**Hair Style Recommendation System: Project Report**

# 1)Introduction

The Hair Style Recommendation System leverages computer vision and machine learning to analyze facial features and recommend hairstyles that complement a users face shape. By pro- cessing an uploaded facial image, the system detects landmarks, classifies the face shape (oval, round, square, heart, or long), and suggests flattering hairstyles based on stylistic guidelines. Built using Python libraries like DLib and face\_recognition, the system is deployed as a web application for user accessibility. This project addresses the challenge of subjective hairstyle selection, offering an automated, data-driven solution for individuals and salon professionals.

# 2)Why This Project and Societal Impact

# Motivation

The project was chosen due to its intersection of computer vision, machine learning, and practi-

cal application in the beauty industry.

## Societal Impact

The system has significant societal implications:

* + - **Enhanced Confidence**: By recommending hairstyles tailored to facial features, the sys- tem boosts users confidence, particularly for special occasions or professional settings.
    - **Time and Cost Efficiency**: It reduces the trial-and-error process in salons, saving time and money for users and stylists.
    - **Inclusivity**: The system caters to diverse face shapes and preferences, promoting inclu- sivity in beauty standards.
    - **Industry Innovation**: Salons adopting such technology can differentiate themselves, appealing to tech-savvy consumers and enhancing service quality.

By addressing personal, professional, and cultural needs in hairstyle selection, the system em- powers users and supports the evolution of the hair care industry.

# 3)Methodology

## System Architecture

The system processes a user-uploaded image through a pipeline of image acquisition, pre- processing, feature extraction, face shape classification, and hairstyle recommendation. The architecture is modular, ensuring scalability and ease of maintenance. Figure [1](#_bookmark0) illustrates the workflow.

Input Image

Preprocessing

Output Recommendations

Recommendation Engine

Face Shape Classifier

Feature Extraction

Figure 1: System Architecture of the Hair Style Recommendation System

## Architecture Explanation

The system operates as follows:

1. **Image Acquisition and Preprocessing**: Users upload a facial image via a web interface. The face\_recognition library crops the face based on eye positions (with a 34% offset) and resizes it to 300x300 pixels, normalizing pixel values for consistent processing.
2. **Feature Extraction**: DLibs 68-point facial landmark detector identifies key points, par- ticularly along the jawline and chin. From these, 23 features are computed, including:
   * Angles between the chin (point 9) and jawline points (18, 1017).
   * Face width (distance between points 1 and 17).
   * Face height (forehead to chin).
   * Ratios: face width to height, jaw width (points 7 to 11) to face width, mid-jaw width (points 5 to 13) to jaw width.

These geometric features capture face shape characteristics.

1. **Face Shape Classification**: A machine learning classifier (e.g., Random Forest or SVM) trained on the extracted features classifies the face shape into oval, round, square, heart, or long.
2. **Recommendation Engine**: Based on the classified face shape and user inputs (e.g., hair length, up-do preference), the system maps to a predefined set of hairstyle recommen- dations (e.g., beachy waves for oval faces, voluminous styles for square faces). Recom- mendations are displayed via the web interface.

The system is deployed using Flask, ensuring a user-friendly experience.

## 4)Dataset Description

The dataset was curated by analyzing 234 celebrity images from 22 fashion and style websites, focusing on face shape classifications. Of these, 33 images had unanimous face shape labels from three or more sources, and 65 had agreement from two or more, ensuring reliable anno- tations. The dataset includes images labeled with one of five face shapes: oval, round, square, heart, or long, along with 23 geometric features derived from facial landmarks.

## Dataset Characteristics

* + - **Size**: 234 images, with 33 high-confidence (3+ sources) and 65 moderate-confidence (2+ sources) labels.
    - **Source**: Publicly available celebrity images from fashion websites, capturing diverse facial features and lighting conditions.
    - **Annotations**: Each image includes a face shape label and 23 features (e.g., jawline an- gles, face width/height ratios) for classifier training.
    - **Usage**: The dataset is split into training and testing sets, with high-confidence images used for validation to ensure robust model performance.

The datasets expert-sourced labels reduce subjectivity, though its small size suggests potential for augmentation or expansion.

