TA’AM: Used-clothes Application

Saleh Adel Saleh
Computer Science Department
Faculty of Computer and Information Sciences, Ain Shams University
Cairo, Egypt
salehadel2001@gmail.com

Shehab Mostafa Fahmy
Computer Science Department
Faculty of Computer and Information Sciences, Ain Shams University
Cairo, Egypt
shehabmostafa2323@gmail.com

Abdelaziz Hossam Abdelaziz
Computer Science Department
Faculty of Computer and Information Sciences, Ain Shams University
Cairo, Egypt
zezohossam08@gmail.com

Abdulrhman Hosny Muhammed
Computer Science Department
Faculty of Computer and Information Sciences, Ain Shams University
Cairo, Egypt
abdulrhmanhosnymuhammed@gmail.com

Muhammed Mahros Muhammed
Computer Science Department
Faculty of Computer and Information Sciences, Ain Shams University
Cairo, Egypt
mhrwsmhmd8882@gmail.com

***Abstract*—The clothing industry contributes heavily to global carbon emissions and wastewater. This project introduces *Ta’am*, a platform promoting the resale and reuse of clothes to reduce environmental impact. *Ta’am* aims to empower users to buy and sell used garments, support local traders by showcasing their products, and improve user experience using advanced machine learning models. This application simplifies selling items by using machine learning to extract details from uploaded images, eliminating the need for users to manually input information. It ensures image quality and offers various search options. When a user uploads a product image, the system automatically creates a post with extracted attributes like category, color, and size, if the image meets quality standards. This process involves assessing image quality, detecting objects in the image, validating the content, estimating key points, calculating size, and detecting colors. Additionally, users can search for items using an image, with the system returning the most relevant matches based on image quality and similarity metrics.**

Keywords—Fashion, DeepFashion2, Agrigorev, BRISQUE, K-Means, ResNet, CBIR

I. INTRODUCTION

The clothing industry ranks fourth in environmental impact, responsible for 10% of global carbon emissions and 20% of wastewater. To address this, extending the lifespan of clothes and choosing second-hand over new ones can significantly reduce the industry's carbon footprint and wastewater contributions.

Creating a platform for users to sell, and buy clothes, particularly used garments, and providing local traders with a dedicated marketplace for showcasing their products is a compelling and innovative concept. Ta’am addresses two vital aspects. Firstly, it empowers individuals to extend the lifespan of their clothing, contributing to a more sustainable fashion industry by promoting the resale and reuse of garments. Secondly, it supports local traders who often lack a dedicated space to showcase their products. By doing so, Ta’am not only fosters economic growth but also strengthens local communities by bringing people together.

Our project relies on cutting-edge Convolutional Neural Networks (CNNs), which are smart computer models. They change how people use our platform to improve their experience. Let's talk about three important things that make our project special: checking image quality, picking out details, and making it easy to search for stuff.

II. RELATED WORK

A. Clothes Category Classification

Bossard et al. (2012) [9] used the ACS dataset to extract features including Histogram of Oriented Gradients (HOG), Speeded Robust Features (SURF), Local Binary Patterns (LBP), and color information. Bossard then used these features to perform multiclass classification with One vs. All SVM, random forests, and transfer forests, achieving average accuracies of 35.03%, 38.29%, and 41.36%, respectively. CNN exceeded these accuracy baselines on the ACS dataset when Lao, Brian, and Karthik A. Jagadeesh (2015) [10] used the standard AlexNet Convolutional network which had been trained using ImageNet and achieved a 50.2% accuracy on the test set.

Tatiana Sennikova (2021) [13] trained a model using the DeepFashion dataset and pre-trained ResNet34 model and achieved a Top-3 Accuracy of 88.6%, which is 6% higher than the benchmark accuracy and a Top-5 Accuracy of 94.1%, which is 4% higher than the benchmark accuracy. This should not come as a surprise as the authors of the original paper used VGG16 architecture as a backbone, which is a less powerful model. On loading 98 user-specified images, the Top-1 Accuracy of the model is 62.4%.

B. Clothes Key-points Estimation

Paulauskaite-Taraseviciene et al. [8] show that the most challenging task is key point detection as it directly depends on segmentation results and different garment types have different key points.

