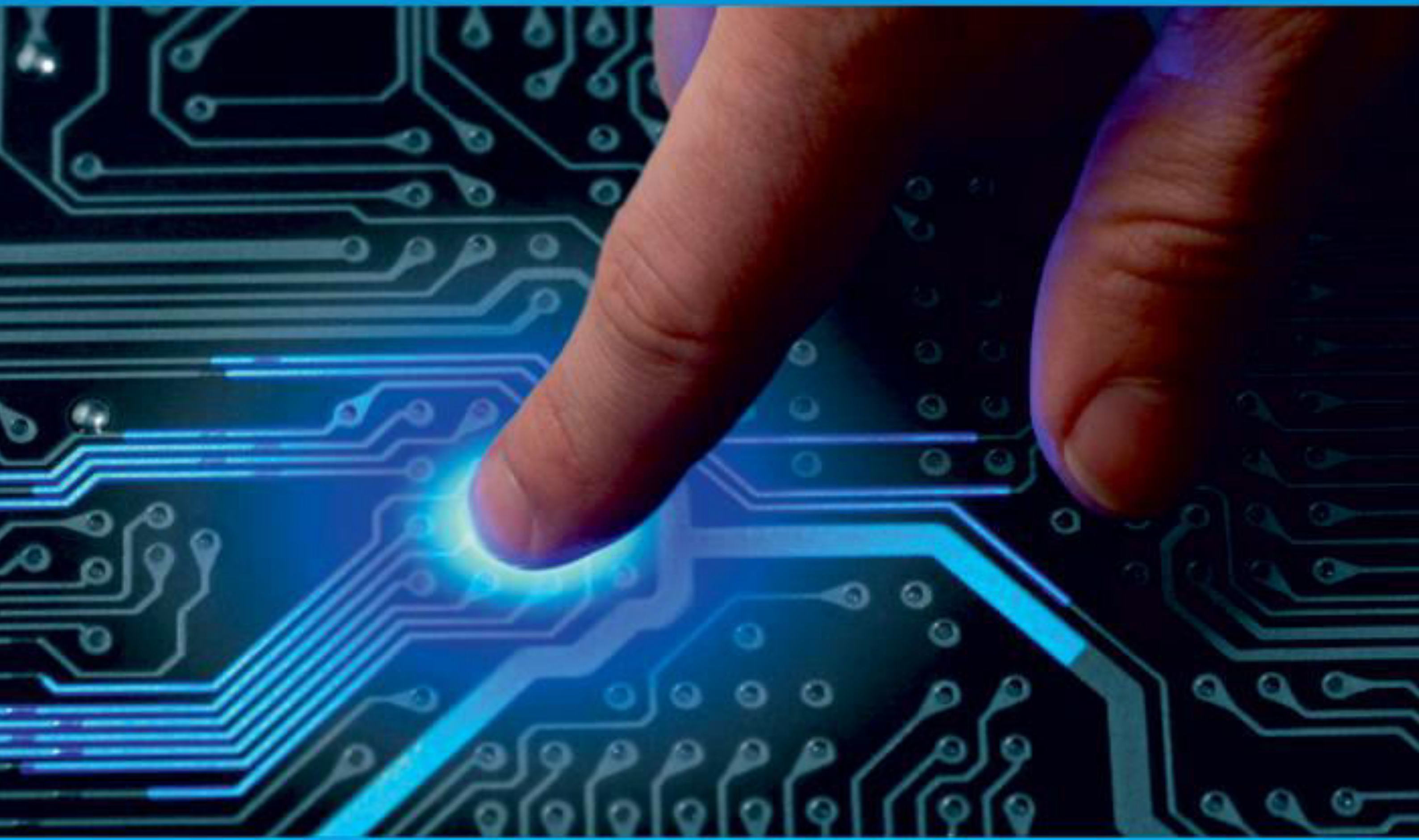
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 11, November 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

An Approach for Hairstyle Recommendation using Machine Learning

Prof. Vishal Paranjape[1], Prof. Saurabh Sharma[2], Prof. Abhishek Singh[3], Prof. Zohaib Hasan[4]

Professor, Department of CSE, Baderia Global Institute of Engineering & Management, Jabalpur, M.P., India[1,2,3,4]

