Problem Statement

Develop a fashion recommendation system that can suggest similar items based on a query image.

Goal: Assist small fashion businesses with personalized recommendation features.

Given: A query image by the user.

Objective: Learn a model that recommends compatible outfits.

Constraints: Limited labeled data, No textual or semantic metadata used, computational power for training models.

Approach/Technology Used:

- Siamese Neural Network with color histogram fusion
- CNN from scratch (I/P: (3,128,128) #Layers: 3)
- Cosine similarity for similar item generation

Motivation: Bridging the gap in fashion e-commerce by offering ML-driven stylistic recommendations.

Application: Fashion retail platforms, boutique businesses, styling assistants, face detection.

Dataset Description

Dataset: Repolyvore Dataset

Train: Repolyvore Fashion Dataset (dress, shoes, pants) [25k images]

Test: Query folder[10 images]

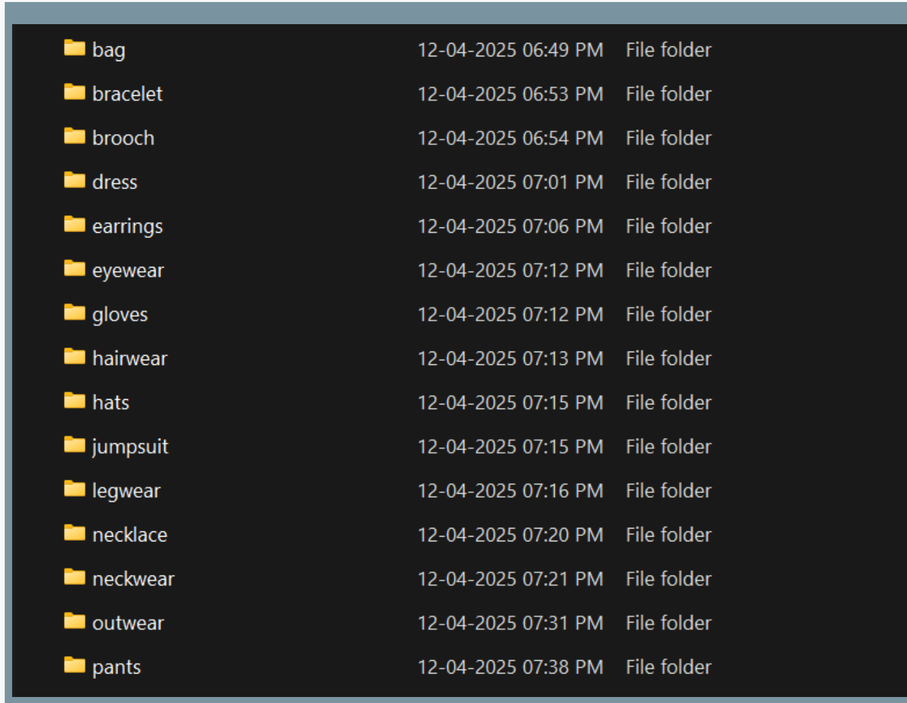
Data Type: Images (.jpg), RGB images

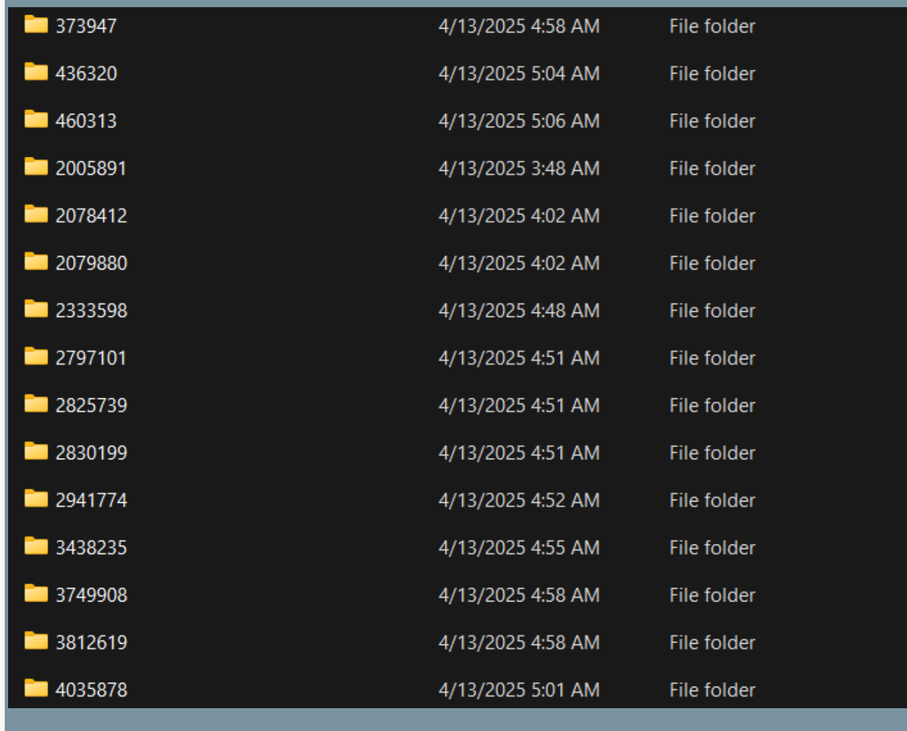
Description: The dataset includes various curated outfits which includes serveral items which look compatible together.