The study outlines a systematic approach to garment key-point estimation, leveraging a set of algorithms that are adaptable to different types of garments through parameter tuning and sensitivity adjustments.

C. Size Calculation using a Reference Object

Adrian Rosebrock. [14] Published an article that presents an approach to automatic size detection using the dimensions of a known-size object at the same image and the same surface.

The article shows that measuring the size of objects is similar to computing the distance from our camera to an object in both cases, we need to define a ratio that measures the number of pixels per a given metric we call it pixel per metric.

III. THE PROPOSED SYSTEM

This system architecture is composed of three layers. They are divided into the Presentation layer, Logic layer, and Data layer.

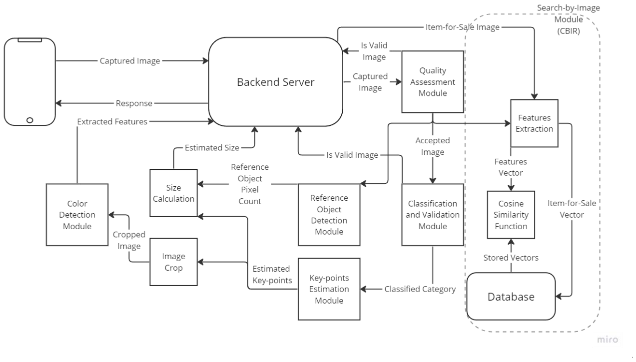

Fig. 1. System Architecture

- Presentation Layer:

  The system functions by receiving images for product posting or image-based search from users. It then transmits this data to the logic layer (server) for analysis. After processing, the system retrieves and presents the analyzed data to the user.

- Logic Layer:

  Information collected from the presentation layer is processed by the included models to apply attribute extraction and facilitation of search functionalities.

  - Scenario One:
    - When the user uploads a product image the system takes the image as input and produces a post of the product with the attributes extracted "Category, Color, Size" in case the image satisfies the image quality assessment criteria.
    - Pipeline:
      - Quality Assessment
      - Reference Object Detection
      - Classification and Validation
      - Key-points Estimation
      - Size Calculation Equation
      - Color Detection
  - Scenario Two:
    - Where users initiate searches using images, the system returns the closest images to the query in case the image satisfies the image quality assessment criteria.
    - Pipeline:
      - Quality Assessment
      - Classification and Validation
      - Feature Extraction
      - Cosine Similarity

- Data Layer:

  The data access is where the information processed by the application is stored and managed. The users' data (username, password, history, etc, are stored using Firebase. There is a local database used as well to store the post details and work with the machine-learning models.

A. *Quality Assessment*

Using the BRISQUE model, the image quality score was computed, and a specific threshold of 30% was applied for image assessment.

The Brisque model is a no-reference image quality assessment algorithm designed to predict the perceived quality of images without comparing them to reference images.

The Brisque model extracts statistical features from images, such as contrast, brightness, and texture. It then utilizes a machine learning algorithm, typically based on Support Vector Regression (SVR), to map these features to perceived image quality scores. This mapping is learned from a large dataset of images manually rated for quality.

B. *Reference Object Detection*

The process of extracting the ID card involves contour analysis, a vital technique in image processing and computer vision. Contour analysis entails enhancing the image using methods like thresholding, edge detection, and morphological operations to isolate regions of interest and reduce noise.

Key steps include:

- Edge Detection: Utilizing methods such as Canny or Sobel to detect edges within the image.
- Morphological Operations (Dilation): Expanding object boundaries by adding pixels, useful for tasks like filling gaps or enlarging objects.
- Gaussian Blur: Reducing image noise and smoothing details by applying a Gaussian function.
- Conversion to HSV Color Space: Representing the image based on Hue, Saturation, and Value components.
- Processing on S Channel: Applying Gaussian Blur and Sobel edge detection on the Saturation channel.
- Dilation: Expanding the edges obtained from the Sobel operation.
- Thresholding: Applying threshold values to the grayscale channel and performing contour analysis to identify regions resembling the ID card.
- Masking: Obtaining the mask of the largest contours, which represents the ID card.