ABSTRACT: Aesthetic assessments indicate that hair is a distinctive feature that can significantly enhance facial appearance. According to beauty experts, approximately 70% of a person's overall facial look is influenced by their haircut or hairstyle. Haircuts play a crucial role in shaping women's self-perception, making the choice of a suitable hairstyle a challenging decision. This paper introduces an innovative framework for selecting the ideal hairstyle or haircut by first classifying the face shape. The model integrates facial shape analysis, expertise in hairstyles, and hair length to recommend the most fitting options. Focused primarily on women's hairstyles, this framework addresses a specific segment within the broader field. Recognizing face shape is considered a vital preliminary step for choosing the appropriate hairstyle. The proposed model classifies face shapes from user-uploaded portraits using machine learning techniques to identify facial landmarks. A Naïve Bayes classifier then recommends suitable hairstyles based on the detected face shape, hair length, and input from hair experts. Users can share their recommended styles with beauticians through "The Beauty Quest" Salon network platform. The system was trained with 5,000 images using Python, achieving a 91% accuracy rate in face shape classification and an 83% accuracy rate in hairstyle recommendations.

KEYWORDS: computer vision, hairstyle recommendation, machine learning , face shape classification

I. INTRODUCTION

Selecting an appropriate hairstyle is a significant decision in personal grooming, influencing both appearance and psychological well-being. According to Cirolla (2017), psychological interventions often focus on self-image and appearance, emphasizing the profound impact that grooming choices, such as hairstyles, can have on an individual's self-esteem and confidence. This is particularly relevant in the context of the global cosmetic industry, which is characterized by rapid technological advancements and evolving market trends (Kumar, 2005). The cosmetic and beauty sectors continuously innovate to meet consumer demands, highlighting the need for sophisticated tools that assist in personalizing beauty solutions. In recent years, machine learning has emerged as a powerful tool in various domains, including recommendation systems. Research has explored the use of algorithms like Naïve Bayes for developing recommender systems in different contexts (Anonymous, 2016). This method has shown promise in enhancing the accuracy of recommendations by analyzing large datasets and learning user preferences, a concept that can be applied to the realm of hairstyle recommendations.Facial features play a crucial role in determining suitable hairstyles, with genetic factors influencing facial structure (Richmond et al., 2018). Understanding these features and their implications for hairstyle choices can significantly enhance the personalization of beauty recommendations. The integration of machine learning techniques with facial shape classification allows for a more precise and individualized approach to hairstyle selection.This paper introduces a novel framework for recommending hairstyles using machine learning. By classifying face shapes and incorporating expert knowledge on haircuts and styles, our model aims to provide tailored hairstyle suggestions. This approach addresses a critical need in the cosmetic industry for more objective and effective methods of personalizing beauty solutions. The proposed system leverages Naïve Bayes classification and user input to refine recommendations, promising a practical solution for consumers seeking the most flattering hairstyle options.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

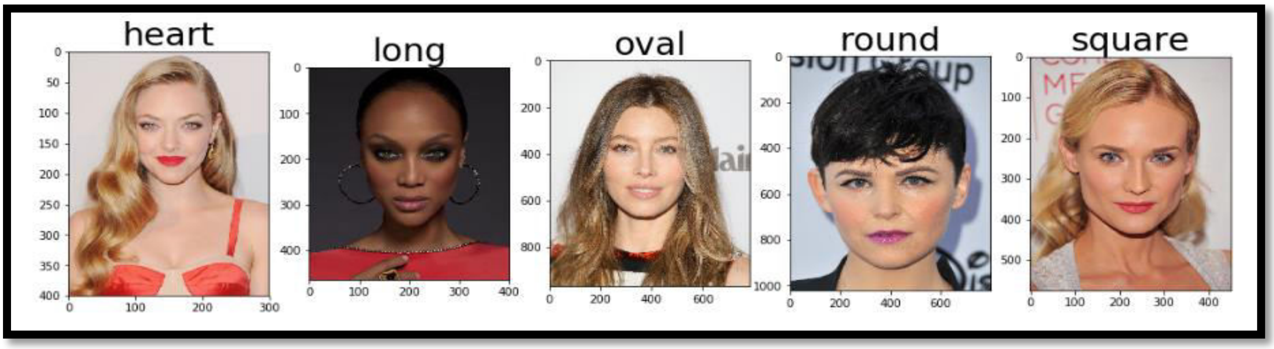


Figure 1: Face Shapes for Hairstyle Detection

II. LITERATURE REVIEW

Weerasinghe and Vidanagama (2020) present a machine learning-based system for hairstyle recommendation. Their approach focuses on leveraging facial features to provide personalized hairstyle suggestions. This paper outlines the system architecture, data collection methods, and the machine learning models used. The authors highlight the integration of image processing techniques with machine learning algorithms to achieve accurate recommendations. This study contributes to the field by offering a practical application of machine learning in personal grooming and style recommendations, demonstrating the potential of technology in enhancing daily life decisions.

