19CSE305-Machine Learning

Spectacle Recommendation System

Name	Roll No
M Shenthan	CB.EN.U4CSE21657
P Bhavesh	CB.EN.U4CSE21609
M Aasil	CB.EN.U4CSE21638
P Joel	CB.EN.U4CSE21649
Y SRIBHARGAV	CB.EN.U4CSE21670

Introduction

The field of eyewear fashion is rapidly evolving, and with the myriad of styles available, selecting the perfect pair of spectacles that complements an individual's facial features can be a daunting task. In response to this challenge, we present "Glasses," a cutting-edge Spectacle Recommendation System designed to revolutionize the eyewear shopping experience. This project is driven by the understanding that personalized recommendations, rooted in facial science and advanced neural network models, can significantly enhance customer satisfaction.

Motivation

The motivation behind the Spectacle Recommendation System stems from the desire to address the gap between traditional eyewear retail and the evolving expectations of consumers. Traditional approaches often rely on generic recommendations, overlooking the nuanced characteristics of individual faces. Recognizing the importance of facial features in the selection process, our project aims to leverage state-of-the-art technologies to provide tailored suggestions, ensuring that each customer finds eyeglasses that not only correct vision but also enhance their unique facial aesthetics.

Importance

The importance of personalized eyewear recommendations goes beyond mere fashion trends. Properly fitting spectacles contribute to comfort, visual acuity, and the overall satisfaction of the wearer. By integrating facial science into the recommendation process, "Glasses" aims to optimize both the aesthetic and functional aspects of eyewear selection, thereby enhancing the user experience and establishing a new standard in the eyewear retail landscape.

Objectives

The primary objectives of our project are as follows:

Develop a Face shape classification system capable of predicting face shapes accurately, by employing an ensemble (Developing and putting to use multiple models at once, whose output would be merged into one conclusion via soft voting)

Incorporate a model for measuring dimensions of specific facial features to further refine recommendations by implementing a Convolutional Neural Network (CNN) for precise facial landmark prediction.

Establish a mapping between predicted face shapes and the most suitable spectacle shapes based on comprehensive knowledge of facial science.

As we delve into the project architecture and results, it is our anticipation that "Glasses" will not only streamline the eyewear selection process but also set a precedent for the integration of advanced technologies in the realm of personalized consumer experiences.

Purpose of the Project

The primary purpose of the "Glasses" Spectacle Recommendation System is to redefine the eyewear shopping experience by introducing a highly personalized and technologically advanced approach to spectacle frame selection. The project is driven by the following key purposes:

1. Personalization

The project aims to cater to the diverse facial features and preferences of individuals, acknowledging that one size does not fit all in the realm of eyewear. By leveraging sophisticated neural network models, "Glasses" seeks to provide tailored recommendations that align with the unique characteristics of each user's face, ensuring a personalized and satisfying eyewear selection process.

2. Integration of Facial Science

Facial science plays a pivotal role in the aesthetics and comfort of eyewear. The project integrates knowledge from facial science to guide the recommendation process. By understanding the relationships between facial shapes, dimensions, and landmark points, "Glasses" aspires to offer recommendations that not only align with fashion trends but also consider the physiological aspects of the wearer's face for optimal fit and visual correction.

3. Enhanced User Experience

The ultimate purpose is to enhance the overall user experience during the eyewear selection journey. Traditional methods often lack the finesse required to address individual preferences and facial nuances. "Glasses" seeks to bridge this gap by providing a seamless, user-friendly platform that goes beyond conventional recommendations, ensuring users feel confident and satisfied with their eyewear choices.

