International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

Face Recognition Techniques: Occluded Face Recognition

Satya Shrivastava*, Prof. M.S. Prasad, Dr.Shivani Verma
Amity Institute of Space Science and Technology
Amity University, AUUP

Submitted: 25-01-2022 Revised: 05-02-2022 Accepted: 08-02-2022

ABSTRACT - Face Recognition is a technique used to detect the faces of individuals. Though other methods of identification can be more accurate, face recognition has always remained a widely used technique because it is simpler, easily implementableand has a wide practical use. The purpose of this paper is to thoroughly examine and evaluate some of the most widely used algorithms in Face Recognition systems. The paper discusses the difficulties that might arise when choosing proper algorithms and includes a case study on occluded face recognition.

Index Terms - Face Recognition, Face Detection, Feature Extraction, classification, Occluded face recognition

I. INTRODUCTION

At its most basic level, facial recognition comprises recognizing or verifying someone based on the characteristics and traits of their face. Humans are extremely good at detecting known faces, but not so good at distinguishing a large number of new ones. Computers are able to overcome human limitations due to their near-infinite memory and calculating speed. Developing an automated system that can recognize faces as well as a person can is a significant undertaking. It's one of the most important uses of image analysis. Figure 1 shows the different applications of Face Recognition System. Biometrics, access control, and information security are just a few of the uses for face recognition.

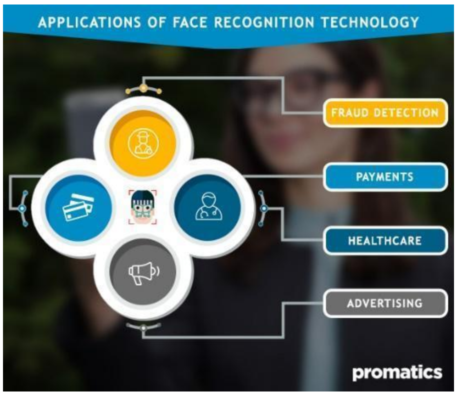

Figure 1: Applications of Face Recognition System

The remainder of this paper is arranged in the following manner. The general face recognition system is discussed in Section II, and each of its components is explained in depth. It also provides a comparative analysis of several designs for each component. Face recognition techniques are discussed in Section III. The Occluded Face Detection issue and associated studies are discussed in Section IV. It also includes a YOLO Algorithm-based Occluded Face Detection Model. The paper concludes with Section V.

II. FACE RECOGNITION

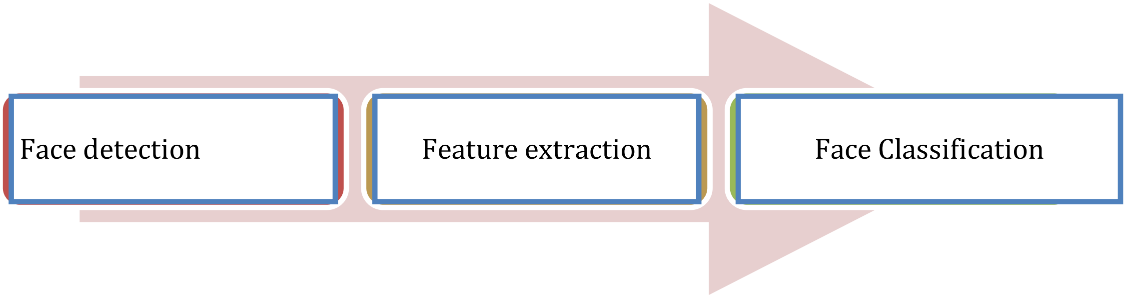

Figure 2: Generic Face Recognition System

Figure 2 shows a generic face recognition system.A face recognition system's input is always a video or picture stream.The method of detecting faces from images is known as face detection. The next stage is to extract relevant facial features from the data, which is called feature extraction.These features could include specific face regions, variations, angles, or measurements that are either human relevant (e.g., the space between the eyes) or not. Finally, the system recognizes the face. In an identification task, the system would report an identify from a database. This step includes the usage of a comparison technique, a classification algorithm, and an accuracy measure.