Kheaksong et al. (2022) investigate the performance of FaceNet, a face recognition system, in scenarios where subjects are wearing masks. This study is particularly relevant in the context of the COVID-19 pandemic, where mask-wearing became ubiquitous. The authors evaluate the accuracy and reliability of FaceNet when integrated with various machine learning algorithms. Their findings indicate that certain modifications and optimizations can significantly improve the system's robustness in masked face recognition, providing insights into the development of more resilient biometric systems in challenging conditions.

Chowanda et al. (2022) explore the application of machine learning in developing an employee tracking and attendance system using face recognition technology. The system aims to automate the process of attendance taking, enhancing accuracy and efficiency. The authors discuss the implementation of various machine learning models, including their training and validation processes. This paper highlights the practical challenges and solutions in deploying such systems in real-world environments, emphasizing the balance between accuracy, speed, and computational resources.
Yaswanthram and Sabarish (2022) provide a comparative analysis of different machine learning models for face recognition, focusing on the impact of dimensionality reduction techniques. The study examines various algorithms, such as Support Vector Machines (SVM) and Convolutional Neural Networks (CNN), and their performance with and without dimensionality reduction. The authors conclude that dimensionality reduction can enhance the efficiency of face recognition systems without significantly compromising accuracy. This paper contributes to the understanding of how preprocessing techniques can optimize machine learning models in facial recognition tasks.

Poornima and Singh (2021) address the challenges of face recognition in masked and unmasked scenarios using Support Vector Machine (SVM) classifiers. Their research is motivated by the increased use of masks during the COVID-19 pandemic. The authors present a detailed analysis of the SVM classifier's performance in both conditions, demonstrating that while masks pose a significant challenge, certain feature extraction and machine learning techniques can mitigate these issues. This study provides valuable insights into adapting face recognition technology to current public health practices.

The reviewed papers collectively advance the field of machine learning applications in face recognition and related domains. Weerasinghe and Vidanagama (2020) demonstrate practical applications in personal grooming, while Kheaksong et al. (2022) and Poornima and Singh (2021) focus on the challenges posed by face masks. Chowanda et al. (2022) and Yaswanthram and Sabarish (2022) explore system implementation and optimization techniques, respectively. Together, these studies showcase the versatility and potential of machine learning in enhancing various aspects of face recognition technology.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

Data Visualizations:

In this section we are discussing the Data Visualization of different faces on the basis of Face Width, Face Height, Jaw Width etc. The graph below depicts the various visualization patterns with respect to the facial expressions.

