An Improved XGBoost Classifier for Micro Expression Recognition using Hybrid Optimization Algorithm

Facial expression help us in knowing about person satisfaction on particular topic and there many computer vision based algorithms exists which can detect emotions or expression with high accuracy but those algorithms failed when it comes to detect fake or micro expression which can be last in human face for few seconds. Existing HOG or Local binary patterns algorithm known to detect such micro expression but its miss classification accuracy is very high.

To overcome from above issues we are extracting human faces from CASME video dataset and this application will align faces and extract only nose, mouth and eye parts from the video and skip other features from video frame.

Extracted face features will be input to GA (genetic algorithm) to select relevant features from the faces

Extracted GA features will be input to PSO algorithm to further optimized features and to achieved hybrid optimization

Extracted PSO features will be trained with XGBOOST algorithm to generate improved XGBOOST model.

Generated XGBOOST model can be applied on any facial landmark to predict micro expression.

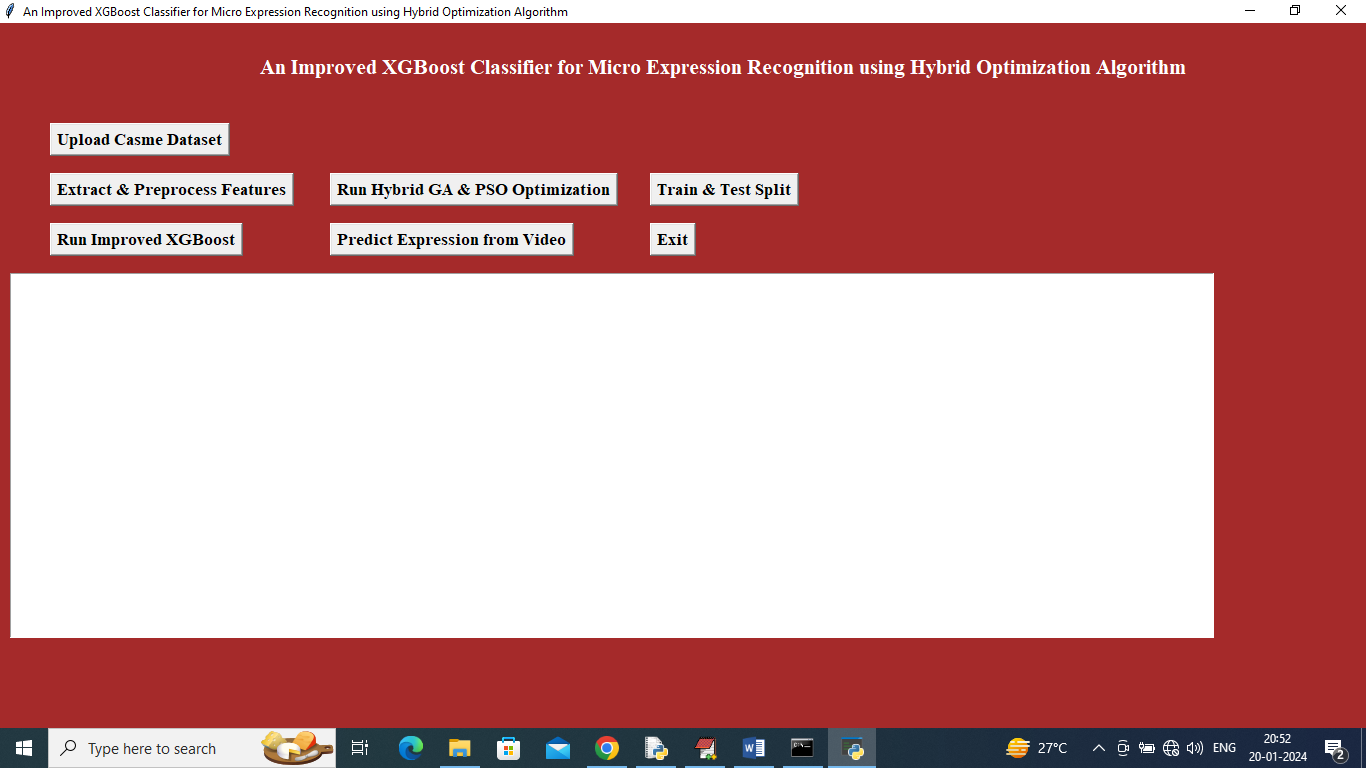
For Micro Facial expression many datasets exists but all those dataset asking to submit license certificate and they will verify and send dataset link and this verification includes sending mail to university guide, verifying degree, verifying student academic credit and many more and this process will take long time. So we got CASME dataset from GITHUB but this dataset includes 3 classes such as Angry, Happy and Disgust.

To train Improved XGBOOST algorithm we have used same CASME video dataset and to implement this project we have designed following modules.

