

**Izmir Democracy University**

**Engineering Faculty**

**Department of Electrical Electronics Engineering**

**Teaching Staff:** Başak Esin Köktürk Güzel

**Group Members:** Onur Mert - Muratcan Avcu

**School Numbers:** 1906102030 – 1906102042

**Project Name:** Clustering The Segmented Clothes by Colors

**Github Link:** <https://github.com/muratcanavcu>

**Abstract**

In this project, we present a novel approach to organize a dataset of human photos with clothing items through the integration of image segmentation and K-means clustering techniques. The dataset, comprising 2032 images, undergoes a two-step process to categorize clothing based on color. Initially, image segmentation is employed to isolate and extract the clothing items from each photograph, producing high-quality PNG images of segmented garments. Subsequently, K-means clustering is applied to the numerical values representing the colors within the segmented clothing, resulting in the formation of distinct clusters.

The outcome of this process is a dataset neatly organized into folders, each containing clothing items of similar or identical color. This organization enables efficient exploration, retrieval, and analysis of clothing data based on color attributes. Such a methodological approach contributes to the field of computer vision and image processing, offering a valuable tool for researchers, fashion enthusiasts, and industry professionals alike.

**Introduction**

The rapid growth of image-based datasets, particularly those involving human subjects and their attire, has necessitated innovative methods for efficient organization and analysis. In response to this demand, our project employs a combination of image segmentation and K-means clustering to create a structured dataset focused on color-based clothing categorization.

The initial step involves the application of image segmentation techniques to extract clothing items from human photographs. This results in the generation of high-resolution PNG images, isolating the clothing items for further analysis. The subsequent application of K-means clustering to the numerical representations of color within these segmented clothing items enables the creation of distinct clusters based on color similarity.

By organizing the dataset into folders representing clusters of similar-colored clothing, our approach facilitates a streamlined exploration of fashion trends, color preferences, and style variations. This project has potential applications in various domains, including fashion industry trend analysis, recommendation systems, and virtual wardrobe creation. Researchers and professionals can leverage this method to efficiently navigate and analyze large clothingdatasets,enhancing the understanding of color patterns and trends within diverse fashion collections.

**Literature Review**

The integration of image segmentation and clustering techniques in the context of clothing analysis has garnered attention in recent years. Researchers have explored various methodologies to enhance the organization and analysis of large-scale clothing datasets. Image segmentation serves as a crucial initial step in isolating clothing items from complex scenes, enabling focused analysis on the garments themselves.

Existing literature highlights the diverse applications of clustering algorithms in the realm of fashion and image analysis. K-means clustering, in particular, has been widely employed for its simplicity and effectiveness in grouping similar data points. While previous studies have applied clustering techniques to clothing datasets, the combination of segmentation and clustering for color-based organization is a relatively novel approach.

Several projects have explored the use of clustering algorithms for garment categorization, often emphasizing attributes such as style, pattern, or overall appearance. However, the focus on color as a primary categorization criterion has been less explored. Our project aims to bridge this gap by leveraging the rich color information obtained through image segmentation, offering a unique perspective on clothing organization and analysis.

**Method**

**Data Collection and Preprocessing:**

We collected a dataset consisting of 2032 human photographs featuring individuals wearing various types of clothing. Each image was carefully annotated to ensure accurate segmentation during the subsequent stages. Prior to segmentation, the images were preprocessed to enhance clarity and reduce noise, involving resizing to a standardized resolution and normalization of pixel values.

giyim, kişi, şahıs, günlük elbise, üst, tepe içeren bir resim

Açıklama otomatik olarak oluşturuldu giyim, kişi, şahıs, moda aksesuar, insan yüzü içeren bir resim

Açıklama otomatik olarak oluşturuldu giyim, kişi, şahıs, üst, tepe, günlük elbise içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Image Segmentation using UNet:**

For the segmentation of clothing items within the images, we employed the UNet architecture, a convolutional neural network (CNN) widely recognized for its effectiveness in semantic segmentation tasks. The UNet model was trained on a subset of the dataset with corresponding ground truth segmentation masks. Training involved optimizing the network's parameters using a combination of binary cross-entropy loss and Adam optimization.

kişi, şahıs, giyim, moda aksesuar, günlük elbise içeren bir resim

Açıklama otomatik olarak oluşturuldu giyim, kişi, şahıs, Spor tişörtü, cep, kese, boşluk içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Generation of Segmented Clothing Masks:**

Once the UNet model was trained, it was applied to the entire dataset to generate segmented clothing masks. The resulting masks represented the precise boundaries of clothing items within each image, enabling the isolation of garments for subsequent analysis.

**K-Means Clustering for Color-Based Categorization:**

The numerical values representing the color information within each segmented clothing item were extracted. These color values were then used as input features for K-Means clustering. We opted for K-Means due to its simplicity and efficiency in grouping similar data points. The number of clusters (K) was determined through empirical validation and experimentation to ensure meaningful color groupings.

**Organization of Dataset into Color-Based Clusters:**

The K-Means clustering algorithm produced distinct color-based clusters, each representing a group of similar-colored clothing items. The dataset was then organized into folders, with each folder corresponding to a specific color cluster. This organization facilitates convenient access to clothing items based on their color attributes for subsequent analysis.

This combined approach of UNet segmentation and K-Means clustering presents a robust methodology for the organization and analysis of large clothing datasets based on color attributes. The integration of these techniques provides a foundation for insightful exploration of fashion trends, color preferences, and style variations within the dataset.

ekran görüntüsü, metin, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

ekran görüntüsü, metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

**References**

[1] Cushen, G. A., & Nixon, M. S. (2012, July). Real-time semantic clothing segmentation. In *International Symposium on Visual Computing* (pp. 272-281). Berlin, Heidelberg: Springer Berlin Heidelberg.

[2] Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, *9*(8), 1295.

[3] Patil, R. V., & Jondhale, K. C. (2010, July). Edge based technique to estimate number of clusters in k-means color image segmentation. In *2010 3rd International Conference on Computer Science and Information Technology* (Vol. 2, pp. 117-121). IEEE.