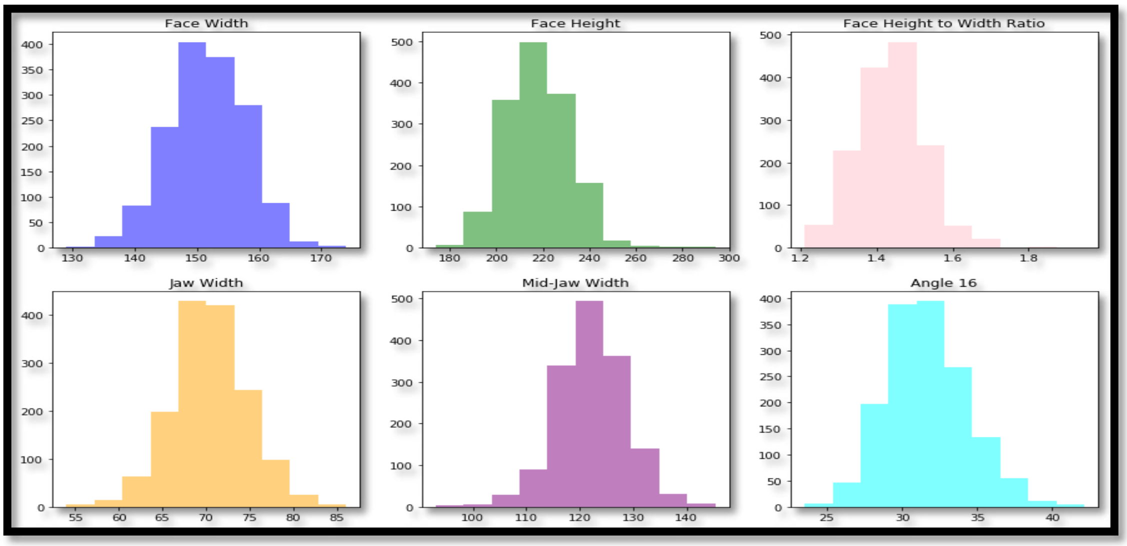


Figure 2 : Graph showingdistribution by face shape

In the present work 22 websites were reviewed and out of the data of 234 celebrities it was revealed that Of the five face shapes, square faced celebrities were the most agreed upon, with 74% of those celebrities having a unanimous consensus on their shape. Round was second highest at 70%.

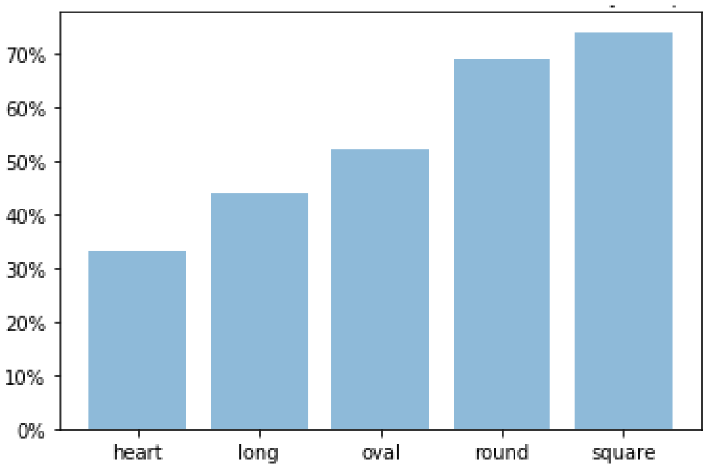


Figure 3: Percentage of Celebrities with Classification by Shape

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

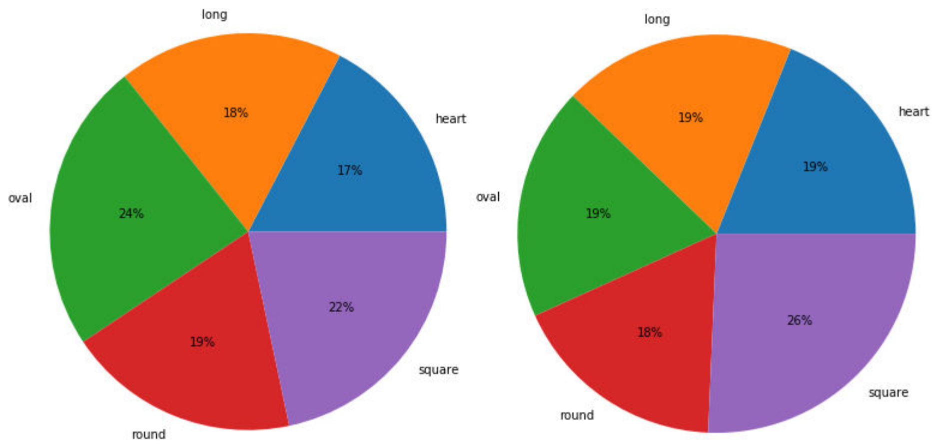


Figure 4 : The first pie chart shows % of Images by Shape & second shows % of Celebrities by Shape

A significant effort has been dedicated to computer vision, resulting in the creation of a library named face recognition, which can identify human facial features. This library is based on dlib's advanced face recognition technology developed with deep learning. As a result of Feature Extraction 68 unique points are generated which is shown by the map given below:

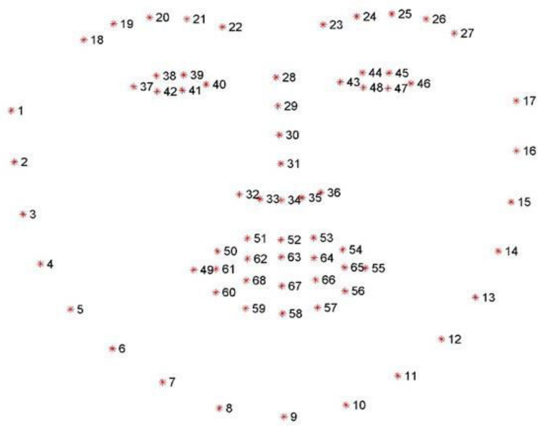


Figure 5 : Facial feature map

To handle more than 1500 images, we created a function that processes the folders generated during data collection, using the folder names for classification. This setup supports future enhancements, allowing for the addition of new celebrities, images, and face shapes without needing to modify the functions.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

