Assignment #2 — Product Recommendation using CBIR via feature extraction

Rhichard Koh (100842848)

Faculty of Science, Engineering & Information Technology, Durham College

Advanced Neural Networks: Deep Learning for Spatial Analysis (COSC 41001)

Dr. Noopa Jagadeesh

October 15, 2024

Dataset and Preprocessing Steps

The dataset used for this project is the Fashion Product Images dataset, which contains a collection of clothing and fashion accessory images. This dataset provides a diverse range of product types that serve as ideal content for building a recommendation system based on visual similarity. Each image was resized to 224x224 pixels to match the input requirements of the pre-trained ResNet50 model. The images were then converted to tensors and normalized using ImageNet's mean and standard deviation values. These transformations ensure that the input data is compatible with the model, which was originally trained on ImageNet images. The dataset was loaded and preprocessed using Pillow for image manipulation and torchvision transformations to prepare the images.

Model and Techniques Used

The ResNet50 model was chosen as the feature extractor for this project. Specifically, a pre-trained version of ResNet50 was used with its classification layer removed, allowing the model to generate a feature vector that represents each image in a high-dimensional space. This approach captures the essential visual features of each image, such as color, texture, and shape, without explicitly classifying the objects. The cosine similarity metric was used to compute the similarity between the feature vectors of images. Cosine similarity measures the cosine of the angle between two vectors, which makes it an effective measure for identifying visually similar products. The top-N most similar images were retrieved based on this similarity score. In summary, feature extraction was performed using ResNet50 without the final classification layer to obtain 2048-dimensional feature vectors for each image, and cosine similarity was used to measure similarity between these feature vectors.

Evaluation Results

The effectiveness of the recommendation system was evaluated qualitatively by visual inspection. The system successfully identified and recommended images that were similar to the query image based on features such as color, texture, and overall design. Quantitative evaluation metrics such as precision or recall could not be directly applied, as this system focuses on visual similarity, and there is no ground-truth similarity score available. However, user testing demonstrated a high level of satisfaction with the recommendations, indicating that the model was effective in retrieving visually similar items.

Challenges Faced and How They Were Addressed

One of the challenges faced was the dataset size, as the dataset contained a large number of high-resolution images, which led to increased processing time and memory

usage. This was addressed by using GPU acceleration for both the feature extraction process and the model's forward pass. Another challenge was the feature extraction time, as extracting feature vectors from all images was time-consuming. To address this, the feature vectors were saved to a file after being computed. This allowed the system to avoid redundant computations and significantly speed up future queries. Additionally, image quality variability was a challenge, as some images in the dataset were of lower quality or had different lighting conditions. To mitigate the impact of these inconsistencies, the model was pre-trained on ImageNet, making it robust to variations in image quality.

Applications and Domains for CBIR

The content-based image retrieval (CBIR) system developed in this project has several potential applications in other domains. In e-commerce, it can be used to recommend similar products to customers based on visual features, such as clothing, furniture, or accessories, enhancing the online shopping experience. In medical imaging, it can be used to retrieve similar medical images (e.g., X-rays, CT scans) from a database to assist doctors in diagnosing medical conditions by finding past cases with similar patterns. In the field of art and design, the system can assist users in finding artworks or design elements with a similar style or visual motif, which can be useful for art historians, designers, and creators. In digital asset management, the CBIR system can help organizations manage large collections of digital assets by enabling efficient search and retrieval based on visual characteristics, such as images in marketing or branding. Finally, in surveillance, it can be used to identify objects or individuals in surveillance footage by comparing them to existing records or past images, aiding in security and law enforcement applications.