2.1 Face Detection

Face identification algorithms must deal with some issues that are common in images acquired in uncontrolled contexts, such as surveillance video systems. These issues can be attributed to a number of factors, including:

1. Pose Variation-The best-case scenario would be to use just frontal photographs for face detection. However, in uncontrolled circumstances, this is impossible. The subject's motions or the camera's viewpoint might cause pose change.
2. Feature occlusion- The existence of beards, spectacles, or hats introduces a lot of variation. Artifacts and other people's faces can also partly hide faces. Various obstructed face detection approaches are shown in Figure 4.
3. Facial expressions- Different facial gestures cause facial characteristics to change dramatically.
4. Imaging conditions-Different cameras and ambient circumstances can impact the quality of a picture, which can affect the look of a face.

Figure 3 shows the block diagram of the face detection process.

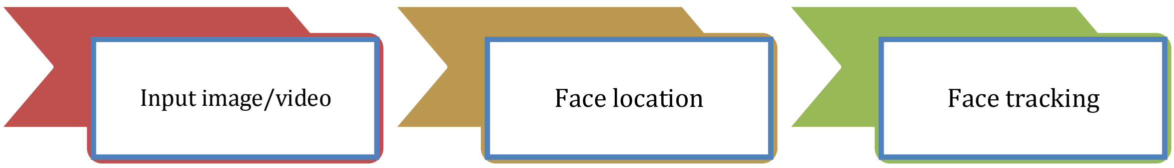

Figure 3: Face detection process

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

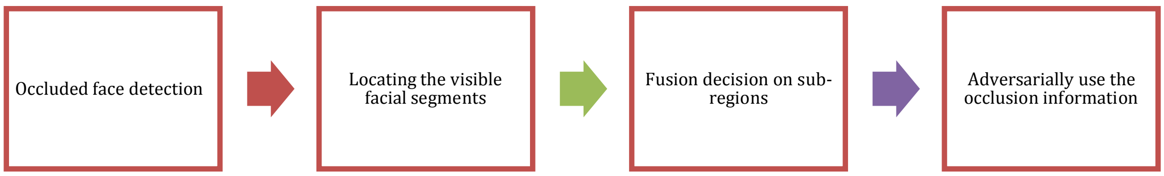

Figure 4:Occluded face detection methods

2.1.1 Face detection Methodologies

Detection can be done in a variety of ways, depending on the situation. It might be in a controlled environment, which is the simplest instance. Photographs are shot with certain lighting, backgrounds, and other factors in mind. Face detection may be done using simple edge detection algorithms [1]. Color pictures [2][3] or images in motion [4] can also be used for detection. Yan, Kriegman, and Ahuja et al. provided a categorization [32] in which approaches were separated into four groups as indicated in figure 5:

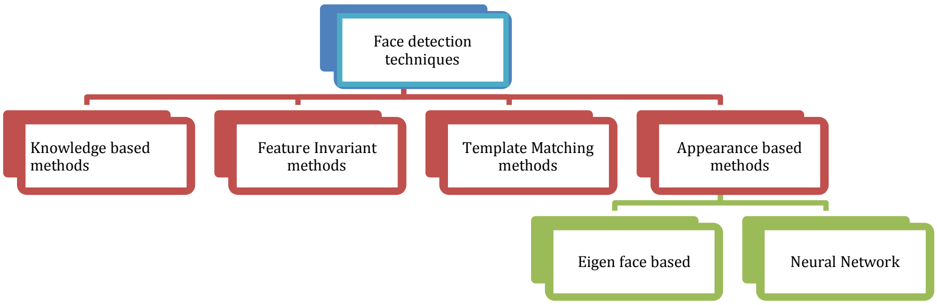

Figure 5: Categories of face detection techniques

1. Knowledge-based methods: These are rule-based methods.Some recent studies employ multiple color models. RGB and HSV, for example, have been successfully combined [5]. In [5], the authors have chosen the following parameters:

$$0.4 \leq r \leq 0.6,\ 0.22 \leq g \leq 0.33,\ r > g > (1 - r)/2 \quad (1)$$

$$0 \leq H \leq 0.2,\ 0.3 \leq S \leq 0.10,\ 0.22 \leq V \leq 0.8 \quad (2)$$

Skin tone can fluctuate considerably as the light conditions change. As a consequence, skin color detection is combined with other approaches like local symmetry or structure and geometry.

