NAME OF PROJECT

Feature-Based Fashion Accessory Clustering

DATE

19/03/2025

FEATURE-BASED FASHION ACCESSORY CLUSTERING

A Traditional Computer Vision Approach

CENTRALSUPELEC

Computer Vision

PRESENTED BY

KARAOGUZ Oguz, PODDAR Karam, RAVELLO Riccardo

01

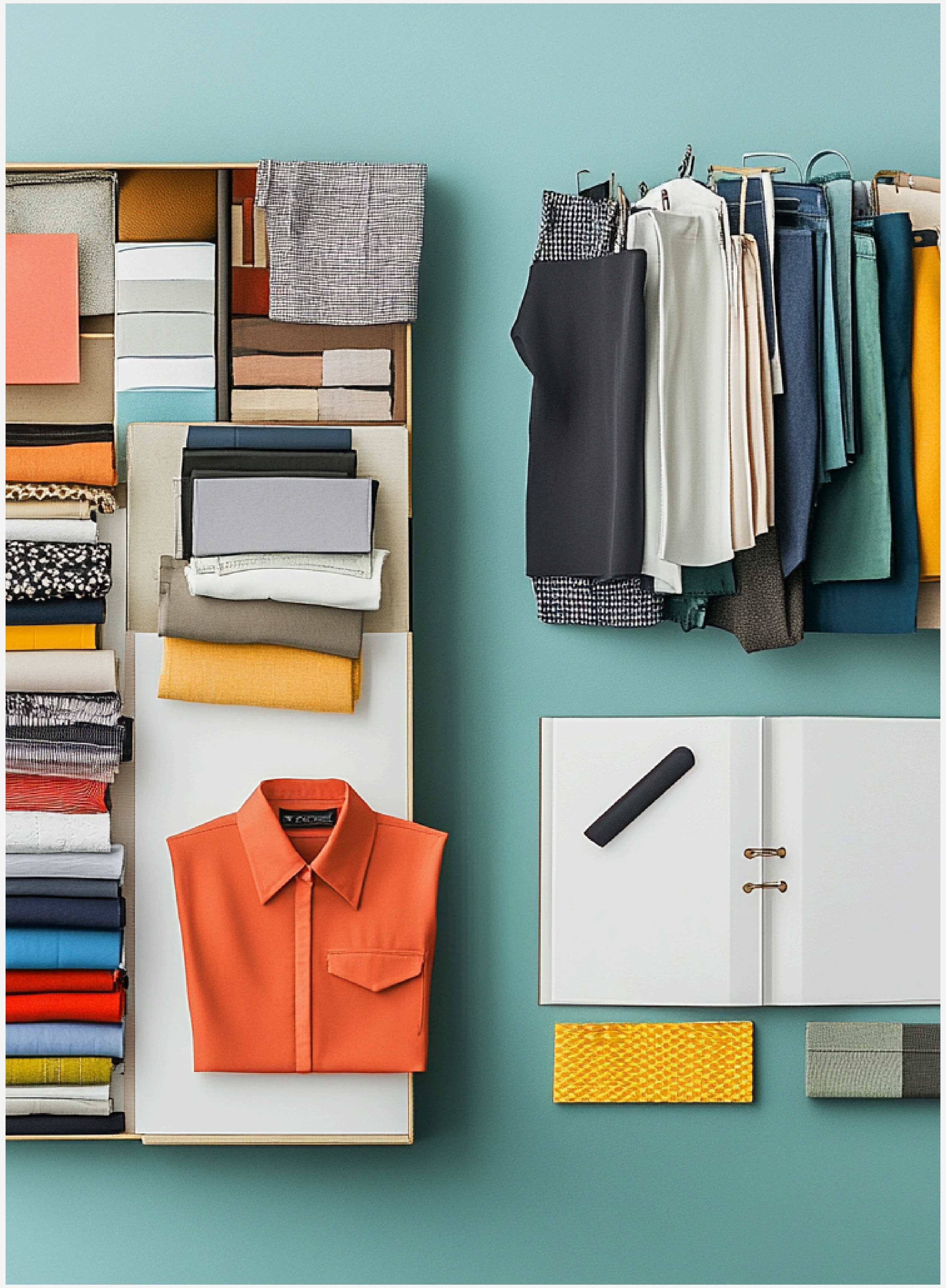
INTRODUCTION & MOTIVATION

Problem:
Organization of fashion accessory images

Challenges:
Manual tagging (labor-intensive), Deep learning (requires training data)