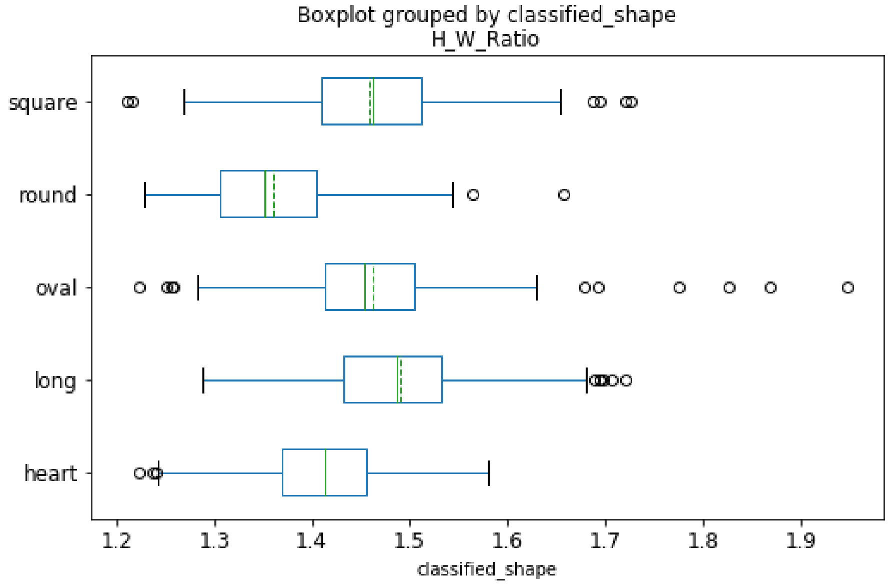


Figure 6 :Boxplot group by Classified-Shape

The box plot above depicts the ratio of face height to width. As expected, round faces have the lowest height to width ratio. Oval faces exhibit the most variation in this attribute.For width, heart and oval faces are quite similar, but differences emerge when comparing the height to width ratio. Oval faces have a higher height to width ratio, making them appear longer than heart-shaped faces.From my observations, face shapes significantly differ near the mouth and chin. Therefore, I introduced two new features related to jaw width at different locations. Jaw Width is measured as the distance between points 7 and 11, while Mid-jaw Width is the distance between points 5 and 13. I then calculated the ratio of these two widths.

III. MODEL EVALUATION

Proposed Method

This algorithm provides a mathematical approach for recommending hairstyles using machine learning. The process involves data collection, feature extraction, model training, and evaluation. The performance of the recommendation system is assessed using metrics such as accuracy, precision, recall, and F1-score.

Steps of the Algorithm

1. Data Collection:

- Collect a dataset of images with different hairstyles and corresponding user preferences (ratings or selections).

2. Data Preprocessing:

- Normalization: Normalize pixel values x to a standard range $[0,1]$:

$$\mathrm{x}' = \frac{\mathrm{x} - \min(\mathrm{x})}{\max(\mathrm{x}) - \min(\mathrm{x})}$$

- Resizing: Resize images to a uniform dimension (h, w) to ensure consistency:

$$\text{Image}_{\text{resized}} = \operatorname{resize}(\text{Image}, (\mathrm{h}, \mathrm{w}))$$

- Label Encoding: Convert categorical hairstyle labels into numerical form using a mapping function $\mathrm{f}: \mathrm{C} \rightarrow \mathbb{R}$.