C. *Classification and Validation*

This model validates whether the image contains clothes or not, based on that, it returns the classified category or "Other" and in that case it is rejected.

D. Key-points Estimation

13 pre-trained ResNet50 models were employed to extract key points for each category, then the height and width of the product were determined by calculating the distances between specific key points.

E. Size Calculation Equation

An EN 13402 Standard for clothing size designation was used to map the measurements from centimeters to real-world measurements “S, M, L, XL”.

Figure 2 shows an example of the equation applied to get the waist of trousers.

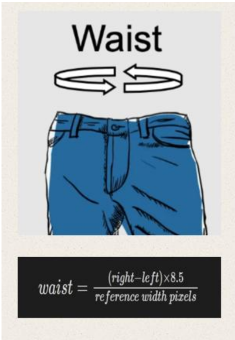

Fig. 2. Size Calculation example

F. Color Detection

The image was cropped using the detected key points to ignore background colors, then K-Means is used to segment the image into regions of similar colors and identify representative colors for each region, forming the color palette.

The execution time to extract the most three dominant colors is approximately 0.1 seconds.

G. Features Extraction

The MobileNet model extracts features from the query image, generating a 1 x 50176-dimensional vector within a time frame of 20ms to 50ms.

Utilizing cosine-similarity metrics, the closest 5 products to the query image are calculated. Comparing the query image with 5000 images, from the closest to the farthest, typically takes approximately 0.7s to 1s.

The vector dimensionality is reduced to 5000 using an AvgPooling Layer followed by a Dense Layer, resulting in the model being 10 times faster.

Comparison is limited to the same category rather than all categories, making the model 13 times faster and taking around 30ms to 50ms to execute.

IV. RESULTS

A. Classification and Validation

We utilized the ResNet34 architecture on Agrigorev’s custom kaggle dataset for category classification, achieving an accuracy of 88.5% after 50 epochs as shown in Table I.

TABLE I. CLASSIFICATION MODELS COMPARISON

Dataset	Architecture	Accuracy
FashionMNIST	Custom CNN Architecture	92.4% (e=10)
Agrigorev’s Custom Kaggle Dataset	ResNet34	88.5% (e=50) 82.3% (e=100)
DeepFashion2	ResNet50	90.5% (e=40)

B. Key-points Estimation

Table II shows the results of each pre-trained ResNet50 model for each category of DeepFashion2 dataset.

TABLE II. MAP AND MSE FOR EACH CATEGORY

Category	mAP	MSE
short sleeve top	0.8111	0.0588
skirt	0.8646	0.0652
vest	0.7933	0.0622
vest dress	0.8271	0.0570
short sleeve dress	0.7920	0.0630
short sleeve outwear	0.6566	0.0989
trousers	0.8624	0.0605
shorts	0.8936	0.0578
sling dress	0.7215	0.0814
long sleeve dress	0.6278	0.1069
long sleeve outwear	0.6930	0.0949
sling	0.6897	0.0840
long sleeve top	0.6906	0.0932
short sleeve top	0.8111	0.0588

C. Color Detection

After estimating the key-points, they are used to crop the image in a way to increase the efficiency of color detection model. Various color detection models were experimented as shown in Table III.

TABLE III. COLOR DETECTION MODELS COMPARISON

	ColorThief	Mini-Batch K-Means
Published In	2017	2018
Execution Time	3 colors: 0.34 sec 10 colors: 0.37 sec 100 colors: 0.4 sec	3 colors: 0.1 sec 10 colors: 0.3 sec 100 colors: 0.4 sec
Accuracy	Accurate	More Accurate

D. Features Extraction

Features extraction model is the cornerstone of Content-Based Image Retrieval (CBIR) or searching by image. Different architectures were used for higher retrieval similarity and lower processing time as shown in Table IV.

TABLE IV. FEATURES EXTRACTION MODELS COMPARISON

	VGG16	ResNet50	MobileNet
Feature Extraction per Image	1s to 2s	20ms to 50ms	20ms to 50ms
Vector Dimension	25088	100352	50176
Cosine-similarity with 5000 Images	0.4s to 0.7s	1.5s to 1.8s	0.7s to 1s

V. Conclusion

This project introduces innovative models designed for extracting features, evaluating image quality, and enhancing search capabilities using images. The feature extraction model, powered by deep learning techniques, aims to accurately identify and classify clothing attributes like category, size, and color. Quality assessment involves using the BRISQUE model to ensure uploaded images meet specified standards before processing.