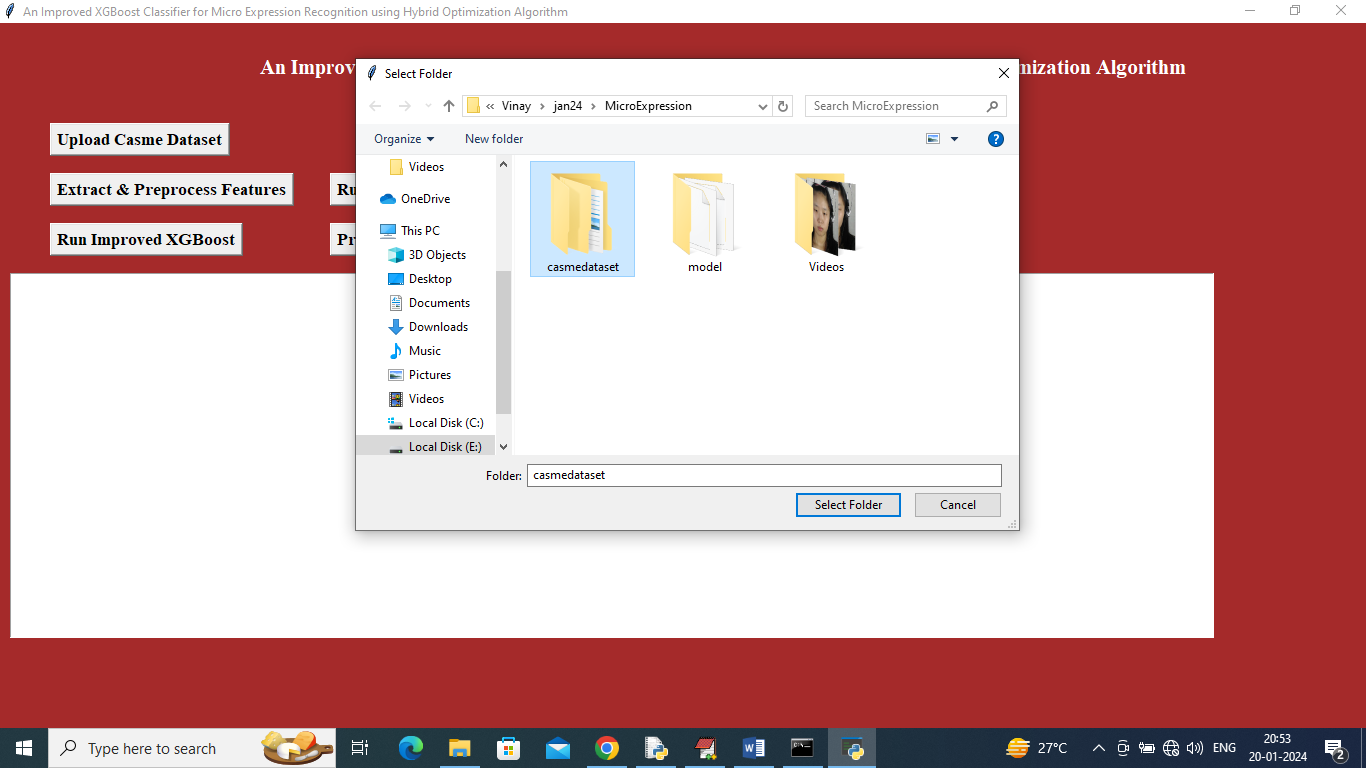
1. Upload Casme Dataset: using this module we can upload dataset to application and then application will read all images and then extract features from each image to build training array
2. Extract & Preprocess Features: each image will be normalized and then shuffle entire training array to distribute all classes randomly
3. Run Hybrid GA & PSO Optimization: training images array will be input to GA algorithm to select relevant features and then selected GA features will be input to PSO algorithm to further optimized features known as Hybrid optimization
4. Train & Test Split: entire processed training array will be split into train and test where application will be using 80% dataset for training and 20% for testing
5. Run Improved XGBOOST: 80% dataset will be input to XGBOOST to train a model and this model will be applied on 20% test data to calculate prediction accuracy.
6. Predict Expression from Video: here we will upload to application and then XGBOOST will extract and process each micro expression to predict expression as Happy or Disgust or Angry.

