Public figures image classification model

12th June 2022

Students

Sujatha Ramesh, Anoushka Jain, Mohammad Tanzil Alam

Matric numbers: 30006034, 30006028, 30006050

1 Executive Summary:

A facial recognition system is a technology capable of matching a human face from a digital image or a video frame against a database of faces, typically employed to authenticate users through ID verification services, works by pinpointing and measuring facial features from a given image. Many people are familiar with face recognition technology through the FaceID used to unlock iPhones (however, this is only one application of face recognition). Typically, facial recognition does not rely on a massive database of photos to determine an individual's identity — it simply identifies and recognizes one person as the sole owner of the device, while limiting access to others.

In 2014, Facebook announced its DeepFace program, which can determine whether two photographed faces belong to the same person, with an accuracy rate of 97.25%. When taking the same test, humans answer correctly in 97.53% of cases, or just 0.28% better than the Facebook program.

In June 2015, Google went one better with FaceNet. On the widely used Labelled Faces in the Wild (LFW) dataset, FaceNet achieved a new record accuracy of 99.63% (0.9963 ± 0.0009).

Using an artificial neural network and a new algorithm, the company from Mountain View has managed to link a face to its owner with almost perfect results. This technology is incorporated into Google Photos and used to sort pictures and automatically tag them based on the people recognized. Proving its importance in the biometrics landscape, it was quickly followed by the online release of an unofficial open-source version known as OpenFace. A study published in June 2019 estimates that by 2024, the global facial recognition market would generate $7billion of revenue, supported by a compound annual growth rate (CAGR) of 16% over 2019-2024.

In this project, face recognition from the images is identified using famous technique and briefly analysed how these images are used to train the classification models with its class label.

2 Introduction:

Today, how to identify people accurately and effectively has always been an interesting topic, both in research and in industry. With the rapid development of artificial intelligence in recent years, facial recognition gains more and more attention. Compared with the traditional card recognition, fingerprint recognition and iris recognition, face recognition has many advantages, including but not limited to non-contact, high concurrency, and user friendly. It has high potential to be used in government, public facilities, security, e-commerce, retailing, education, and many other fields. Recent development in machine learning, with the advancements in camera and computer vision technologies, have accelerated the design, development, testing, deployment, and operation of facial recognition systems.

Beyond Facial recognition, the “Facial recognition” is sometimes described as encompassing facial characterization. This is also called facial analysis - systems, which detect facial attributes in an image, and then sort the faces by categories such as hair colour, gender, or race. We do not consider such systems to be a part of facial recognition systems because they are not used to predict the identity of a person.

Some of use cases in real life – facial recognition:

a. Security – Law enforcement

Facial recognition is used when issuing identity documents and, most often, combined with other biometric technologies such as fingerprints (preventing ID fraud and identity theft).

Face match is used at border checks to compare the portrait on a digitized biometric passport with the holder's face.

Face recognition CCTV can be used to enable police to track and identify past criminals suspected of perpetrating an additional infraction. Police can also take preventive actions. By using an image of a known criminal from a video or an external picture (or a database), operators can detect matches in live video and react before it’s too late.

Facial recognition can be used to find missing persons and victims of human trafficking. Suppose missing individuals are added to a database. In that case, law enforcement can be alerted as soon as they are recognized by face recognition — whether it is in an airport, retail store, or other public space.