# 5) Result Comparison

The performance of the face shape classifier was evaluated by comparing five models: Mul- tilayer Perceptron (MLP), K-Nearest Neighbors (KNN), Random Forest (proposed), Gradient Boosting, and Linear Discriminant Analysis (LDA). The accuracy for each face shape and overall is presented in Table [1](#_bookmark1).

The table shows varied performance across face shapes:

* **MLP**: Achieved the highest overall accuracy (68%), excelling in long (80%) and round (74%) face shapes.
* **KNN**: Lowest overall accuracy (58%), performing poorly on most shapes except round (79%).
* **Random Forest**: Moderate overall accuracy (61%), with balanced performance across shapes, highest in round (76%).

Table 1: Performance Comparison of Face Shape Classification Models (Accuracy in %)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Shape | MLP | KNN | Random Forest | Gradient Boosting | LDA |
| Heart | 65.0 | 56.0 | 60.0 | 59.0 | 59.0 |
| Long | 80.0 | 63.0 | 54.0 | 57.0 | 50.0 |
| Oval | 72.0 | 51.0 | 59.0 | 66.0 | 71.0 |
| Round | 74.0 | 79.0 | 76.0 | 76.0 | 69.0 |
| Square | 62.0 | 49.0 | 61.0 | 68.0 | 67.0 |
| Overall | 68.0 | 58.0 | 61.0 | 65.0 | 64.0 |

* **Gradient Boosting**: Strong overall accuracy (65%), particularly for square (68%) and oval (66%) shapes.
* **LDA**: Competitive overall accuracy (64%), performing well on oval (71%) and square (67%) shapes.

The choice of Random Forest in the system balances consistency across shapes, though MLP and Gradient Boosting show higher overall accuracy.

# 6)Justification

* + - **MLP**: Excelled due to its ability to model non-linear relationships in the 23 geometric features. Its high accuracy for long (80%) and round (74%) shapes suggests it effec- tively captured complex patterns in these categories, likely benefiting from its layered architecture that learns hierarchical feature representations.
    - **Gradient Boosting**: Performed well (65% overall) by iteratively improving weak learn- ers, particularly for square (68%) and oval (66%) shapes. Its strength lies in handling small datasets and focusing on misclassified instances, enhancing generalization.
    - **LDA**: Achieved strong results for oval (71%) and square (67%) shapes by maximizing class separability in a linear subspace, which works well when features (e.g., geometric ratios) are linearly separable.

## Why Others Did Not Perform Well

* + - **KNN**: Underperformed (58% overall) due to its sensitivity to high-dimensional data (23 features). Its distance-based approach struggled with noisy or overlapping features, as seen in low accuracy for long (63%) and square (49%) shapes, where geometric features may not be distinct enough.
    - **Random Forest**: While consistent, its overall accuracy (61%) was lower than MLP and Gradient Boosting, possibly due to limited feature diversity in the dataset. It struggled with long (54%) shapes, where feature variance might be high, reducing ensemble effec- tiveness.

The small dataset size (234 images) favored models like MLP and Gradient Boosting, which adapt well to limited data through non-linear learning and iterative optimization. KNNs poor performance highlights its unsuitability for high-dimensional, small datasets without significant feature engineering.

# 7)Conclusion

The Hair Style Recommendation System, adapted from hussein073s GitHub repository, effec- tively automates hairstyle selection by classifying face shapes and recommending personalized styles. The systems architecture, leveraging DLib and face\_recognition, ensures efficient pro- cessing, while the dataset of 234 celebrity images provides a reliable foundation. The ablation study reveals MLPs superior overall accuracy (68%), with Gradient Boosting (65%) and LDA (64%) also performing well, while KNN (58%) struggled with the high-dimensional features. Random Forest (61%) was chosen for its balanced performance across face shapes, though future iterations could explore MLP or Gradient Boosting for higher accuracy. The system en- hances user confidence, promotes inclusivity, and streamlines salon processes, demonstrating its value in the beauty industry. Future work could involve expanding the dataset, integrating deep learning models like CNNs, and adding hair texture analysis for more comprehensive recommendations.

**TEAMS:**

**1.M Natrajan (2023510041)**

**2.S Selvamani(2023510035)**

**3.U Subramani(2023510024)**

**4.A Heldon leo(2023510029)**