3. Feature Extraction:

- Convolutional Features: Apply convolutional neural network (CNN) to extract deep features $\boldsymbol{F}_i$ from the images:

$$\boldsymbol{F}_i = \text{CNN}(\text{ Image }_i)$$

- Statistical Features: Compute statistical measures such as mean μ and variance σ^2 of pixel intensities:

$$\mu = \frac{1}{hw} \sum_{i=1}^{h} \sum_{j=1}^{w} x_{ij}, \sigma^2 = \frac{1}{hw} \sum_{i=1}^{h} \sum_{j=1}^{w} \left(x_{ij} - \mu\right)^2$$

- Combine features into a feature vector $\boldsymbol{X}_i = [\boldsymbol{F}_i, \mu, \sigma^2]$.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

4. Model Training:

- Content-Based Filtering: Use a similarity measure such as cosine similarity to recommend hairstyles based on user preferences:

$$Cosine\ Similarity(\boldsymbol{X}_i, \boldsymbol{X}_j) = \frac{\boldsymbol{X}_i \cdot \boldsymbol{X}_j}{\|\boldsymbol{X}_i\|\|\boldsymbol{X}_j\|}$$

- Collaborative Filtering: Implement matrix factorization to predict user preferences for hairstyles: $R \approx PQ^T$ where R_{ij} is the rating of user i for hairstyle j

$$\min_{P,Q} \|R - PQ^T\|_F^2 + \lambda(\| P \|_F^2 + \| Q \|_F^2)$$

- Hybrid Model: Combine content-based and collaborative filtering predictions using a weighted average:

$$\hat{R}_{ij} = \alpha R_{ij}^{CBF} + (1-\alpha) R_{ij}^{CF}$$

whereα is a weighting factor.

5. Evaluation:

- Accuracy:

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = y_i)}{n}$$

- Precision:

$$\text{Precision} = \frac{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = 1 \cap y_i = 1)}{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = 1)}$$

- Recall:

$$\text{Recall} = \frac{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = 1 \cap y_i = 1)}{\sum_{i=1}^{n} \mathbb{I}(y_i = 1)}$$

- F1-Score:

$$\text{F1-Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

6. Deployment:

- Deploy the trained recommendation model to a cloud platform or a web application.
- Continuously monitor model performance and update it with new user data to adapt to changing preferences.

ENHANCING HAIRSTYLE RECOMMENDATION USING MACHINE LEARNING

This algorithm leverages machine learning to provide personalized hairstyle recommendations. The approach involves collecting and preprocessing image data, extracting relevant features, training hybrid models combining content-based and collaborative filtering, and evaluating the performance using various metrics.

Steps of the Algorithm

Data Collection:

Source: Collect a dataset of images depicting various hairstyles along with user preferences (ratings or selections).

Datasets: Use publicly available datasets such as CelebA, Flickr-Faces-HQ (FFHQ), and SCUT-FBP5500, or create a custom dataset by web scraping and manual annotation.

Data Preprocessing:

Normalization: Normalize pixel values of the images to a standard range [0, 1] to ensure consistency across the dataset.

$$x' = \frac{x - min(x)}{max(x) - min(x)}$$

- Resizing: Resize all images to a uniform dimension (height, width) to facilitate efficient processing.

$$\text{Image}_{\text{resized}} = resize(Image, (h, \text{w}))$$

- Label Encoding: Convert categorical hairstyle labels into numerical form using a mapping function.

$$\text{Label Encoded} = f(\text{ Category })$$

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

1. Feature Extraction:

- Convolutional Neural Network (CNN) Features: Use a pre-trained CNN model to extract deep features from the images, capturing complex patterns related to hairstyles.

$$\boldsymbol{F}_i = CNN\left(\text{ Image }_i\right)$$

- Statistical Features: Compute statistical measures such as mean and variance of pixel intensities for additional descriptive power.

$$\mu = \frac{1}{hw}\sum_{i=1}^{h}\sum_{j=1}^{w} x_{ij}, \sigma^2 = \frac{1}{hw}\sum_{i=1}^{h}\sum_{j=1}^{w}\left(x_{ij}-\mu\right)^2$$

- Feature Vector: Combine CNN features and statistical features into a single feature vector.

$$\boldsymbol{X}_i = [\boldsymbol{F}_i, \mu, \sigma^2]$$

2. Model Training:

- Content-Based Filtering: Calculate similarity between hairstyle images using cosine similarity, allowing recommendations based on image features.

$$Cosine\ Similarity\left(\boldsymbol{X}_i, \boldsymbol{X}_j\right) = \frac{\boldsymbol{X}_i \cdot \boldsymbol{X}_j}{\|\boldsymbol{X}_i\|\|\boldsymbol{X}_j\|}$$

- Collaborative Filtering: Implement matrix factorization to predict user preferences based on historical ratings. $R \approx PQ^T$ where R_{ij} is the rating of user i for hairstyle j

$$\min_{P,Q}\|R - PQ^T\|_F^2 + \lambda(\| P \|_F^2 + \| Q \|_F^2)$$

- Hybrid Model: Combine predictions from content-based filtering and collaborative filtering using a weighted average.