Our approach:
Traditional feature-based clustering

The Dataset
750 men's accessories from Kaggle Fashion Product Images dataset



02

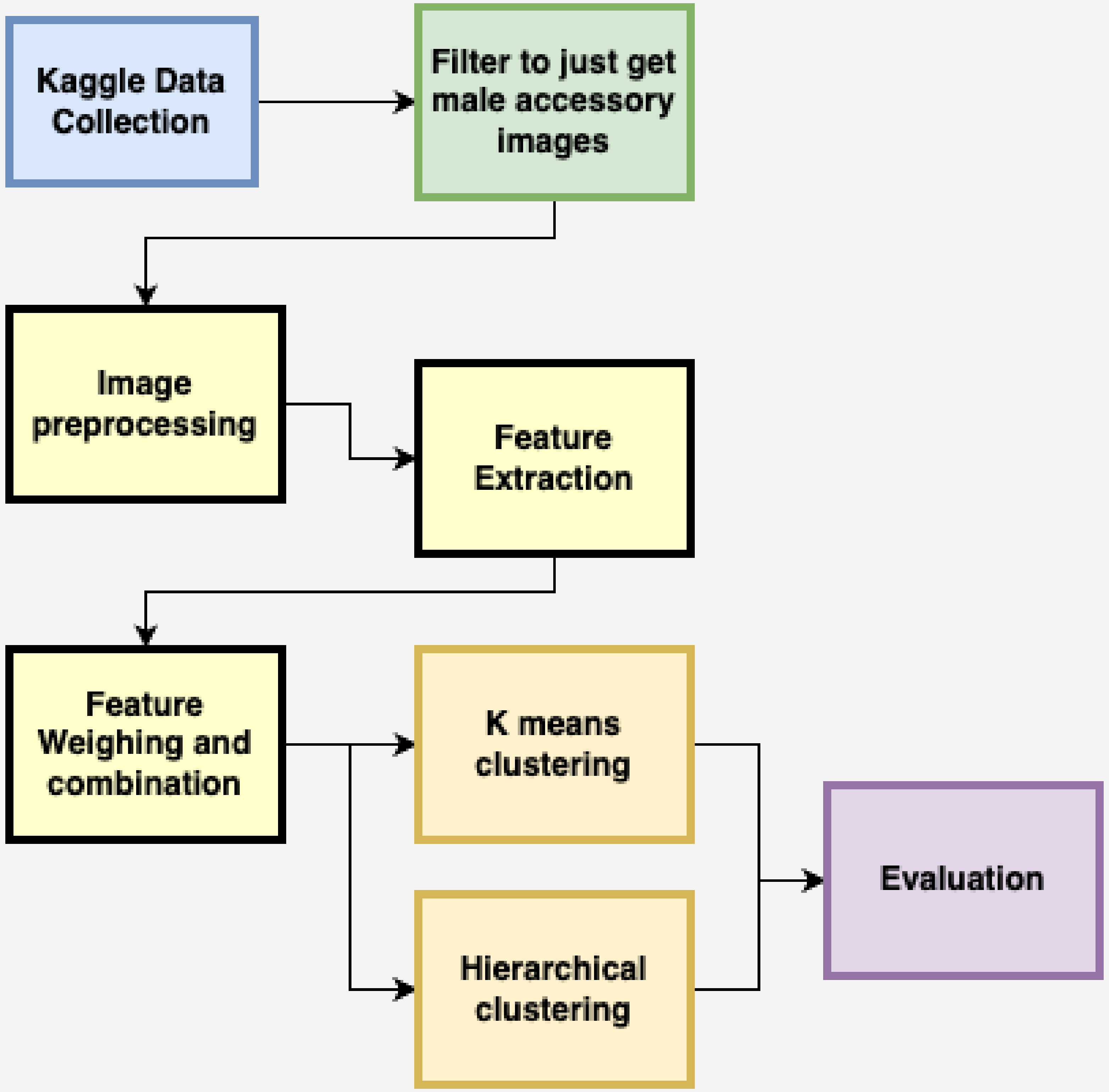
KEY RESEARCH QUESTIONS

- How effectively can traditional feature extraction identify meaningful patterns?
- Can these features discriminate between different accessory categories?
- How does K-means compare to hierarchical clustering for this domain?
- What's the optimal balance between computational efficiency and accuracy?

RELATED WORK

03

- **Feature extraction:** SIFT (Lowe), Harris corner detection (Harris & Stephens)
- **Colour and texture representation:** Duda et al.'s work on pattern classification, histogram-based features