2. Feature-invariant methods: Algorithms that strive to uncover invariant aspects of a face that don't vary with angle or position.
3. Template matching algorithms compare input photos to previously recorded face or feature patterns. Methods for matching templates attempt to define a face as a function. These techniques are only applicable to unoccluded frontal faces.
4. Appearance-based methods are a type of template matching method that learns its pattern database from a set of training photos. The following are the most important methods of the appearance-based method:

a. Eigenface-based-PCA was used by Sirovich and Kirby [6, 7] to effectively represent faces. The goal of this approach is to represent a face as a coordinate system. The eigen pictures were the vectors that made up this coordinate system. [8].
b. Neural Networks-Neural networks have successfully solved many pattern recognition problems such as object recognition, character recognition, and so on. These systems are frequently used in face detection in a variety of ways. Some of the first researchers used neural networks to discover face and non-face patterns. [9]

2.2 Feature Extraction

Feature extraction may be preceded by Dimensionality reduction techniques as shown in Figure 6 [34].

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

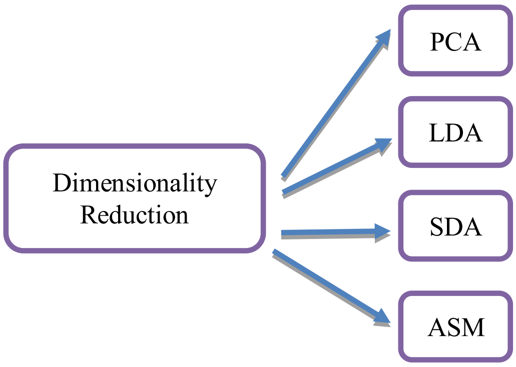

Figure 6: Dimensionality Reduction techniques

2.2.1 Dimensionality Reduction

PCA, or Principal Component Analysis, is a dimensionality-reduction approach for reducing the dimensionality of large data sets by transforming a large collection of variables into a smaller one that retains the majority of the information in the original set. Despite the fact that PCA reduces dimensions, when dealing with multi-class data, it's vital to reduce dimensions in such a way that inter-class separation is preserved. LDA is a method that can be employed in this situation.

A dimensionality reduction approach is Linear Discriminant Analysis (LDA) [26]. Dimensionality reduction approaches, as the name indicates, lower the number of dimensions (i.e., variables) in a dataset while preserving as much information as feasible. In reality, the covariance matrix of each class may not be properly computed if there are insufficient training samples. Deng Cai et al. (2007) presented Semi-supervised Discriminant Analysis (SDA) [24] as a solution to this problem.

SDA is a linear dimensionality reduction algorithm. It can make efficient use of both labeled and unlabeled data points. Thedata points with label are used to maximize the discriminating power, while the unlabeled data points are used to maximize the locality preserving power. LDA.

SDA is a dimensionality reduction approach that is linear. It can use both labelled and unlabeled data points effectively. The discriminating power is maximized by using labelled data points, while the locality preserving power is maximized by use of the unlabeled data points.

The Active Shape Model (ASM) [28] model is trained from manually drawn contours (surfaces in 3D) in training images. Principal Component Analysis is used by the ASM model to determine the key variations in the training data, which enables the model to automatically recognize if a contour is a possible/good object contour. The ASM modes also include matrices that describe the texture of lines perpendicular to the control point, which are utilized to rectify positions during the search process [27].

Figure 7 gives different schemes for feature extraction methods in Occlusion. The aim is to extract features that are less affected by occlusions while preserving the discriminative capability.

