Data: We created pairs of MNIST images.

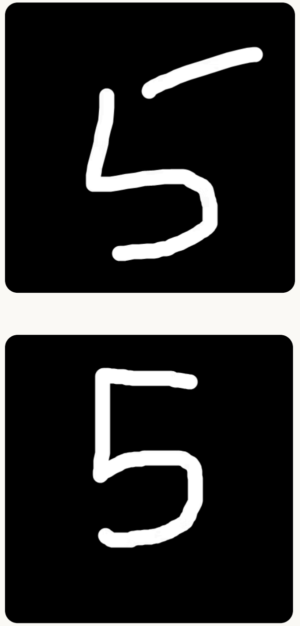
- Positive: (5, 5), (2, 2)
- Negative: (7, 1), (3, 8)

Model: A Siamese Network with a simple CNN base.

Training: Used Contrastive Loss to learn embedding distance.



Not Similar



Similar

Result: A Streamlit app. Users upload two digits. The model outputs a distance score. A low score indicates the digits are the same, demonstrating effective similarity learning.

Key Takeaways & Applications

What We Learned

- Siamese Networks are powerful for verification.
- Excel at One-Shot Learning with few examples.
- Key: learning meaningful embedding spaces.

Real-World Applications

- Face Recognition: Person identity verification.
- Signature Verification: Authenticating documents.
- Duplicate Detection: Finding similar content.
- Recommendation Engines: Visual product similarity.