Facial recognition is used to identify when known shoplifters, organized retail criminals, or people with a history of fraud enter stores. Photographs of individuals can be matched against large databases of criminals so that loss prevention and retail security professionals can be notified when shoppers who potentially represent a threat enter the store.

b. Banking

Biometric online banking is another benefit of face recognition. Instead of using one-time passwords, customers can authorize transactions by looking at their smartphone or computer. With facial recognition, there are no passwords for hackers to compromise. Face recognition could make debit cards and signatures a thing of the past.

c. Marketing & advertising

Marketers have used facial recognition to enhance consumer experiences. For example, frozen pizza brand DiGiorno used facial recognition for a 2017 marketing campaign where it analysed the expressions of people at DiGiorno-themed parties to gauge people’s emotional reactions to pizza. Media companies also use facial recognition to test audience reaction to movie trailers, characters in TV pilots, and optimal placement of TV promotions. Billboards that incorporate face recognition technology – such as London’s Piccadilly Circus – means brands can trigger tailored advertisements.

d. Healthcare

Hospitals use facial recognition to help with patient care. Healthcare providers are testing the use of facial recognition to access patient records, streamline patient registration, detect emotion, and pain in patients, and even help to identify specific genetic diseases. AiCure has developed an app that uses facial recognition to ensure that people take their medication as prescribed.

e. Monitoring Casino floors

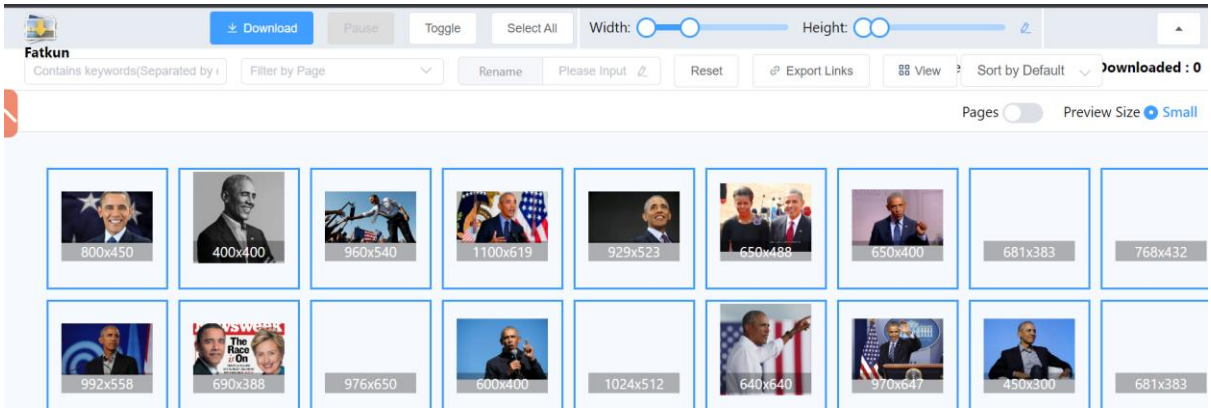
Facial recognition can help gambling companies protect their customers to a higher degree. Monitoring those entering and moving around gambling areas is difficult for human staff, especially in large, crowded spaces Facial recognition technology enables companies to identify those who are registered as gambling addicts and keeps a record of their play so staff can advise when it is time to stop.

3 Aim:

In this project, face recognition from the images is identified using OpenCV and Haar cascade and how these images are used to train the classification models to find the accuracy with its class labels.