- **Background removal:** Inspired by Rother's "GrabCut" but adapted for automatic processing
- **Fashion image analysis:** Contrasts with DeepFashion (Liu et al.) which relies on deep learning
- **Clustering:** K-means (Jain), Hierarchical clustering approaches
- **Our contribution:** Specialized preprocessing, weighted feature representation, comparative analysis for fashion accessories



04

METHODOLOGY OVERVIEW

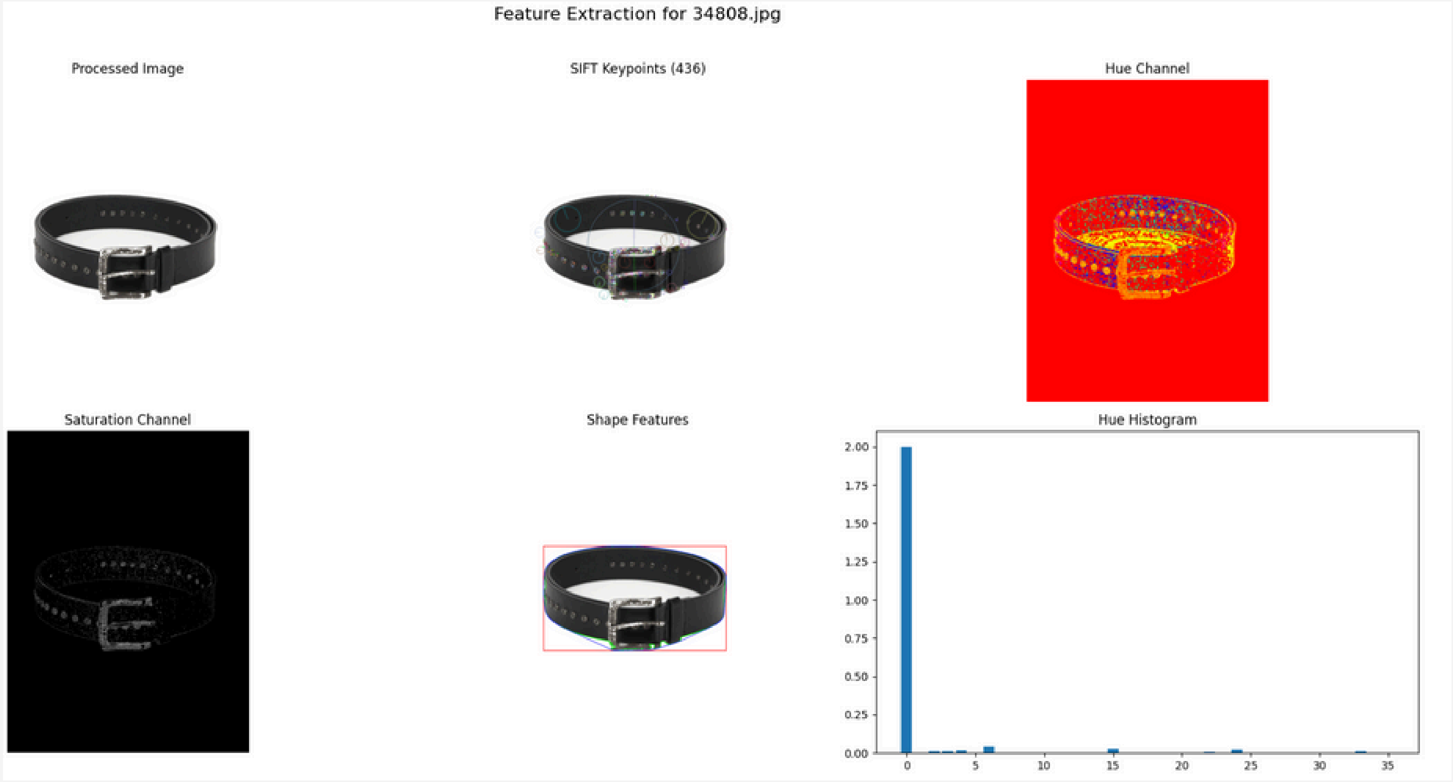
- Data collection (Fashion Product Images)
- Preprocessing
- Feature extraction (SIFT, shape, color, texture)
- Feature weighting and combination
- Clustering (K-means and hierarchical)
- Evaluation