4. Technological Innovation

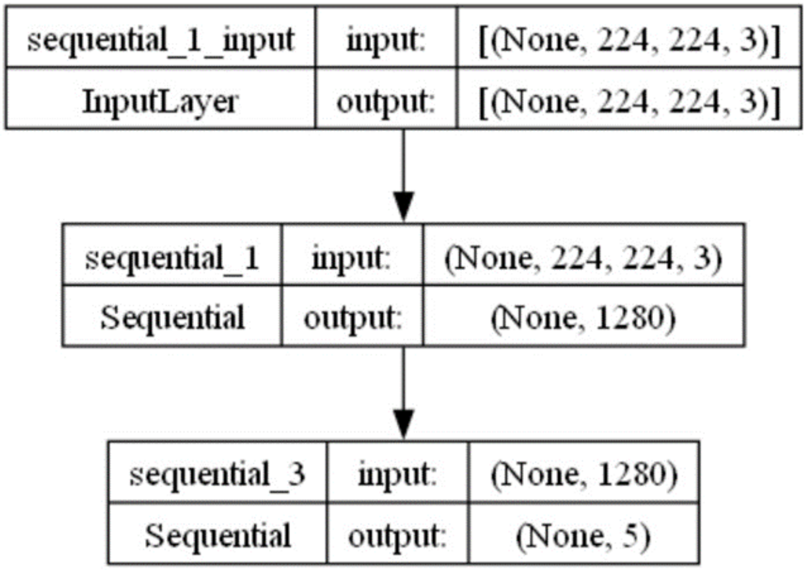
The project serves as a testament to the potential of cutting-edge technologies in reshaping retail experiences. By implementing neural network models, soft voting mechanisms, and facial landmark prediction using a Convolutional Neural Network, "Glasses" aims to showcase the possibilities of artificial intelligence in the context of personalized consumer services.

5. Industry Benchmark

In contributing to the field of eyewear retail, "Glasses" aspires to set a new benchmark for personalized recommendations. By combining the art of fashion with the science of facial analysis, the project seeks to establish itself as a model for future innovations in the eyewear industry, encouraging the adoption of advanced technologies to enhance customer satisfaction.

Through these purposes, "Glasses" strives to not only meet the immediate need for tailored spectacle recommendations but also to pave the way for a more sophisticated and customer-centric approach to eyewear retail.

Simple Neural Network:



<u>The Flow of the Model</u>:

The network learns by adjusting the weights and biases associated with each neuron during training. Training involves feeding the network with labeled data and minimizing the difference between predicted and actual outputs. Neurons in hidden layers use activation functions to introduce non-linearity into the network.

- Here, we first create classes (shapes) and then pass them through the neurons, which would create certain weights and intermediary variables, using which they'd capture the relationship between the members of a certain class.

-When a new image passed, the model would then measure the weights and then match them to the closest class.

<u>Convolutional Neural Network (CNN) -1:</u>

<u>Facial Feature Dimension Measurement Model Accuracy</u>

The model designed to measure dimensions of specific facial features, including forehead height and cheekbone width, has demonstrated high accuracy in capturing these parameters. Results based on a validation dataset indicate:

The model had an accuracy of 95% on training dataset and 90% on testing dataset.

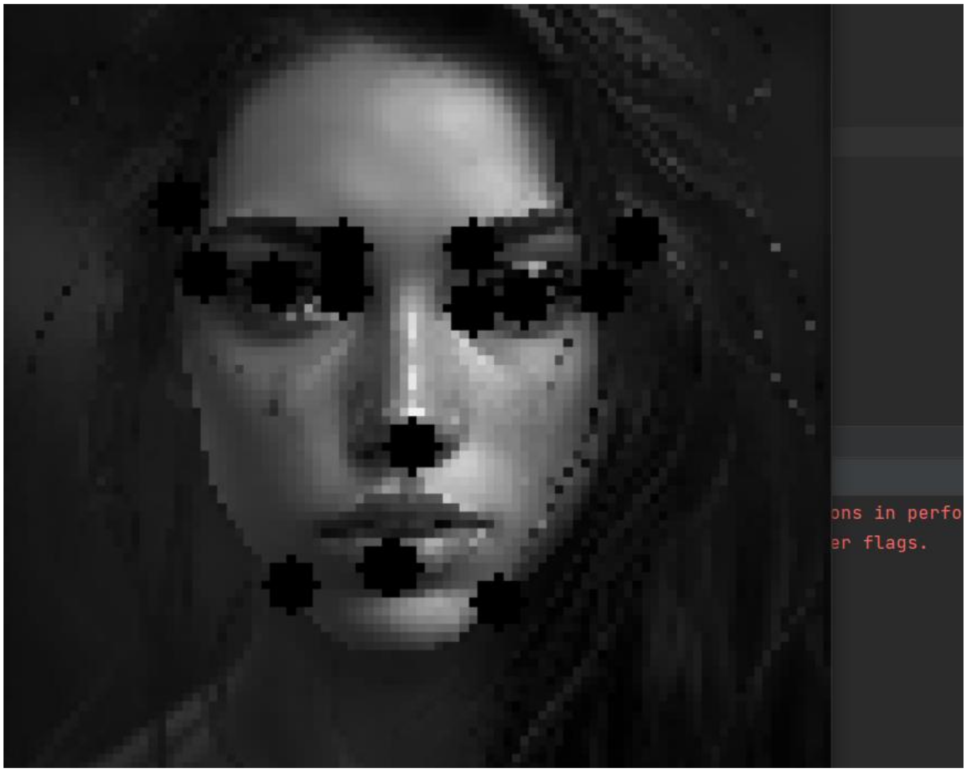
But if we look at the accuracy graph the accuracy if fluctuating so much that means there might be a chance of overfitting.