4 Data background:

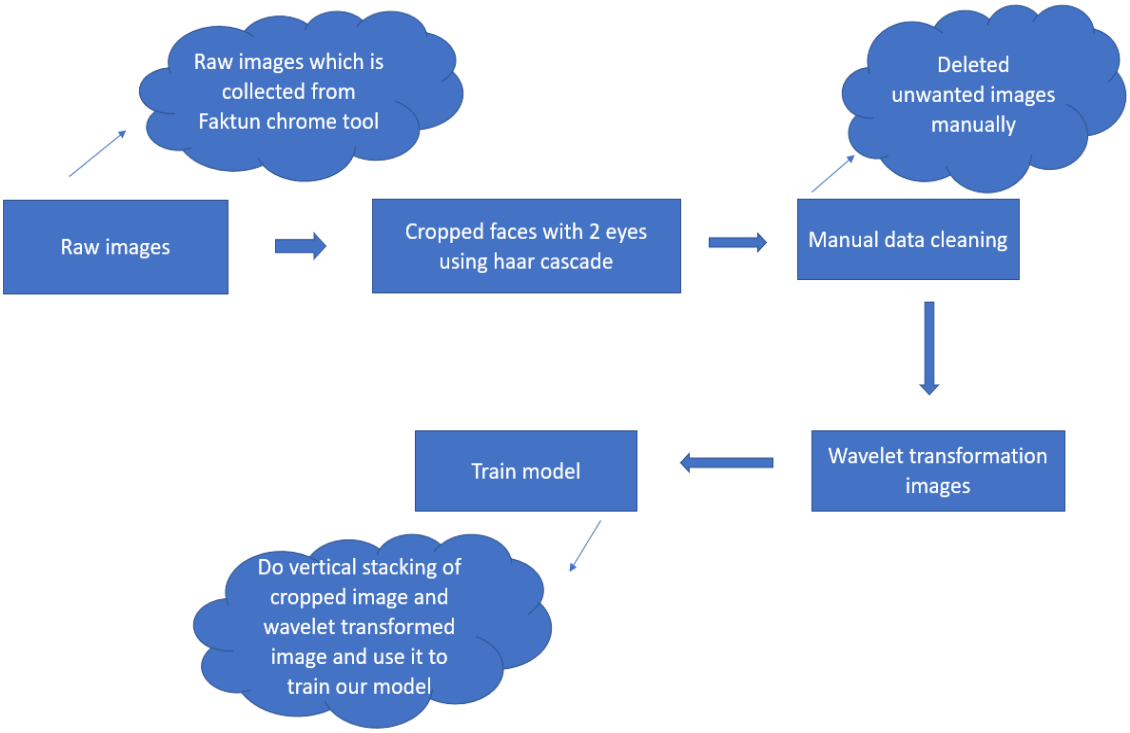
The image dataset is collected from Faktun chrome tool and limited to six public figures namely Barack Obama, Bill gates, Dalai lama, Lionel Messi, Roger Federer and Indira Nooyi with the test images of Jeff Bezos. And some images dataset downloaded from https://www.kaggle.com/code/tanishgupta26/famous-personalities-image-classifier/data



5 Methodology:

5.1 Understanding the project:

This project is about how the images are used to train SVM classifier, Logistic Regression and Random Forest classifier algorithms. To do this supervised machine learning, truth data is needed. This truth data is considered as the images X or dependent variable and the class label Y or target variable which tells whether it is the correct person. i.e., whether it is Bill Gates or Barack Obama. To achieve this goal, we need lot of images to train our classifier. As discussed before, these images are collected from Faktun chrome tool and for this project our data is limited to 6 public figures to train our model. The project plan is,



5.2 Data Cleaning:

The transformation of the raw dataset into a comprehensible format is known as data preprocessing. Data preprocessing is a critical step in data mining that improves data efficiency. Images data might have some issues when it is collected directly from google. To detect the person in the photo, the face is targeted to recognise. To detect the face and two eyes, OpenCV image processing library is used and for specific detection the famous technique called Haar Cascade is used. The images are cropped using Haar cascade and saved as an individual folder.

In this project, data cleaning is done on cropped images. The pictures are deleted manually if the two eyes and face are not clearly visible.

For instance,

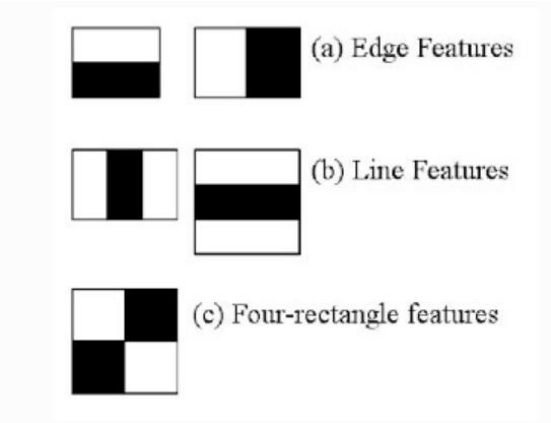
In the below images, two eyes are not properly visible in the picture of Obama and Messi.