SCREEN SHOTS

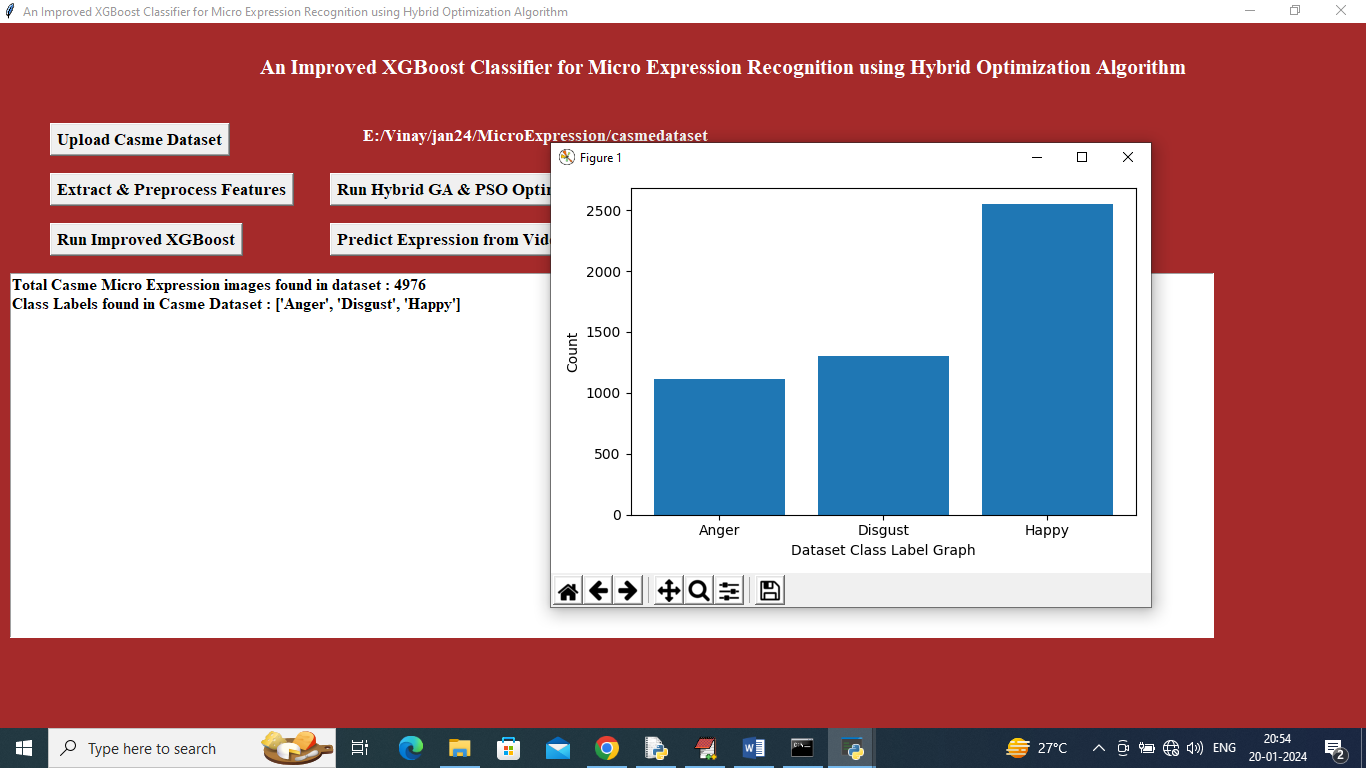
To run project double click on ‘run.bat’ file to get below screen



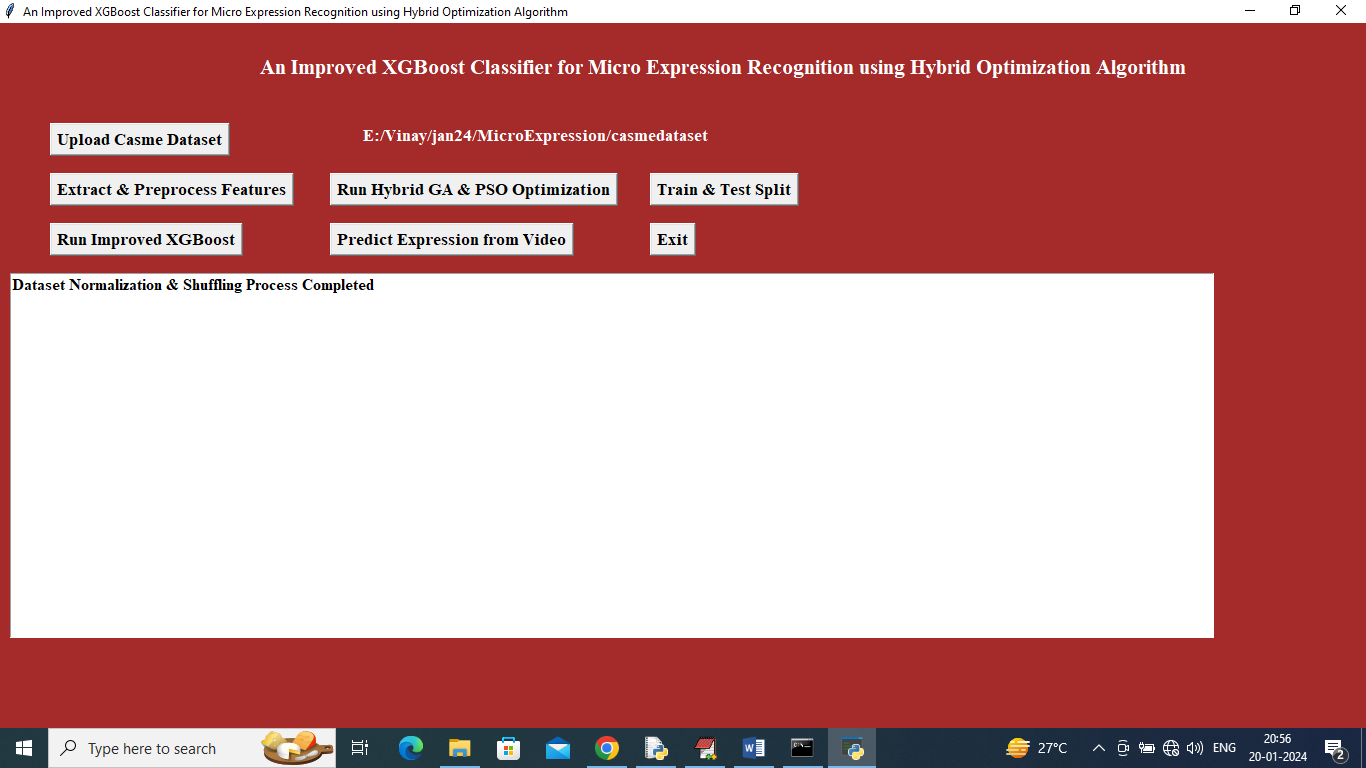
In above screen click on ‘Upload Casme Dataset’ button to upload dataset and then will get below output