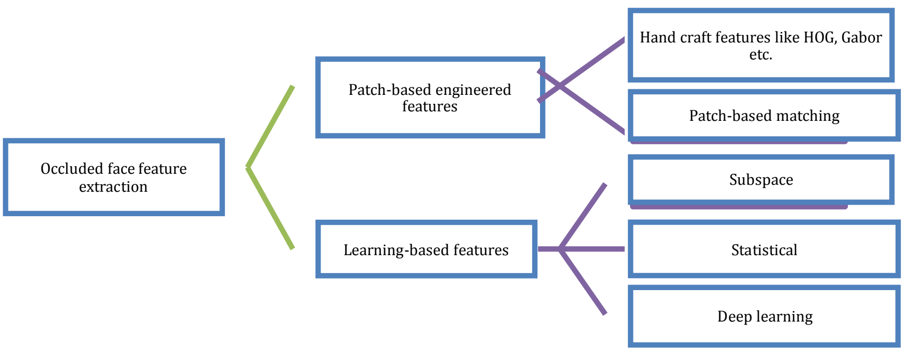

Figure 7: Feature extraction methods in occlusion

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

2.2.2 Feature selection

The feature selection algorithm's purpose is to choose a subset of the extracted features that results in the least amount of classification error. There has been a resurgence of interest in adopting feature selection methods as a result of the following circumstances:

1) Multi sensor fusion
2) Integration of multiple data models

The simplest approach to the feature selection problem would be to:
1) examine all feasible subsets of size m, and
2) choose the subset with the most relevant criteria function value.

By obtaining bounds on the eventual criterion value using intermediate results, this approach avoids an exhaustive search. Because the number of feature subset evaluations for large feature sizes can quickly become prohibitive, a variety of suboptimal selection approaches are proposed, which effectively tradeoff the optimality of the chosen subset for computing speed.
The majority of well-known feature selection strategies proposed in the literature [19] are listed in Table 1.

Table 1: Feature Selection methods

Method	Definition	Comments
Exhaustive Search	Evaluate all possible subsets of features	Optimal, complex. Guaranteed to find optimal subset, not feasible for even moderately large value of m and d
Branch and bound	Use branch and bound	Guaranteed to find an optimal subset provided the criterion function satisfies the monotonicity property; the worst-case complexity of this algorithm is exponential.
Best individual feature	Evaluate and select features individually.	Not very effective. Simple algorithm
Sequential Forward Selection (SFS)	Evaluate growing feature sets (starts with the best feature).	Retained features can't be discarded. Faster than SBS.
Sequential Backward Selection (SBS)	Evaluate shrinking feature sets (starts with all the features)	Deleted features can't be reevaluated Requires more computation than SFS
"Plus l-take away r" selection	First do SFS then SBS.	Avoids the problem of feature subset "nesting" encountered in SFS

		and SBS Must choose l and r values. (l>r)
Sequential Forward Floating Search (SFFS)) and Sequential Backward Floating Search and dynamic update. (SBFS)	Like Plus l-take away r" selection, but l and r values automatically pick and dynamically update	Close to optimal. Affordable computational cost.

The primary two strategies in Table 1 are the only ones that guarantee an optimal subset. Because the simplest pair of features does not have to contain the smallest single feature, all other techniques are suboptimal [17]. In general, exceptional, little feature sets aren't often included in good, larger feature sets. As a result, the simple strategy of selecting only the most basic individual features may fail miserably. However, selecting a few particular good features as a first step in reducing very big feature sets might still be advantageous (e.g., many features). More advanced methods that evaluate feature interdependence must be used to make an additional selection. These work by analysing either rising (ahead selection) or diminishing (reverse selection) feature sets (backward selection). SFS (SBS) is a simple sequential approach that add (delete) one feature at a time. The "Plus l - deduct r" strategy, and hence the SFFS and SBFS [21] Sequential Floating Search methods, are more complex procedures. These approaches go backwards in time as long as they identify improvements over earlier feature sets of comparable size. In almost every big feature selection job, these techniques outperform straight sequential searches.

Several feature selection techniques have been compared in terms of classification error and run time by Ferri et al. [18] and Jain and Zongker [19]. Somol et al. [22] created an adaptive version of the SFFS algorithm that has been demonstrated to perform better.