Some images in the cropped folder are not clear and which we do not need for our model to train. This image data is cleaned manually by deleting such pictures after cropping them using Haar cascade.

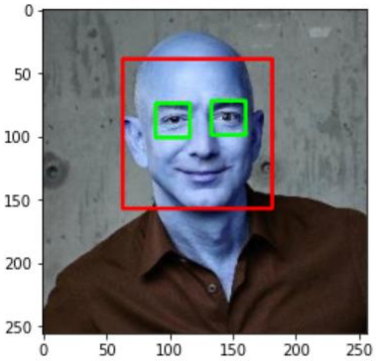
Haar Cascade:

It is an Object Detection Algorithm used to identify faces or face detection in an image or a real time video. The algorithm uses edge or line detection. Initially, the algorithm needs a lot of positive images (images of faces) and negative images (images without faces) to train the classifier. Then we need to extract features from it. For this, haar features shown in below image are used. They are just like our convolutional kernel. Each feature is a single value

obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle.



OpenCV act as a trainer as well as detector. To train the classifier for any object like car, planes etc, can use OpenCV to create one. If faces are found, it returns the positions of detected faces as Rect (x, y, w, h). Once it locates the face, should create a ROI for the face and apply eye detection on this ROI (since eyes are always on the face).



Computer Vision:

Computer Vision is research in the AI discipline that enables computers to see and understand the world around them. It allows for computers to recognize objects, scenes, faces and much more from digital images taken by cameras (Video or Picture). It can be trained on huge datasets. It facilitates the understanding & allows computers to react to the world in real time.

5.3 Feature Engineering:

Feature engineering is the process of transforming raw data into features that better represent the underlying problem to the predictive models, resulting in improved model accuracy on unseen data. In this project, feature engineering is done on the cropped images by the technique called wavelet transformation to extract the important features of the image which is useful for the classification model.

Wavelet Transformation:

Wavelet transform will extract the feature or temporal information in a form of signal from the image. This wavelet transform image helps machine to understand each and every feature in the image. It transforms the original image like,

6 Train our model:

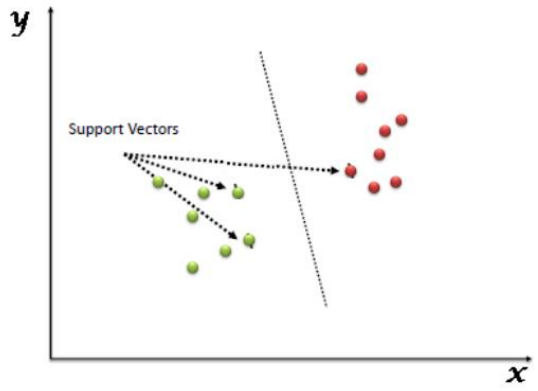
To train our model, the classification models like Support Vector Machine (SVM), Logistic regression and Random Forest classifier are used. For the initial step, the model is fit using validation set.

6.1 Support Vector Machine (SVM):

Support Vector Machine can be used for both classification and regression problems. Mostly it is used in non-linear separation problems. In brief, the kernel does some sophisticated data transformations before determining how to split the data based on the labels or outputs specified. Handwriting recognition, intrusion detection, face identification, email classification, and web page generation all use SVMs. SVMs are used in machine learning for this purpose. It is capable of handling classification and regression on both linear and non-linear data.

Another reason we employ SVMs is that they may discover intricate associations between business data without requiring you to perform numerous adjustments on your own. It's an excellent choice when working with micro datasets with tens to hundreds or even thousands of features. Because of their ability to handle small, complex information, they often produce more accurate results than other algorithms

In SVM, each data item is plotted as a point in n-dimensional space and the classification is performed by finding the hyper-plane that distinguishes the two classes.