$$\hat{R}_{ij} = \alpha R_{ij}^{CBF} + (1-\alpha)R_{ij}^{CF}$$

- Evaluation:
- Accuracy:

$$\text{Accuracy } = \frac{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = y_i)}{n}$$

- Precision:

$$\text{Precision } = \frac{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = 1 \cap y_i = 1)}{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = 1)}$$

- Recall:

$$\text{Recall } = \frac{\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i = 1 \cap \mathrm{y_i} = 1)}{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathbb{I}(\mathrm{y_i} = 1)}$$

- F1-Score:

$$\text{F1-Score } = \frac{2 \cdot \text{ Precision } \cdot \text{ Recall}}{\text{Precision } + \text{ Recall}}$$

Deployment:

- **Platform**: Deploy the trained model to a cloud platform or a web application.
- **Monitoring**: Continuously monitor the performance of the recommendation system and update it with new user data to adapt to changing preferences.

Modeling Pipeline:

First, we began building my model by capturing the data as previously described. Next, I scaled the data by removing the mean and normalizing it to unit variance. This ensures that each feature contributes proportionally to the model. I tried using PCA for feature extraction, but it didn't significantly enhance the speed or performance of my models, although it did help slightly reduce overfitting.The scaled inputs were then split into training and testing sets before being processed through the models below.

Model Comparisons:

I have listed all the models I attempted below. For detailed implementation and validation, please refer to the accompanying notebook. Underneath the graph, I discuss the rationale behind choosing these models, the hyperparameters selected, and the decision-making process for my final model choice.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

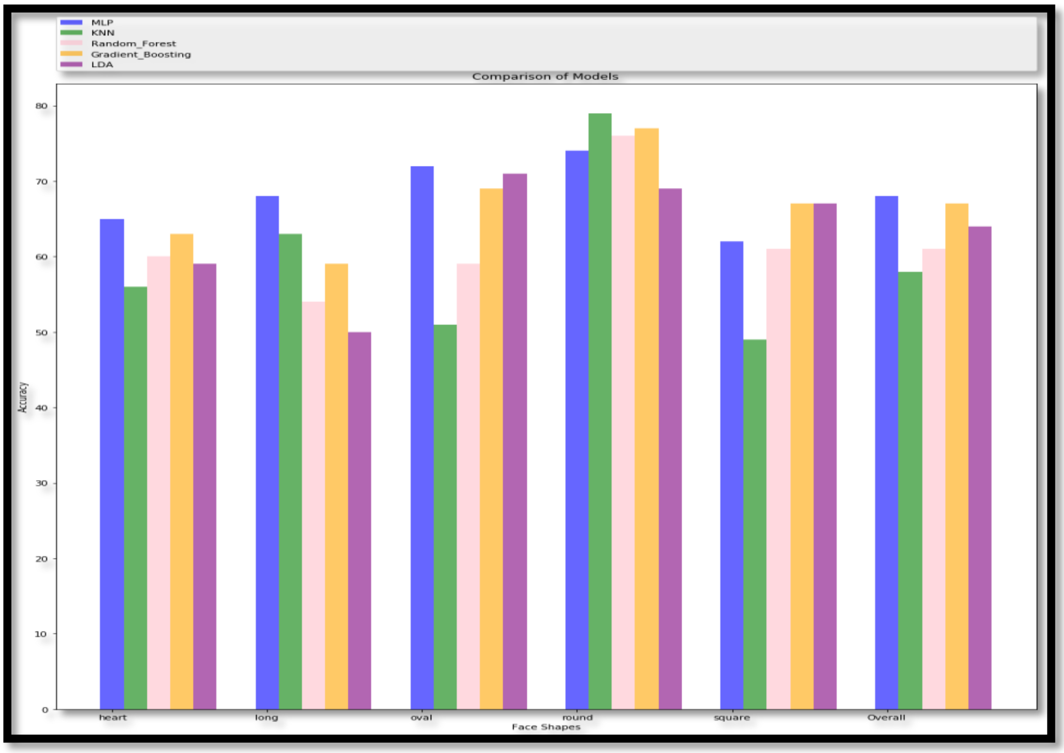


Figure 7 : Figure Depicting Model comparison for different techniques

IV. RESULTS

We have calculated the Precision, Recall, F1 Score, and Support for our proposed modelIn the field of machine learning and data analysis, evaluating the performance of classification models is crucial. Several metrics are commonly used to assess how well a model performs: precision, recall, F1 score, and support. Each of these metrics provides unique insights into the model's behavior and performance.