05

PREPROCESSING PIPELINE

Key steps:

- Grayscale
- Thresholding
- Morphological operations
- Contour detection
- Skin detection



06

FEATURE EXTRACTION

SIFT features (weight: 1.0)
- Scale-Invariant keypoints that capture distinctive local patterns invariant to rotation and scale
- Represented as 128-dimensional vectors organized into a Bag of Visual Words (500 dimensions)

Shape features (weight: 6.0)
- Five key metrics: aspect ratio, extent, solidity, equivalent diameter, and convexity
- Captures geometric properties of accessories like elongation and boundary complexity

Color features (weight: 5.0)
- HSV color space histograms with 36 bins for hue, 32 for saturation, 32 for value
- Includes dominant color extraction and spatial color distribution across a 4×4 grid

Texture features (weight: 0.8)
- Histogram of Oriented Gradients (HOG) computed on 8×8 pixel cells
- Captures edge information and surface patterns with 9 orientation bins spanning 180°

07 CLUSTERING APPROACHES

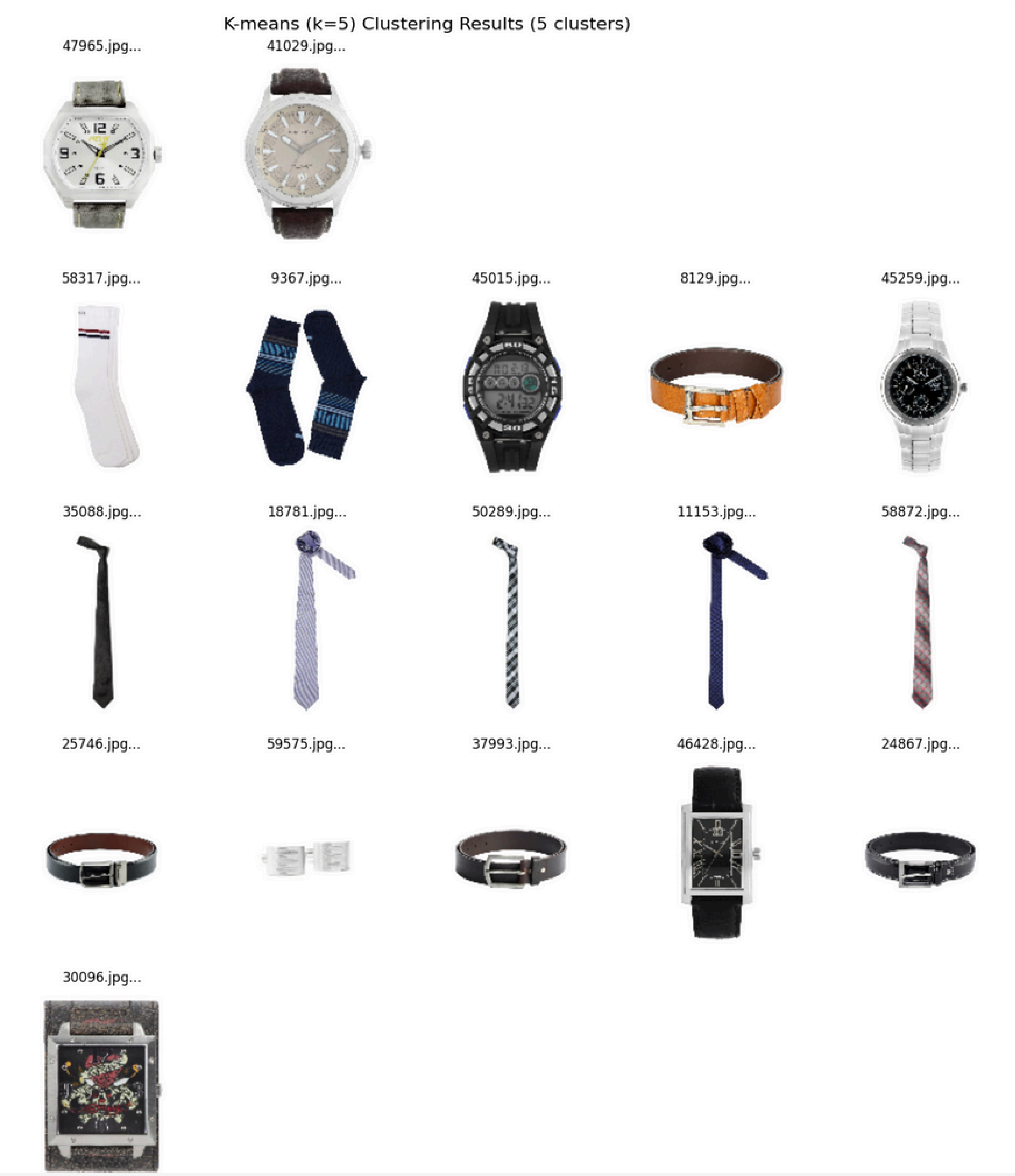
K-means

- Iterative algorithm that partitions data into k clusters by minimizing within-cluster variance
- Sensitive to initialization - we used multiple random starts to avoid local minima
- Requires pre-specifying number of clusters (k)
- Optimal cluster determination using:
  - Silhouette score: measures how similar objects are to their cluster compared to other clusters