Importing necessary libraries,

```
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
from sklearn.pipeline import Pipeline
from sklearn.metrics import classification_report
```

Furthermore, the classification report displays the detailed statistics of the performance of the model.

	precision	recall	f1-score	support
0	0.82	0.91	0.86	35
1	1.00	0.43	0.60	7
2	0.71	0.71	0.71	7
3	0.89	0.96	0.92	25
4	0.90	1.00	0.95	9
5	1.00	0.25	0.40	4
accuracy			0.85	87
macro avg	0.89	0.71	0.74	87
weighted avg	0.86	0.85	0.84	87

From the above classification report, 0 to 5 (in the left) is the class label for our 6 selected public figures (can refer from the below figure) and it displays with its precision and recall score. In the report, precision value for 1 and 5 class is 1.00. It means the model's precision is correct.

```
class_dict = {}
count = 0
for Publicfigures_name in Publicfigures_file_names_dict.keys():
    class_dict[Publicfigures_name] = count
    count = count + 1
class_dict
```

```
{'Barack_Obama': 0,
 'Bill_Gates': 1,
 'Dalai_Lama': 2,
 'Indira_Nooyi': 3,
 'lionel_messi': 4,
 'roger_federer': 5}
```

F1 score:

F1 score is used in classification problems with more than two classes (Multiclass classification). The F1 score is the harmonic mean of the precision and recall. The highest possible value of an F-score is 1.0, indicating perfect precision and recall, and the lowest possible value is 0, if either the precision or the recall is zero. Better the F1 score indicates that you have a low number of false positives and false negatives, indicating that you are correctly identifying serious threats and are not bothered by false alarms. When the F1 score is 1, the model is deemed ideal, but when it is 0, the model is considered a complete failure. Where dealing with real-world classification problems and when there is an imbalanced class distribution, the F1 Score is a superior statistic to use.

As discussed before, from the classification report, precision value for 1 and 5 class is 1.00. It means the model's precision is correct.

GridSearchCV:

Grid search technique is used to find the optimal hyperparameters to increase the model performance. The best hyperparameters values gives the perfect prediction of our model. Grid Search calculates the performance for each combination of all the supplied hyperparameters and their values, and then chooses the optimum value for the hyperparameters. Hyperparameters are variables that the user often specifies when developing the Machine Learning model. Thus, hyperparameters are specified before parameters, or we can say that hyperparameters are utilized to evaluate optimal model parameters. Most importantly, hyperparameters is that their values are determined by the user creating the model.

In this project, the GridSearchCV is run for a different model SVM, Random Forest and Logistic Regression with a different parameter. Hence, the validation set is used to decide the performance of the model by hypertuning it.

```
model_params = {
    'svm': {
        'model': svm.SVC(gamma='auto',probability=True),
        'params' : {
            'svc__C': [1,10,100,1000],
            'svc__kernel': ['rbf','linear']
        }
    },
    'random_forest': {
        'model': RandomForestClassifier(),
        'params' : {
            'randomforestclassifier__n_estimators': [1,5,10]
        }
    },
    'logistic_regression' : {
        'model': LogisticRegression(solver='liblinear',multi_class='auto'),
        'params': {
            'logisticregression__C': [1,5,10]
        }
    }
}
```

```
scores = []
best_estimators = {}
import pandas as pd
for algo, mp in model_params.items(): # iterate through the models above
    pipe = make_pipeline(StandardScaler(), mp['model']) # pipeline is first
    clf =  GridSearchCV(pipe, mp['params'], cv=5, return_train_score=False)
    clf.fit(X_train, y_train)
    scores.append({
        'model': algo,
        'best_score': clf.best_score_,
        'best_params': clf.best_params_
    })
    best_estimators[algo] = clf.best_estimator_

df = pd.DataFrame(scores,columns=['model','best_score','best_params'])
df
```