When a multilayer feed-forward network is employed for pattern classification, the node-pruning approach concurrently determines both the optimal feature subset and, as a result, the ideal network classifier [16][20]. First, train the network, then delete the fewest number of important nodes (in input or hidden layers). After that, the reduced network is retrained, and one more least significant node is removed. This process is repeated until the desired trade-off between classification error and network size is achieved.

Jain and Zongker [19] present a two-class classification problem with Gaussian class-conditional densities in 20 dimensions (the same data was also employed by Trunk [23] to demonstrate the curse of dimensionality phenomenon). Needless to say, the standard of the chosen feature subset is poor for small training sets but improves as training set size increases.

For example, with 20 patterns in the training set, the branch-and-bound algorithm chose a subset of 10 features that shared only five features with the perfect subset of 10 features (when densities were known). With 2,500 patterns in the training set, the branch-and-bound procedure chose a ten-feature subset with only one incorrect feature. Figure 8 depicts an example of the feature selection procedure for the PCA features within the digit dataset using the floating search technique for two different training set sizes.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

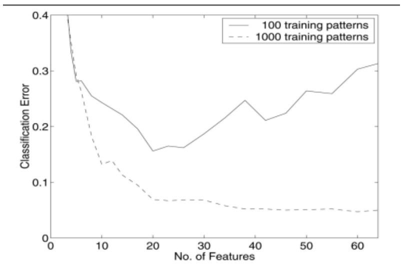

Figure 8: Classification error vs. the number of features using the floating search feature selection technique

Feature extraction is a sort of dimensionality reduction in which a large number of pixels in an image are efficiently represented in such a way that the image's most interesting sections are effectively captured.

2.3 Face Classification

The image must be categorized once the features have been retrieved and selected. Face classification algorithms that are based on appearance use a variety of categorization techniques. To enhance results, two or more classifiers are sometimes combined. On the other hand, most model-based algorithms match the samples to the model or template. The algorithm can then be improved via a learning process. Face recognition relies largely on classifiers in some form or another.

2.3.1 Classifiers

According to Jain, Duin, and Mao [13], three ideas are critical in the development of a classifier: similarity, probability, and decision boundaries.

1) Similarity: Similar patterns should be allocated to the same class. This method is implemented in the facial recognition algorithms that will be employed in the future. The objective is to identify a similarity metric as well as a representation of samples from the same class.
2) Probability: Some classifiers are built using a probabilistic method. The Bayes decision rule is commonly used in decision making. The Bayes error is the best criteria for assessing features, and Bayesian decision processes can generate the best classifier. As a result, an a posteriori probability function might be the best choice.
3) Decision boundaries can be compared to a Bayesian classifier in this technique. It is dependent on the metric chosen. The major goal of this method is to reduce the criteria (a measurement of inaccuracy) between the candidate and testing patterns.

2.3.1.1 Classifier combination

Finding a combination function that receives N-dimensional score vectors from M classifiers and generates N final classification scores is known as the classifier combination problem [14]. Many different classifier combinations are shown in Figure 9.

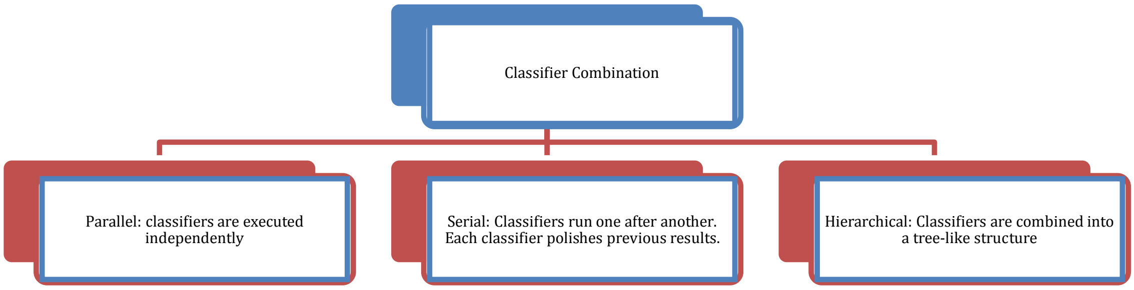

Figure 9: Classifier Combinations

Table 2 brings out a few possible classifier combinations and also provides information if they are trainable or not.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