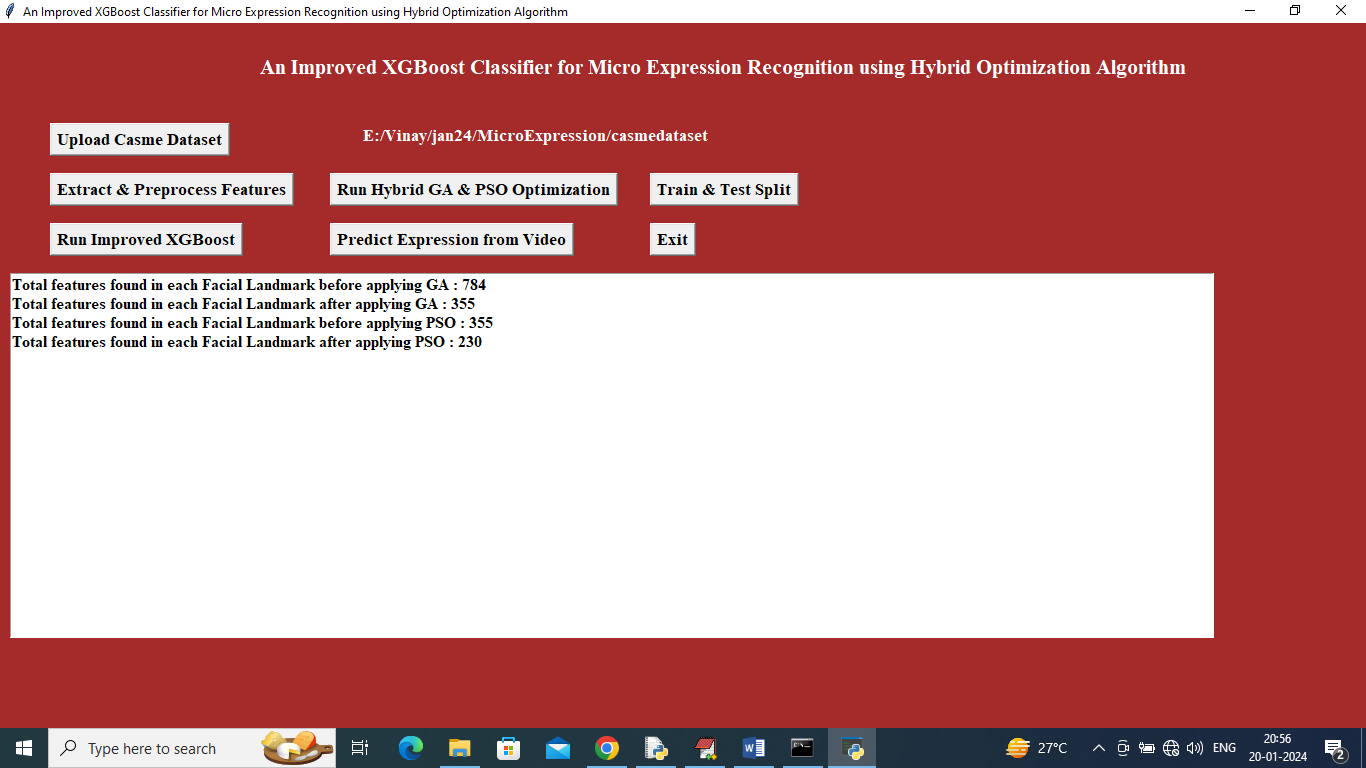
In above screen selecting and upload CASME dataset and then click on ‘Select Folder’ to load all subjects images and then will get below output