Machine Specification:

- CPU: AMD Ryzen 5 5600H
- RAM: 16 GB
- GPU: NVIDIA RTX 3050 (used for training)
- Programming Language: Python (PyTorch-CUDA, TorchVision)
- Platform: Jupyter Notebook

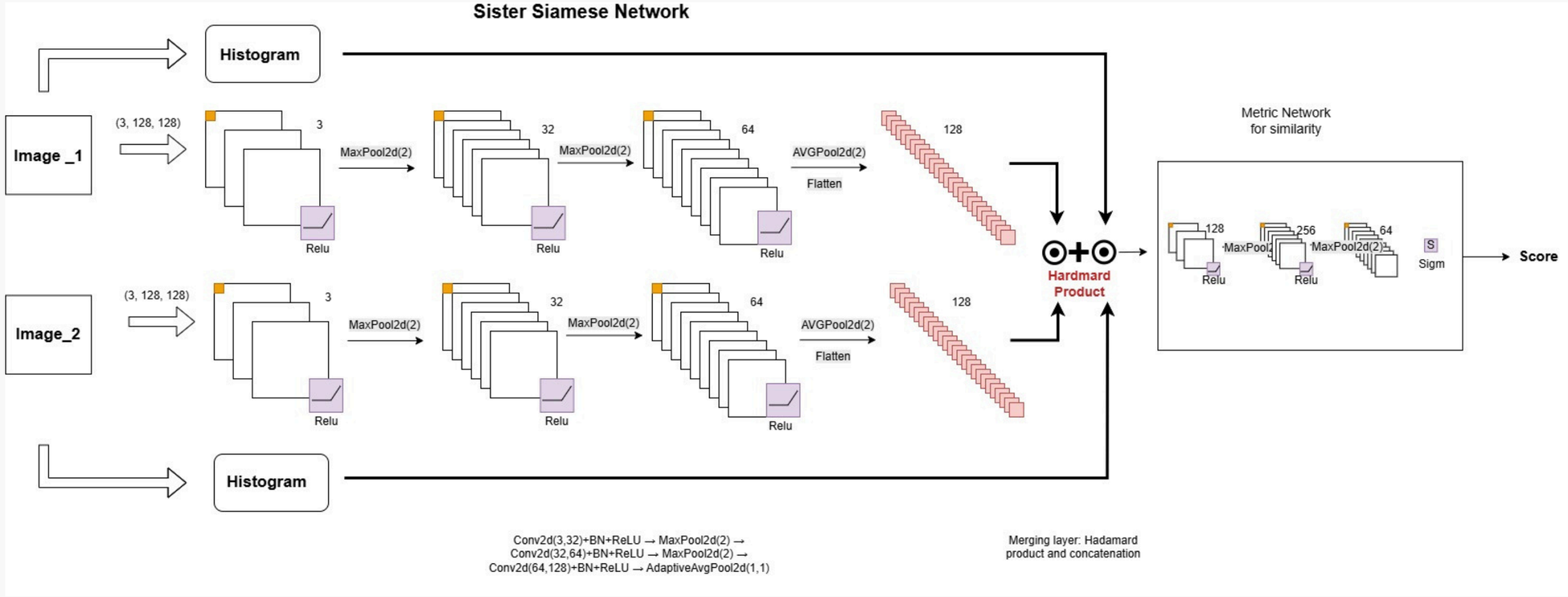
1. archive/images/<compatible-set-id>/<class-id>.jpg
2. Re-Polyvore/<class-name>/<compatible-set-id>_<class-id>.jpg





Neural Network Architecture

Model: Custom CNN with siamese framework



Training Process:

1. Siamese sister network (for feature extraction)
   - [Input: 3 channels of 128x128 Pair of like and unlike images] ,
   - [Output: Vector embedding of lenght 128]
2. Metric Network (for similarity training and score)
   - Hadmard Product of feature embeddings of images and histogram
   - Concatenated feature embeddings of images and histogram