6.2 Random Forest Classifier:

A random forest classifier consists of larger number of decision trees that operates as an ensemble. It increases predicted accuracy and control over-fitting by fitting a number of decision tree classifiers on various sub-samples of the dataset. In our daily lives, we use the same random forest technique. Whether we are purchasing household equipment, choosing a vacation site, or going to see a movie, we consult our relatives for their thoughts and make our final decision based on them. So, the same thing applies to random forest. Suppose a person wants to purchase a coat after going through different shops he is not able to decide which shop he should consider making the purchase. As a result, he decides to consult with a variety of people, including his family members, cousins and friends. Finally, after talking with several people, he chooses the shop to purchase the coat. As a result, random forest works in the same way. It produces output based on most of the output.

6.3 Logistic Regression Classifier:

Logistic regression is a classification model uses a logistic function to model the dependent variable. Logistic regression has a probability between 0 and 1, and the interpretation of the weights differs from the interpretation of the weights in linear regression. The likelihood is no longer influenced linearly by the weights. The logistic function converts the weighted sum to a probability. As a result, we must reformulate the equation for interpretation so that only the linear term appears on the right side.

From our image classification model, best parameters and the best model got from validation set are given follows,

```
{'svm': Pipeline(steps=[('standardscaler', StandardScaler()),
                        ('svc',
                         SVC(C=1, gamma='auto', kernel='linear', probability=True))]),
 'random_forest': Pipeline(steps=[('standardscaler', StandardScaler()),
                        ('randomforestclassifier',
                         RandomForestClassifier(n_estimators=10))]),
 'logistic_regression': Pipeline(steps=[('standardscaler', StandardScaler()),
                        ('logisticregression',
                         LogisticRegression(C=1, solver='liblinear'))])}
```

	model	best_score	best_params
0	svm	0.854427	{'svc__C': 1, 'svc__kernel': 'linear'}
1	random_forest	0.678229	{'randomforestclassifier__n_estimators': 10}
2	logistic_regression	0.839115	{'logisticregression__C': 1}

So, the best model is SVM and Logistic regression with its parameters shown above. To check our model for the test data set using all 3-classification model, the accuracy score will be,

```
best_estimators['svm'].score(X_test,y_test) # done on test set
```

```
0.8275862068965517
```

```
best_estimators['random_forest'].score(X_test,y_test)
```

```
0.6551724137931034
```

```
best_estimators['logistic_regression'].score(X_test,y_test)
```

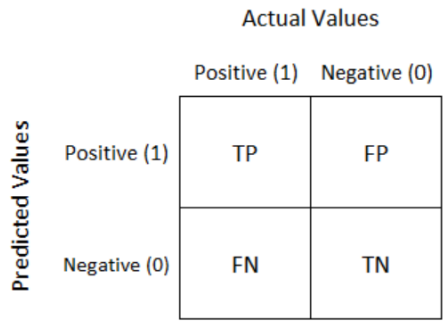
```
0.8275862068965517
```

From the validation set and the test set, found that SVM performs good. So, Support Vector Machine (SVM) is used to predict the images with its correct class label using confusion matrix.

Confusion Matrix:

Confusion matrix is used to describe the performance of a classification model on a set of test data for which the true values are known. To understand the model, get the output in probabilities, better effectiveness and better the performance of the model, confusion matrix is used. Each row of the matrix represents the instances in an actual class while each column represents the instances in a predicted class, or vice versa. The class is a variable in this table that has been created as a set of "actual" and "predicted" identical classes. Which is also known as the Error Matrix, is a performance assessment for machine learning

classification. In supervised learning, a table arrangement is used to visualise the algorithm's performance, whereas in unsupervised learning, it's referred to as a matching matrix.