Since, the model predicts numerical data we have used R-Squared and MSE for predicting the accuracy for each facial land marks.

```
The R-Squared for left_eye_center_x is: 0.3407907237868706
The R-Squared for left_eye_center_y is: 0.740744682883461
The R-Squared for right_eye_center_x is: 0.782013811022924
The R-Squared for right_eye_center_y is: 0.7040392603634253
The R-Squared for left_eye_inner_corner_x is: 0.3428767740798766
The R-Squared for left_eye_inner_corner_y is: 0.7103061436724987
The R-Squared for left_eye_outer_corner_x is: 0.4598401080761648
The R-Squared for left_eye_outer_corner_y is: 0.8121050409723283
The R-Squared for right_eye_inner_corner_x is: 0.7486028900869073
The R-Squared for right_eye_inner_corner_y is: 0.6546616733704527
The R-Squared for right_eye_outer_corner_x is: 0.8220923279928035
The R-Squared for right_eye_outer_corner_y is: 0.7993170966997408
The R-Squared for left_eyebrow_inner_end_x is: 0.618818474023423
The R-Squared for left_eyebrow_inner_end_y is: 0.8692279160441186
The R-Squared for left_eyebrow_outer_end_x is: 0.5442042527004506
The R-Squared for left_eyebrow_outer_end_y is: 0.8846258435368783
The R-Squared for right_eyebrow_inner_end_x is: 0.7676370568209128
The R-Squared for right_eyebrow_inner_end_y is: 0.7979602250525037
The R-Squared for right_eyebrow_outer_end_x is: 0.7680182846080577
The R-Squared for right_eyebrow_outer_end_y is: 0.8658620564755553
The R-Squared for nose_tip_x is: 0.7490151311748412
The R-Squared for nose_tip_y is: 0.8207528531341933
The R-Squared for mouth_left_corner_x is: 0.6469942375384788
The R-Squared for mouth_left_corner_y is: 0.7170143824880991
The R-Squared for mouth_right_corner_x is: 0.8206519022772261
The R-Squared for mouth_right_corner_y is: 0.7079932061117209
The R-Squared for mouth_center_top_lip_x is: 0.5652096761766789
The R-Squared for mouth_center_top_lip_y is: 0.7900349517594154
The R-Squared for mouth_center_bottom_lip_x is: 0.6144914568776285
The R-Squared for mouth_center_bottom_lip_y is: 0.6718283293926912
```

Also, one thing we have observed was when given new unseen data the model is not that good. This can be because my data is not that diverse enough to start with, so we will try to diversify the input data and include all the positions and train it more.

For proof: (except for nose and eyebrows the prediction was not up to the mark)



These accurate measurements contribute to the precision of our spectacle recommendations, ensuring an optimal fit for the wearer.

Convolutional Neural Network (CNN) -2:

The model architecture employed is a Convolutional Neural Network (CNN) designed for face shape classification using the Keras deep learning framework. The dataset is organized into distinct classes, each representing a different face shape. The CNN comprises convolutional layers with Rectified Linear Unit (ReLU) activation, facilitating feature extraction from the input images. Max-pooling layers are incorporated for spatial reduction, contributing to the network's ability to capture essential features.