Optimiser: *Adam Opt. (Learn Rate= 1e-3)*

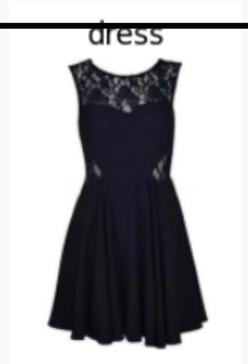
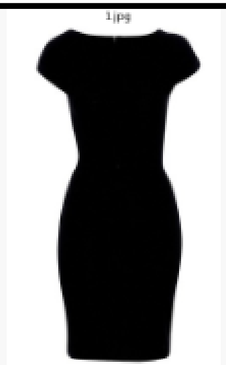
Batch size: *32*

Epochs: *10*

Loss function: Binary Cross Entropy

$$L = -\left[y \cdot \log(\hat{y}) + (1-y) \cdot \log(1-\hat{y})\right]$$

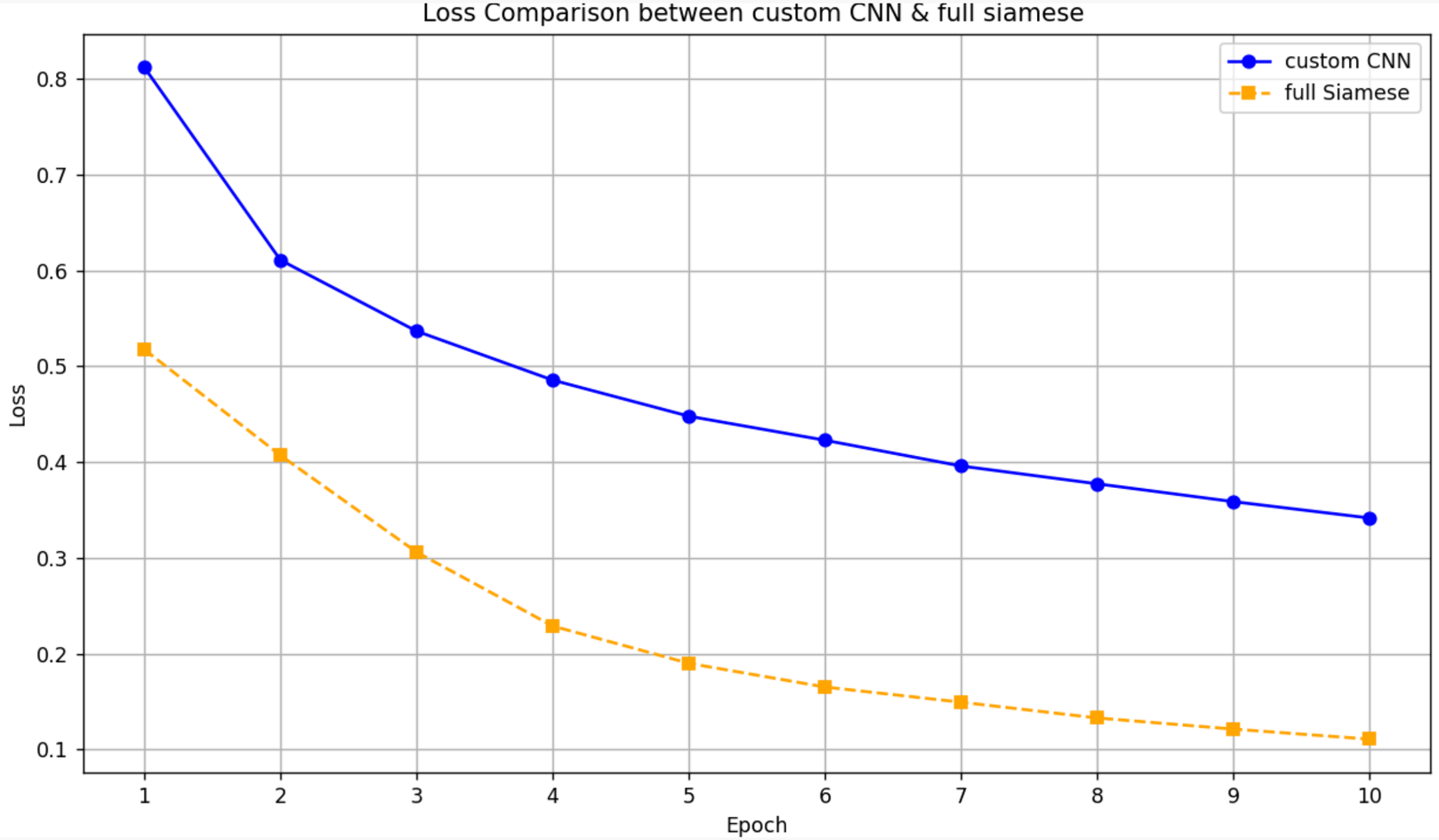
Result- Qualitative- Similarity Search

	Image 1	Image 2	Image 3	Image 4	Image 5
Naive CNN					
Siamese & Color Hist.					