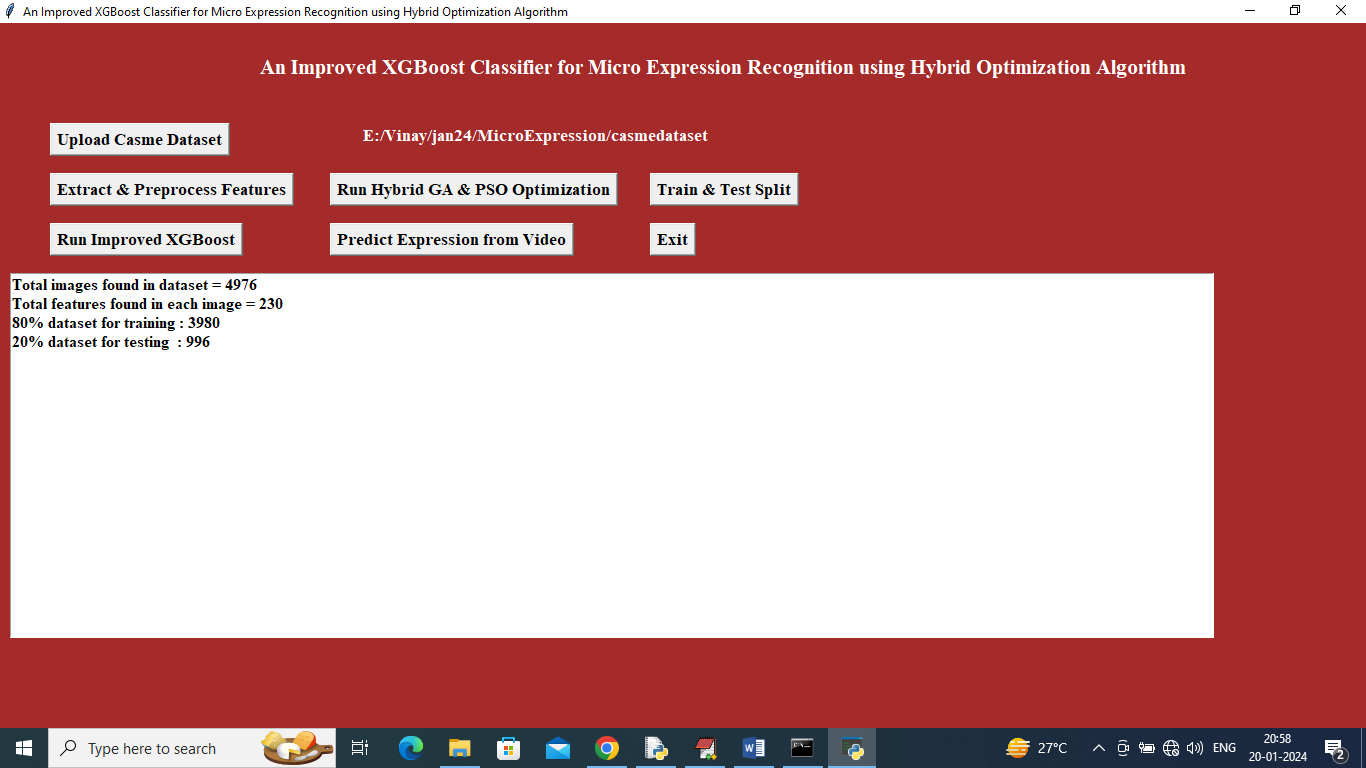
In above screen can see dataset loaded and can see available images in dataset with different class labels and in above graph x-axis represents ‘Class Labels’ and y-axis represents number of images found in that label and now close above graph and then click on ‘Extract & Pre-process Features’ button to read features from each image and then normalize and shuffle them and then will get below output