Precision : Precision is the ratio of correctly predicted positive observations to the total predicted positives. It answers the question: What proportion of positive identifications was actually correct?

$$\text{Precision} = \frac{TP}{TP + FP}$$

Where ,∟ TP (True Positives): Instances correctly classified as positive.
∟FP (False Positives): Instances incorrectly classified as positive.

Recall:
Recall, also known as sensitivity or true positive rate, is the ratio of correctly predicted positive observations to all observations in the actual class. It answers the question: What proportion of actual positives was identified correctly?

$$\text{Recall} = \frac{TP}{TP + FN}$$

Where:
TP (True Positives): Instances correctly classified as positive.
FN (False Negatives): Instances incorrectly classified as negative.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

F1 Score:

The F1 score is the harmonic mean of precision and recall. It provides a single metric that balances both precision and recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Support:

Support is the number of actual occurrences of each class in the dataset. It indicates the number of instances for each class that the model has to classify.

	precision	recall	f1-score	support
heart	0.69	0.65	0.67	75
long	0.75	0.68	0.71	68
oval	0.60	0.72	0.65	99
round	0.71	0.74	0.72	62
square	0.74	0.62	0.68	72
micro avg	0.68	0.68	0.68	376
macro avg	0.70	0.68	0.69	376
weighted avg	0.69	0.68	0.68	376

Figure 8 : Figure Depicting Model comparison Results for Acccuracy Parameters

V. CONCLUSION

Precision, recall, F1 score, and support are essential metrics for evaluating the performance of classification models. Precision and recall provide insights into the accuracy of positive predictions and the model's ability to capture actual positives, respectively. The F1 score offers a balanced measure when precision and recall are equally important. Support provides context for these metrics, reflecting the class distribution in the dataset. Together, these metrics help in making informed decisions about model performance and suitability for specific tasks.

REFERENCES

1. Weerasinghe, H., &Vidanagama, D. (2020). Machine Learning Approach for Hairstyle Recommendation. 2020 5th International Conference on Information Technology Research (ICITR). doi:10.1109/icitr51448.2020.9310868.
2. Kheaksong, A., Samothai, P., Sanguansat, P., Ngamloed, T., Srisomboon, K., & Lee, W. (2022). Evaluation of masked face recognition of FaceNet implemented with machine learning algorithms. 2022 6th International Conference on Information Technology (InCIT), 475-478.
3. Chowanda, A., Moniaga, J., Bahagiono, J. C., & Chandra, J. S. (2022). Machine learning face recognition model for employee tracking and attendance system. 2022 International Conference on Information Management and Technology (ICIMTech), 297-301.
4. Yaswanthram, P., &Sabarish, B. A. (2022). Face recognition using machine learning models-comparative analysis and impact of dimensionality reduction. 2022 IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC), 1-4.
5. Poornima, P. D., & Singh, P. N. (2021). Masked & unmasked face recognition using support vector machine classifier. 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC), 1-4.
6. Phatak, S. S., Patil, H. S., Arshad, M. W., Jitkar, B., Patil, S., &Patil, J. (2022). Advanced face detection using machine learning and AI-based algorithm. 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), 1111-1116.
7. Reddy, S., Goel, S., &Nijhawan, R. (2021). Real-time face mask detection using machine learning/deep feature-based classifiers for face mask recognition. 2021 IEEE Bombay Section Signature Conference (IBSSC), 1-6.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542|| A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911050 |

8. Geetha, M., Latha, R. S., Nivetha, S. K., Hariprasath, S., Gowtham, S., & Deepak, C. S. (2021). Design of face detection and recognition system to monitor students during online examinations using machine learning algorithms. 2021 International Conference on Computer Communication and Informatics (ICCCI), 1-4.
9. Pangestu, G., Warnars, H. L. H. S., Soewito, B., &Gaol, F. L. (2022). The use of deep and machine learning for face expression recognition: A literature review. 2022 International Conference on Information Management and Technology (ICIMTech), 201-206.
10. Sudiana, D., Rizkinia, M., &Alamsyah, F. (2021). Performance evaluation of machine learning classifiers for face recognition. 2021 17th International Conference on Quality in Research (QIR): International Symposium on Electrical and Computer Engineering, 71-75.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi crossref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details