Scheme	Architecture	Trainable	Info-level
Voting	Parallel	No	Abstract
Sum, mean, median	Parallel	No	Confidence
Product, min, max	Parallel	No	Confidence
Generalized ensemble	Parallel	Yes	Confidence
Adaptive weighting	Parallel	Yes	Confidence
Stacking	Parallel	Yes	Confidence
Borda count	Parallel	Yes	Rank
Behavior Knowledge Space	Parallel	Yes	Abstract
Logistic regression	Parallel	Yes	Rank
Class set reduction	Parallel/Cascading	Yes	Rank
Dempster-Shafer rules	Parallel	Yes	Rank
Fuzzy integrals	Parallel	Yes	Confidence
Mixture of Local Experts	Parallel	Yes	Confidence
Hierarchical MLE	Hierarchical	Yes	Confidence
Associative switch	Parallel	Yes	Abstract
Random subspace	Parallel	Yes	Confidence
Bagging	Parallel	Yes	Confidence
Boosting	Hierarchical	Yes	Abstract
Neural tree	Hierarchical	Yes	Confidence

Table 2: Classifier Combination Schemes

III. DIFFERENT APPROACHES TO FACE RECOGNITION

Face recognition is a discipline that is always evolving and improving. Computer vision, optics, pattern recognition, neural networks, machine learning, psychology, and other fields of study are all involved in face identification algorithms.The various ways are as follows:

3.1 Geometric/Template Based approaches

Local face characteristics and their geometric correlations are analysed using geometry feature-based approaches. [29] This method is also known as the feature-based method. The template-based approaches compare the input image to a collection of templates. Statistical technologies such as Support Vector Machines can be used to create the set of templates (SVM).

3.2 Template/statistical/neural network approaches

3.2.1 Template matching [30]: Patterns are represented by samples, models, pixels, curves, textures. The recognition function is usually a correlation or distance measure. This process can be understood clearly by Figure 10.

3.2.2 Statistical approach [31]: Patterns are represented as features. The identification function is a discriminant function.

3.3.3 Neural networks: The representation may vary. There is a network function at some point.

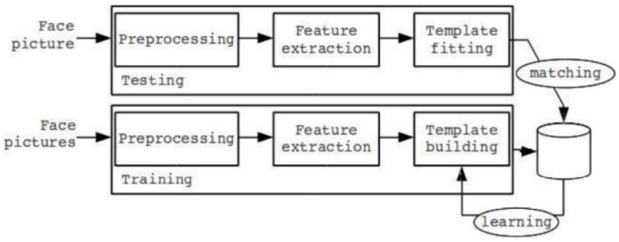

Figure 10: Template-matching algorithm diagram

3.3 Neural Network approach

Geometry feature-based techniques are used to analyze local facial traits and their geometric correlations. [29] This approach is also known as the feature-based approach. Template-based methods compare the input image to a set of templates. To construct the set of templates, statistical technologies such as Support Vector Machines might be used (SVM).

IV. OCCLUDED FACE DETECTION

In current biometric technologies, the human face is a research hotspot. Face collection is influenced by unpredictable circumstances in reality, such as expression change, posture change, illumination, occlusion, and so on, lowering the face recognition system's performance. Face recognition with occlusion has been a long-standing and difficult topic. In comparison to other challenges, there are few studies on occlusion issues. Face recognition with occlusion has a lot of application potential right now, so it's critical to solve the occlusion problem properly. Facial appearance changes substantially due to occlusion, as illustrated in Figure 11.