Query Image

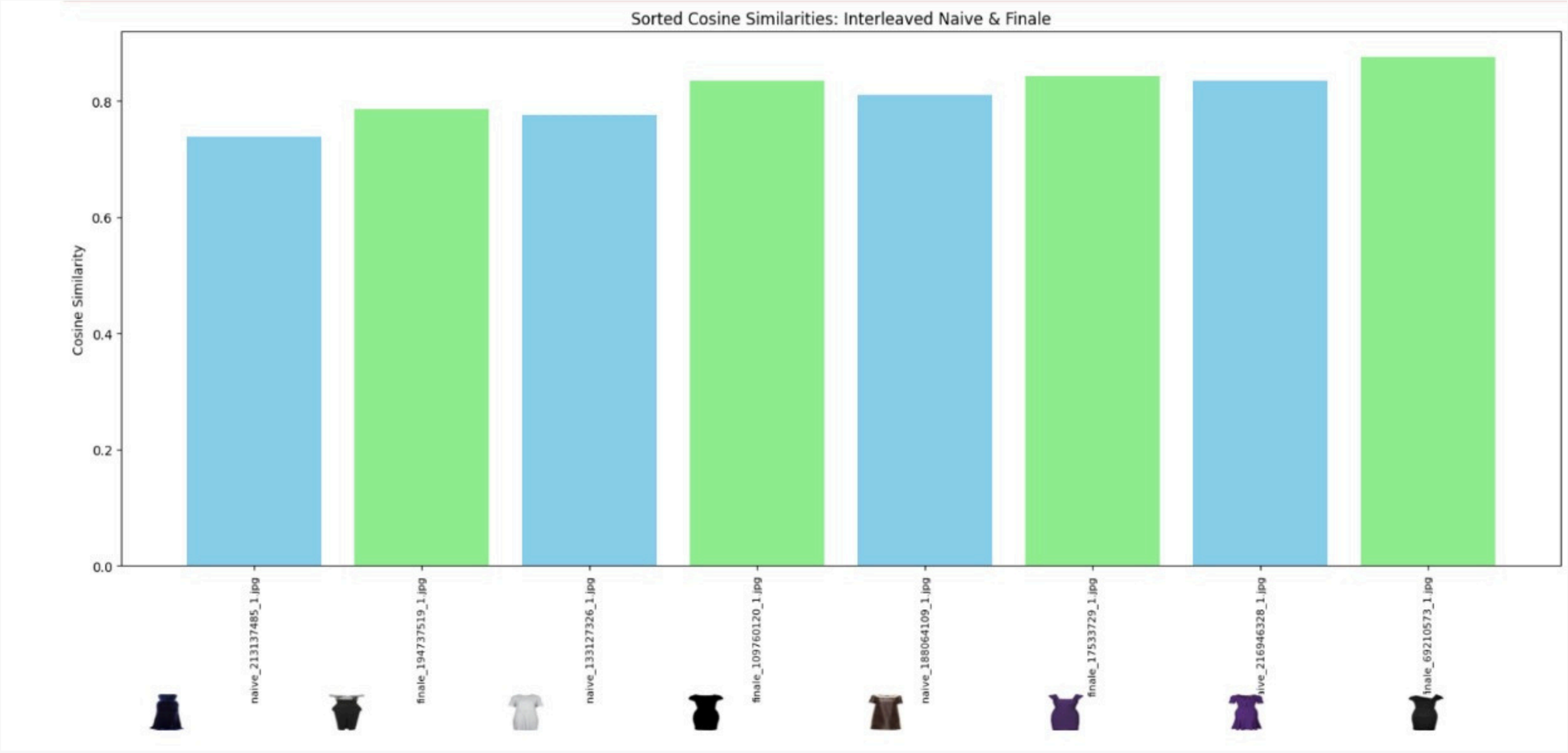
Quantitative Results

Loss Comparison



Quantitative Results

Cosine Similarity Comparison



References

Polyvore Dataset: *https://www.kaggle.com/datasets/dnepozitek/maryland-polyvore-images/*

Citation of Paper: *LEARNING FASHION COMPATIBILITY ACROSS APPAREL CATEGORIES FOR OUTFIT RECOMMENDATION Luisa F. Polan´ıa, Satyajit Gupte: [2019 IEEE International Conference on Image Processing (ICIP)]*

Articles on Siamese Networks:

https://medium.com/@rinkinag24/a-comprehensive-guide-to-siamese-neural-networks-3358658c0513

https://builtin.com/machine-learning/siamese-network

Siamese Networks for Face Detection - https://youtu.be/IXgr63eRU5U?si=-3DI1iocaaN3LyrB

Thank you