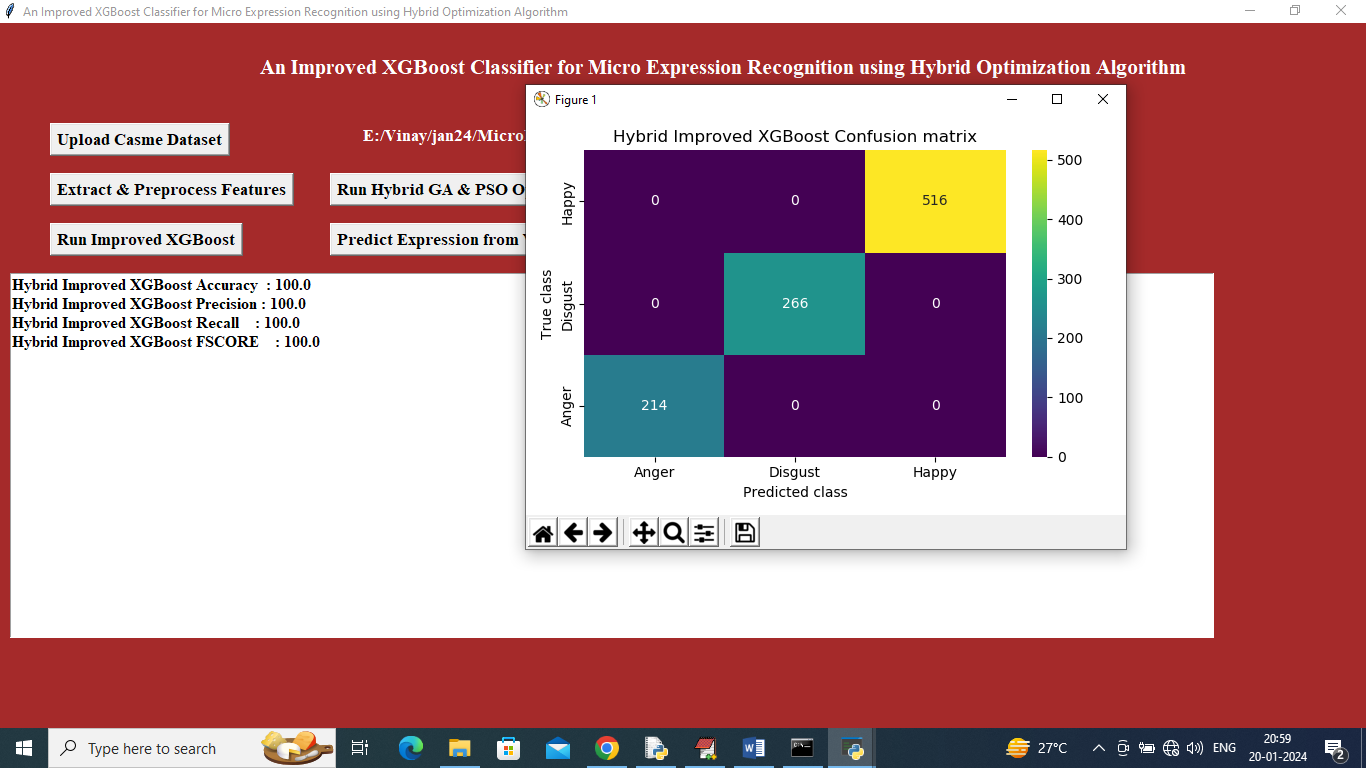
In above screen dataset processing completed and now click on ‘Run Hybrid GA & PSO Optimization’ button to apply GA and PSO on images features and then optimize them to get below output