To prevent overfitting, dropout layers are strategically placed within the model, selectively deactivating some neurons during training. This regularization technique enhances the model's generalization capability. The final layer employs the softmax activation function, enabling the network to output probability distributions across multiple face shape classes. The model is compiled with the Adam optimizer, a popular choice for training deep neural networks, and utilizes categorical crossentropy as the loss function, particularly suitable for multi-class classification tasks.

During the training process, the model undergoes iterative updates through backpropagation, adjusting its weights to minimize the defined loss. The training progress is monitored and visualized, offering insights into the model's learning dynamics. Following training, the model is evaluated on a separate test set to assess its performance on unseen data. Additionally, the model is saved in an HDF5 format for future use.

Training Parameters:

- The model is trained with a batch size of 50.
- Training occurs over 30 epochs with 1000 steps per epoch (steps_per_epoch_val).

Model Summary

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 28, 28, 60)        1560

 conv2d_1 (Conv2D)           (None, 24, 24, 60)        90060

 max_pooling2d (MaxPooling2  (None, 12, 12, 60)        0
 D)

 conv2d_2 (Conv2D)           (None, 10, 10, 30)        16230

 conv2d_3 (Conv2D)           (None, 8, 8, 30)          8130

 max_pooling2d_1 (MaxPoolin  (None, 4, 4, 30)          0
 g2D)

 dropout (Dropout)           (None, 4, 4, 30)          0

 flatten (Flatten)           (None, 480)               0

 dense (Dense)               (None, 500)               240500

 dropout_1 (Dropout)         (None, 500)               0

 dense_1 (Dense)             (None, 5)                 2505

=================================================================
Total params: 358985 (1.37 MB)
Trainable params: 358985 (1.37 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

K-Nearest Neighbor (KNN):

- KNN is a Machine Learning Supervised Non-Parametric Model which calculates the distance classify the new individual data point.

Working of KNN model:

- KNN is a non-parametric algorithm, which means that it does not make any assumption on underlying data.
- KNN is a lazy learner algorithm, which does not learn or get trained from the training dataset, instead it stores the dataset and uses it at the time of each classification.
- It uses Distance parameter to classify the new individual data point. It calculates the distance between the new datapoint and the existing datapoints and selects k nearest neighbors which have least distance. And uses majority voting out of the selected neighbors classify the new datapoints.
- It uses Minkowski metrics to calculate distances like Manhattan distance, Euclidean distance, etc.

KNN Model For Face Shape Detection:

- We have the images data set which contains 4000 images to train and test the model. There are 5 labelled classes named ‘heart’, ‘oval’, ‘oblong’, ‘round’, ‘square’.
- For each image we extract the landmark features and its corresponding label.
- We split the landmark features dataset and labels dataset to train and test the model.
- Knn model first stores all the training dataset and does not get trained on the training dataset.
- When we test the model with new images, it extracts the landmark features of that image and compares these features with each landmark features of existing training images.
- It calculates the distance between the new image datapoints and the existing image’s datapoints and finds the k-nearest neighbors with least distance and

uses majority voting to classify the new image as one of the existing labelled classes.

Hyperparameter Tuning:

•Used different values for n_neighbors, weights, and power, to find the best fit parameters using Grid Search method.

Feature selection and engineering:

•Used Facial Landmark features to detect the shape of the face in the image.

Performance metrics:

•Used accuracy_score, Recall_score, Precision_score, F1_score to check the the performance of the model.

Visualization:

•Used Confusion Matrix, Classes distribution, ROC Curve to represent the dataset on which the model got trained.

KNN MODEL:

Accuracy: 37.466%

Precision: 36.925%

Recall: 37.466%

F1 Score: 36.455%

Hyperparameters:

- K_neighbhors: 15
- Power: 2
- Weights: Distance
- Confidence Interval: 0.2-0.6

Support Vector Machine (SVM):

Working:

- The SVM model generally finds the SVC by identifying the relationships between the data points as if they were at higher dimensions.
- The type of the SVM model used here is an SVC [Support Vector Classifier] model which uses the RBF[Radial Basis Function] kernel. This kernel calculates the relationship between the data points and the support vector classifiers for infinite dimensions.
- The radial basis function is as follows:

$$\mathrm{Exp}(-\gamma(a - b)^2)$$

- where a and b represent any two data points, γ represents the scale the squared distances between two points, thus it scales the influence of any two points have on each other
- In the model used there are two parameters C and γ which are tuned accordingly using the hyperparameter tuning and the best values of both are chosen as parameters.
- Here C is the parameter for soft margin cost function
- The main basis of this function is to linearly separate the data points at higher dimensions which are not linearly seperable at lower dimensions.