  - Elbow method: analyzes distortion (sum of squared distances) drop-off across different k values

Hierarchical:

- Bottom-up approach that starts with each sample as its own cluster and progressively merges
- Creates a dendrogram showing nested clustering structure at different levels
- Does not require pre-specifying number of clusters (can cut dendrogram at different heights)
- Evaluated different linkage criteria:
  - Ward's method: minimizes variance increase after merging
  - Average linkage: uses average distance between all pairs of points in different clusters
  - Complete linkage: uses maximum distance between any points in different clusters

Feature-Based Fashion
Accessory Clustering



08

RESULTS
K-MEANS CLUSTERING

K-means results with k=5 (silhouette score: 0.0298):

- Cluster 1: Watches with light-colored faces - primarily analog watches with white/silver dials
- Cluster 2: Mixed accessories including socks, digital watches, and belts with darker colors
- Cluster 3: Ties with various patterns and colors - showing strong shape-based grouping
- Cluster 4: Belts with different buckle styles and a watch - leather accessories with similar color profiles
- Cluster 5: A distinctive decorative watch with unique visual appearance

Statistical metrics favored k=2 (silhouette score: 0.7943) but resulted in 99.9% of images in a single cluster

Visual coherence demonstrates feature weighting effectiveness despite lower silhouette score