Additionally, the system leverages advanced deep learning techniques to detect and analyze key points critical for tasks such as estimating garment sizes based on a reference object. This approach enhances accuracy and improves user experience by automating these processes through sophisticated machine-learning models.

The models implemented in this project were trained on comprehensive datasets: ResNet34, trained on Agrigorev's Custom dataset, focuses on classification and validation tasks, while ResNet50, trained on the DeepFashion2 dataset, specializes in key-point estimation. These datasets provide a robust foundation for training models that effectively handle diverse clothing items and scenarios.

Overall, this research contributes to advancing automated systems for image-based clothing monitoring, offering practical benefits to local traders and users by facilitating the sale of surplus clothing items with greater ease and accuracy.

Acknowledgment

We express our deepest gratitude to Allah for His strength, wisdom, and perseverance throughout our journey. We are also profoundly thankful to our families for their unwavering support and belief in us. Special thanks to our supervisors, *Dr. Salsabil Amin* and *TA Ahmed Hatem*, for their mentorship and guidance. Finally, we acknowledge our friends and all contributors for their support and resources, which have been invaluable to our research and documentation process.

References

[1] "What is Computer Vision?" IBM, [Online]. Available: https://www.ibm.com/topics/computer-vision

[2] "What are Convolutional Neural Networks Vision?" IBM, [Online]. Available: https://www.ibm.com/topics/convolutional-neural-networks

[3] Kristina Libby, "What is Object Classification and How Can We Use it" Shutterstock, [Online]. Available: https://www.shutterstock.com/blog/object-classification-in-computer-vision

[4] Petru Potrimba, "What is Key-Point Detection" roboFlow, [Online]. Available: https://www.shutterstock.com/blog/object-classification-in-computer-vision

[5] "A Complete Guide to Data Augmentation" DataCamp, [Online]. Available: https://www.datacamp.com/tutorial/complete-guide-data-augmentation

[6] "What Is Transfer Learning" BuiltIn, [Online]. Available: https://builtin.com/data-science/transfer-learning

[7] Alinder, Helena. "Semantic Image Segmentation on Clothing Imagery with Deep Neural Networks." (2020).

[8] Paulauskaite-Taraseviciene, Agne, et al. "An intelligent solution for automatic garment measurement using image recognition technologies." Applied Sciences 12.9 (2022): 4470.

[9] Bossard, Lukas, et al. "Apparel classification with style." Computer VisionACCV 2012. Springer Berlin Heidelberg, 2013. 321-335.

[10] Lao, Brian and Karthik A. Jagadeesh. "Convolutional Neural Networks for Fashion Classification and Object Detection." (2015).

[11] S. S. Islam, E. K. Dey, M. N. A. Tawhid, and B. M. M. Hossain, "A CNN Based Approach for Garments Texture Design Classification", Adv. technol. innov., vol. 2, no. 4, pp. 119–125, May 2017.

[12] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[13] Tatiana Sennikova. "Clothes Classification with the DeepFashion Dataset and Fastai" towardsdatascience, [Online]. Available: https://towardsdatascience.com/clothes-classification-with-the-deepfashion-dataset-and-fast-ai-1e174cbf0cdc

[14] Adrian Rosebrock. "Measuring size of objects in an image with OpenCV". PylImageSearch. [Online]. Available: https://pyimagesearch.com/2016/03/28/measuring-size-of-objects-in-an-image-with-opencv/

[15] Calorie Me, "Image-based Calorie Estimator System." FCIS-GP 2023.

[16] Ge, Y., Zhang, R., Wu, L., Wang, X., Tang, X., & Luo, P. (2019). A Versatile Benchmark for Detection, Pose Estimation, Segmentation, and Re-Identification of Clothing Images. CVPR.

[17] OLOLO. (2020). "Clothing dataset (full, high resolution)". Kaggle. Available: Clothing dataset (full, high resolution) (kaggle.com)