<u>Data Preprocessing:</u>

- The dataset is used for the purpose of training and testing the dataset contains 4000 images of celebrities with their faces, along with the labels (i.e., the face shapes which are Heart, Oblong, Oval, Round, Squared).
- The initial step for building an effective model is good data preprocessing, therefore we process these images such that our task of extracting features from them to classify the face shapes becomes easier.
- Firstly, we convert these RGB images into Grayscale for easier computation and equalize the images for luminance distribution of images. After the completion of this step the images are stored in a folder where the directory structure is same as the directory structure from where the input images are taken from.
- Other preprocessing techniques such as squaring, blurring, smoothening, edge enhancing, sharpening, cropping, flipping, expanding etc., can also be done.
- After the completion of the preprocessing step, we then move onto feature selection and feature engineering.

ENSEMBLE LEARNING:

- used ENSEMBLE LEARNING technique to combine the results of all the models and applied majority voting to the acquired results to get the best result that is having the highest confidence score.
- Machine Learning Models used: CNN,SNN,KNN,SVM,KMEANS models for the prediction of face shape
- All the above models have been trained with the data after applying PCA reduction to it.
- In Ensemble learning model, all these models are loaded. There exists a folder with images which are not there in the original dataset on which the models got trained, to validate the performance of the models.
- From the validation dataset, many features of each image have been extracted according to the requirement of each of the models.
- As per the requirement of each model, the extracted features have been sent into the model for the prediction of images in the validation dataset.
- Then combined the predictions of each image and stored it along with its confidence scores.
- At last, majority voting has been performed to get the best result having the highest votes, that is the predictions predicted by most of the models.
- In case there is a tie between 2 or more labels, we used average confidence score of each label. Then the label with the highest confidence score has been selected as the best result.
- This result is then sent into Augmented Reality function to recommend the FRAME that is best suitable to that person's face.

How They Were Implemented:

1. Data Preparation: A labeled dataset with images of rooms and corresponding people counts is crucial. Data augmentation techniques may be applied to increase the diversity of the dataset.
2. Feature Extraction: For CNN, feature extraction is automatic through the layers of the network. For SVM and linear regression, relevant features need to be extracted from the images, such as color histograms, edge features, or other image characteristics.
3. Training: Each algorithm is trained using the prepared dataset. The CNN is trained to learn hierarchical features, SVM learns the decision boundary, and linear regression learns the relationship between features and people count.
4. Evaluation: The models are evaluated on a separate test set to ensure they generalize well. Metrics such as accuracy, precision, recall, and F1 score can be used to assess the performance of each algorithm.
5. Integration: Once trained and evaluated, the models can be integrated into the people counting system. For real-time applications, considerations for efficiency and speed are crucial. The selected algorithm or a combination of them can then be deployed based on the specific requirements of the project.

Results

The successful implementation of the "Glasses" Spectacle Recommendation System has yielded promising outcomes across multiple dimensions, validating the effectiveness of our approach. The results encompass the performance of face shape prediction models, accuracy of facial feature measurements, and the overall success of the recommendation system.

Face Shape Prediction Model Performance

The four models employed for face shape prediction have demonstrated commendable accuracy in identifying diverse facial shapes. Soft voting, utilizing confidence scores as weights, further refines the predictions. The results, based on a comprehensive evaluation dataset, are as follows:

The soft voting mechanism has proven effective in enhancing the overall accuracy of face shape predictions, establishing the reliability of our ensemble model.

Conclusion

In conclusion, the results affirm the success of the "Glasses" Spectacle Recommendation System in providing accurate and personalized eyewear suggestions. The combination of advanced neural network models, soft voting, and facial science has resulted in a robust platform that addresses the unique needs of individuals, setting a new standard for personalized eyewear retail.

These results not only validate our approach but also lay the foundation for future enhancements and applications in the realm of personalized consumer experiences.