09

RESULTS
HIERARCHICAL CLUSTERING

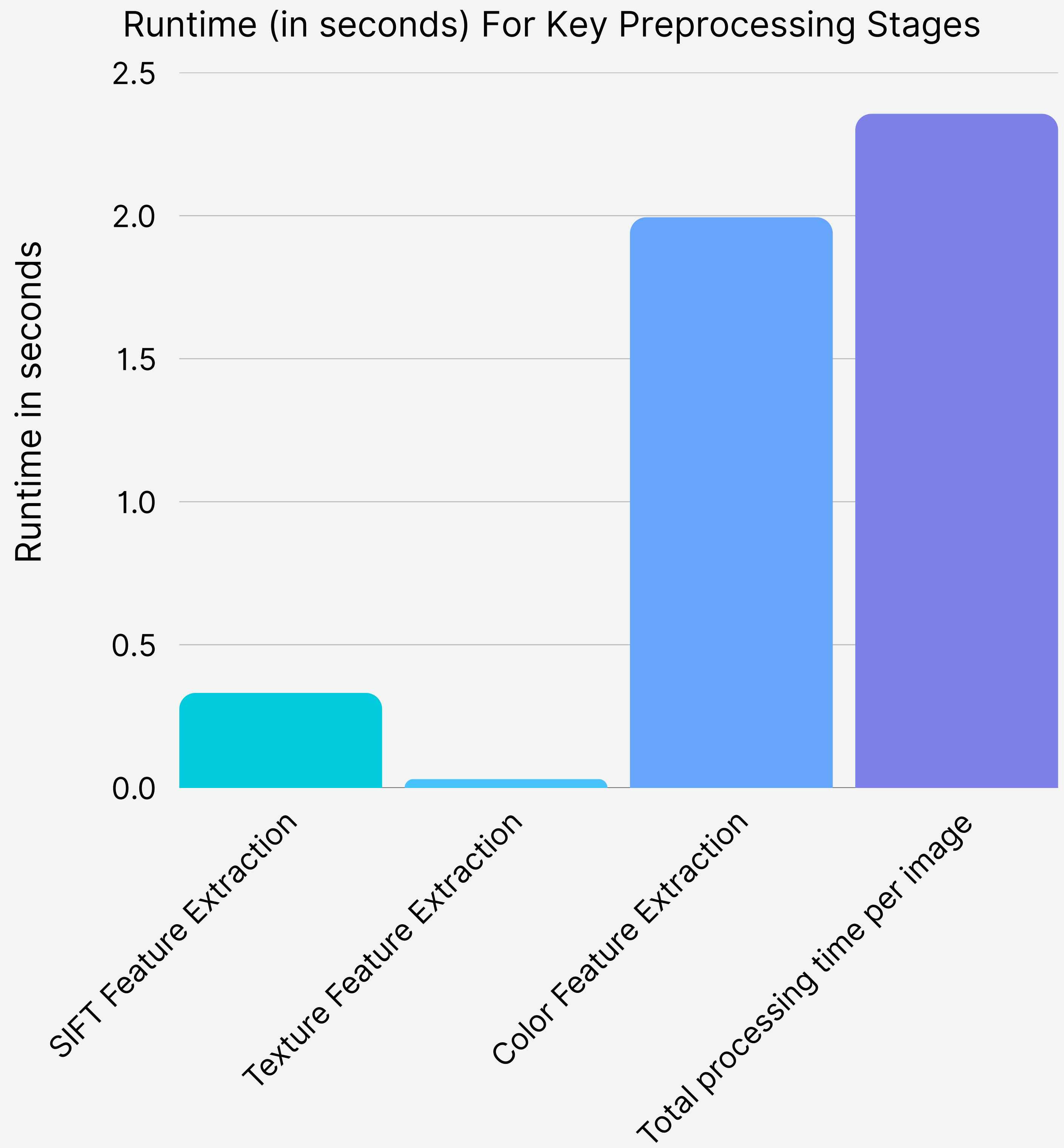
Hierarchical clustering with k=3 (silhouette score: 0.0853):