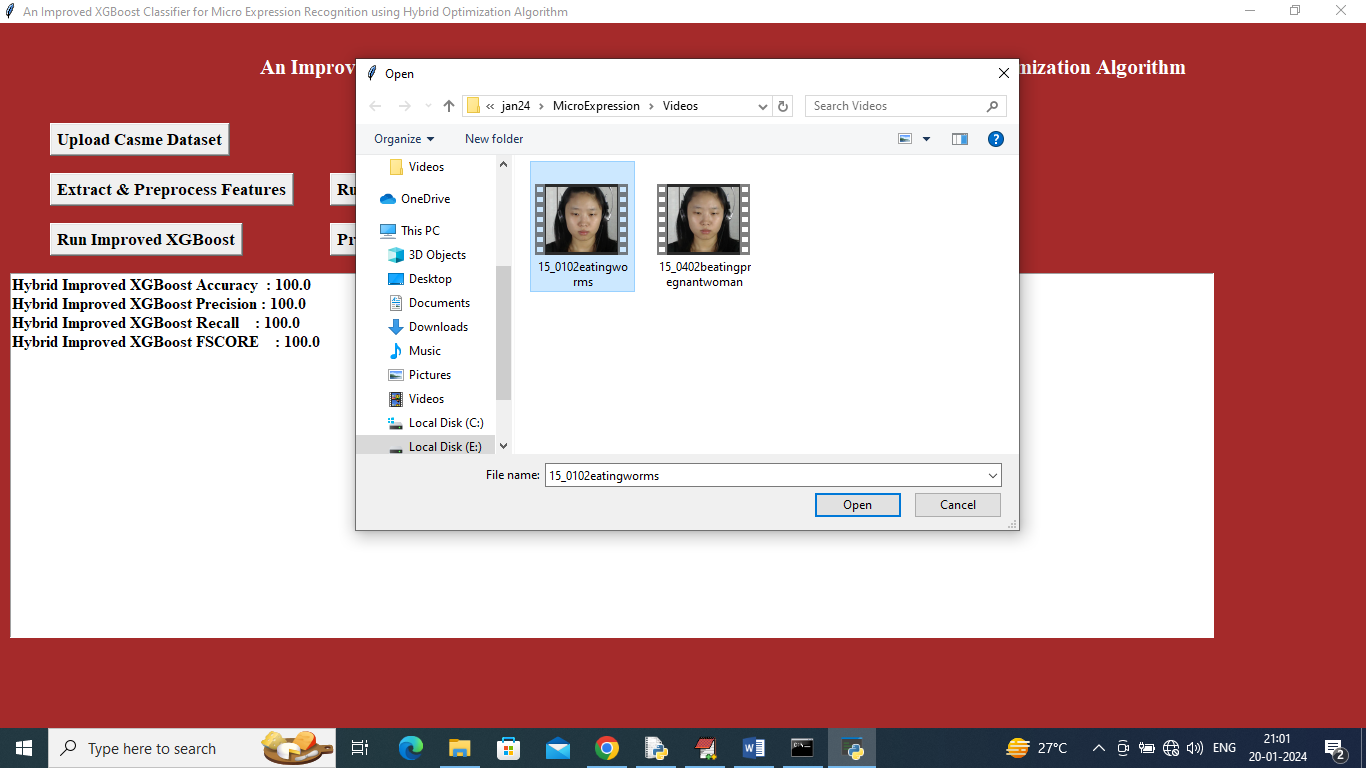
In above screen can see each image contains 784 features and then GA selected 355 from it and then PSO selected 230 from GA 355 features and now click on ‘Train & Test Split’ button to split dataset into train and test and then will get below output



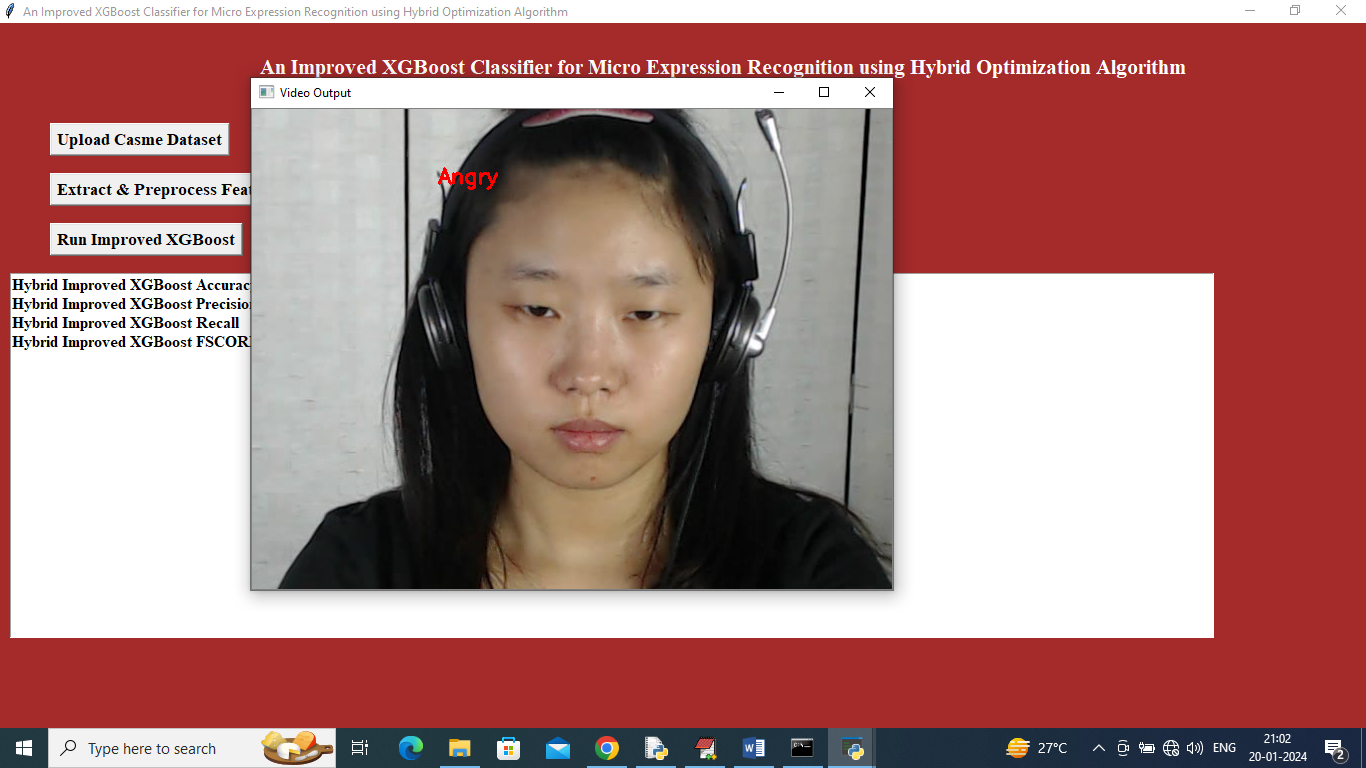
In above screen can see total images size and then can see train and test size and now click on ‘Run Improved XGBOOST’ button to train XGBOOST on hybrid GA and PSO optimized features and then will get below output