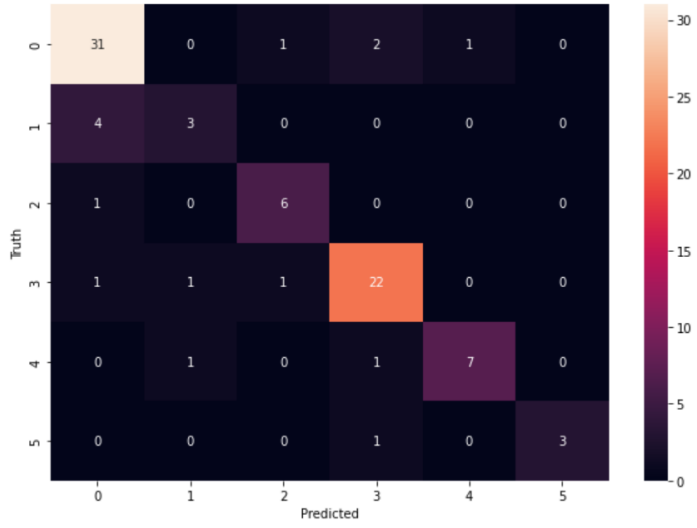


As discussed before, our model is good with SVM. So, the confusion matrix is used to describe our model.

```
best_clf = best_estimators['svm']
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, best_clf.predict(X_test))
cm
```

The class label is,

```
{'Barack_Obama': 0,
 'Bill_Gates': 1,
 'Dalai_Lama': 2,
 'Indira_Nooyi': 3,
 'lionel_messi': 4,
 'roger_federer': 5}
```



Interpretation - confusion matrix:

From the above visualization, 0 is class label of Barack Obama. The model is predicted 31 times as Barack Obama correctly so the model is good and 2 times as Indira Nooyi, 1 time as class label 2 and 4. i.e., Dalai Lama and Messi. Similarly, 3 is class label of Indira Nooyi and the model predicted 22 times as same and so on.

7 Conclusion:

SVM classifier and Logistic Regression classifier model performed good with the accuracy of 85% and 83% in recognising the face with the trained image. Here SVM is choose to predict the model using confusion matrix. As a future work, neural network can also be applied to train the model to get more accuracy.

8 Appendix:

https://github.com/Sujiremy86/SujathaRamesh

9 References:

https://www.researchgate.net/publication/246548380_Wavelet_Transform_based_Feature_Extraction_for_Face_Recognition

https://arxiv.org/ftp/arxiv/papers/1110/1110.1485.pdf

http://leap.ee.iisc.ac.in/sriram/teaching/MLSP_16/refs/W24-Wavelets.pdf

https://medium.com/@koushikc2000/2d-discrete-wavelet-transformation-and-its-applications-in-digital-image-processing-using-matlab-1f5c68672de3

https://towardsdatascience.com/the-wavelet-transform-e9cfa85d7b34

https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/

https://en.wikipedia.org/wiki/F-score

https://www.mygreatlearning.com/blog/gridsearchcv/#:~:text=GridSearchCV%20tries%20all%20the%20combinations,one%20with%20the%20best%20performance.

https://www.analyticsvidhya.com/blog/2021/06/tune-hyperparameters-with-gridsearchcv/

https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

https://www.freecodecamp.org/news/svm-machine-learning-tutorial-what-is-the-support-vector-machine-algorithm-explained-with-code-examples/

https://www.naukri.com/learning/articles/random-forest-algorithm-python-code/#real-life-analogy

https://www.analyticsvidhya.com/blog/2021/06/tune-hyperparameters-with-gridsearchcv/

https://christophm.github.io/interpretable-ml-book/logistic.html

https://insideaiml.com/blog/What-is-Confusion-Matrix%3F-1127

https://aiforkids.in/qa/845/deduce-the-formula-of-f1-score-what-the-need-its-formulation

https://github.com/opencv/opencv/tree/master/data/haarcascades

https://towardsdatascience.com/viola-jones-algorithm-and-haar-cascade-classifier-ee3bfb19f7d8

https://opencv24-python-tutorials.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html