- Group 1: Primarily watches with varied faces and bands - contains analog watches with different dial colors and strap styles
- Group 2: A distinctive decorative watch - isolated as its own cluster due to unique visual features
- Group 3: Eyewear, including multiple styles of sunglasses - successfully grouped functionally similar items

Ward's linkage method provided the most coherent groupings compared to average and complete linkage

Higher silhouette score (0.0853) than K-means at 5 (0.0298) indicates better natural grouping

Hierarchical approach demonstrated particular strength in grouping functionally similar items while maintaining visual coherence



09

PERFORMANCE ANALYSIS

Runtime breakdown:

- **SIFT Feature Extraction:** 0.3316 seconds
- **Texture Feature Extraction:** 0.0302 seconds
- **Color Feature Extraction:** 1.9944 seconds
- **Total processing time per image:** 2.3562 seconds

10 LIMITATIONS & CHALLENGES

FEATURE-BASED FASHION ACCESSORY CLUSTERING

- Highest silhouette score (0.7943) achieved with k=2 but resulted in imbalanced clusters (99.9% in one cluster)
- Mismatch between numerical evaluation metrics and human-perceived meaningful groupings
- Highlights need for both quantitative and qualitative evaluation

CROSS-CATEGORY VISUAL SIMILARITY

- Items from different functional categories (belts, watches) clustered together due to similar materials and colors
- Black leather items appeared virtually identical regardless of their actual category
- Feature extraction couldn't distinguish between same-material accessories with different functions

LIMITED COLOR DISCRIMINATION FOR DARK ACCESSORIES

- High concentration of dark-colored accessories in the dataset (as seen in hue histograms)
- Most values concentrated near 0 in hue channel
- Reduces effectiveness of color-based features despite their high weighting (5.0)

PREPROCESSING CHALLENGES FOR COMPLEX ITEMS

- Background removal occasionally struggled with fragile accessories or items with colors similar to background
- Skin detection sometimes misclassified non-skin regions with similar color profiles
- Accessories worn by models presented particular difficulties for clean isolation

"CURSE OF DIMENSIONALITY" IN HIGH-DIMENSIONAL FEATURE SPACE

- Combined feature vectors created a high-dimensional space where distance measures become less meaningful
- Made it challenging to find natural groupings in the data

CONCLUSIONS & KEY FINDINGS

TRADITIONAL FEATURES PROVIDE SUFFICIENT DISCRIMINATION

Combined SIFT, shape, color, and texture features effectively captured visual similarities without deep learning

DIFFERENTIAL WEIGHTING CRUCIAL (SHAPE: 6.0, COLOR: 5.0)

Empirical testing showed that overweighting shape and color features by 5-6× improved clustering quality dramatically

OPTIMAL CLUSTERS: K-MEANS (K=5), HIERARCHICAL (K=3)

Visual inspection revealed more meaningful groupings than statistical metrics alone would suggest

HIERARCHICAL CLUSTERING BETTER FOR FUNCTIONAL GROUPING

Ward's linkage method created clusters that aligned better with accessory categories (watches, eyewear, belts)

Overall finding

Traditional computer vision techniques remain valuable for fashion retail applications where labeled training data may be limited, achieving reasonable computational efficiency (2.36 seconds per image)

12 FUTURE WORK

NEURAL NETWORK INTEGRATION FOR SPECIFIC COMPONENTS

CNN-based background removal could improve preprocessing while maintaining the interpretability of the clustering stage

ADAPTIVE FEATURE WEIGHTING

Develop methods to automatically determine optimal feature weights for different accessory types through a semi-supervised approach with minimal labeled examples

HIERARCHICAL CATEGORY DETECTION

Implement a two-stage approach that first identifies broad functional categories (watches, belts, eyewear) before performing finer-grained visual similarity clustering within each category

COLOR FEATURE OPTIMIZATION

Optimize color feature extraction (currently 85% of processing time at 1.99 seconds per image) through more efficient algorithms or parallel processing

THANK YOU