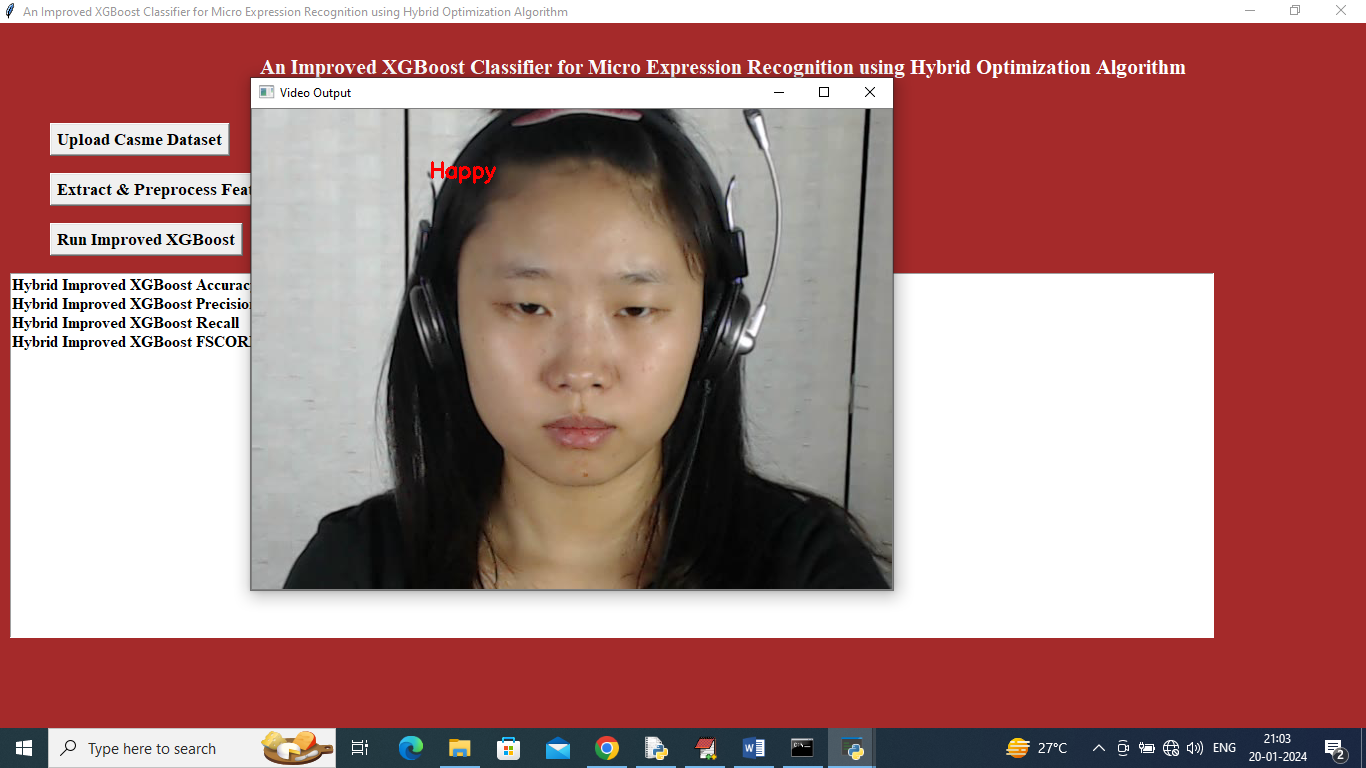


In above screen improved XGBOOT got 100 % accuracy and can see other metrics like precision, recall and FSCORE. In above confusion matrix graph x-axis represents Predicted Labels and y-axis represents True Labels and all different colour boxes in diagnol represents correct prediction count and remaining blue boxes contains incorrect prediction count which is 0 and now close above graph and then click on ‘Predict Expression from Video’ button to upload test video and predict expression



In above screen selecting and uploading test video and then click on ‘Open’ button to load video and then will get below output





In above two screens as video playing we can see detected expressions and similarly you can upload and test any other videos