Figure 11: Occluded face images from MAFA dataset

4.1 Research Work

In [36], the authors discuss the primary reason behind the low performance obtained after testing the face recognition algorithms on partially occluded face images. In the case of upper face occlusion, a little loss in performance is observed due to missing discriminative information which can be improved after using a generic local appearance-based face recognition algorithm. When a local appearance-based face representation is applied in the case of a lower face occlusion, however, there is only a minor drop in performance. [37] proposes an efficient methodology of first analyzing the presence of potential occlusion on a face and then conducting face recognition on the non occluded facial regions based on selective local Gabor binary patterns.The authors of [38] proposed a unique technique for dealing with extreme facial occlusion by utilising the Omega shape created by the person's head and shoulder for head localization. This is done by constructing a novel energy function for elliptical head contour detection. [39] makes use of the OpenCV cascade classifier to detect the human eye and the mouth. This is followed by the face detection of the occlusion according to the relationship between the human eye and the human face. This paper also compares the accuracy of large-area occlusion of various occludes. Song et al [40] propose a novel PDSN (pairwise differential Siamese network) framework to explicitly find correspondence between occluded facial blocks and corrupted feature elements for deep CNN models. Further, based on the PDSN, a robust face recognition system for occlusions has been proposed. In [41], Zhang et al have introduced general face detection and occluded face detection methods. In [42] Anand et al. propose methods to detect the face in a still image. The paper discusses the experimental work conducted on images taken from the public dataset AR face dataset and Color FERET dataset having problem of face occlusion and non-uniform illumination.

4.2 Case Study: Detection Model of Occluded Face Based on YOLO Algorithm

In this study, a detection model of occluded face based on YOLO (You Only Look Once) was proposed to detect faces under three categories:

- Mask;
- No Mask; and
- Occluded Face

YOLO is a real-time object identification technique that uses neural networks. Because of its speed and precision, this algorithm is very popular. In a number of applications, it has been used to recognise traffic lights, pedestrians, parking metres, and animals. The YOLO approach uses convolutional neural networks (CNN) to recognise objects in real time. To detect objects, the approach just takes a single forward propagation through a neural network, as the name suggests. This indicates that a single algorithm run is used to forecast the entire image. The CNN is used to forecast multiple bounding boxes and class probabilities at the same time. There are several variations of the YOLO algorithm. Tiny YOLO, YOLOv3 and YOLOv5 are popular examples.

The Case study was conducted in the following two phases:

(i) Define YOLOv5 Model Configuration and Architecture; and

(ii) Training Custom YOLOv5 Detector

The smallest and fastest YOLOv5 base model was selected for the case study. Other YOLOv5 versions from which the base model could be chosen include:

1. YOLOv5s
2. YOLOv5m
3. YOLOv5l
4. YOLOv5x

The structure of the network can be edited in this step, though rarely it is needed.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

To kick off training, the training command was executed with the following options:
img: define input image size
batch: determine batch size
epochs: define the number of training epochs.
data: set the path to our yaml file
cfg: specify our model configuration
weights: specify a custom path to weights. name: result names
nosave: only save the final checkpoint
cache: cache images for faster training
During training, the mAP@0.5 is observed so as to see how the detector is performing.
4.3 Results

Figure 12: Training Images

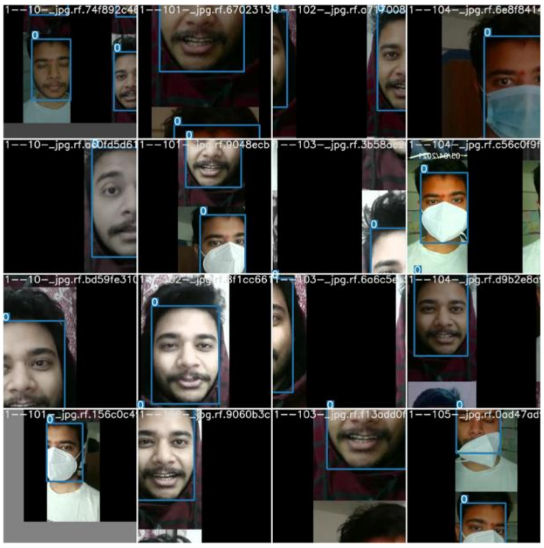

Figure 13: Mosaic Augmented training Images

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 40-53 www.ijaem.net ISSN: 2395-5252

Figure 14:89% Confidence

Figure 15:86% Confidence

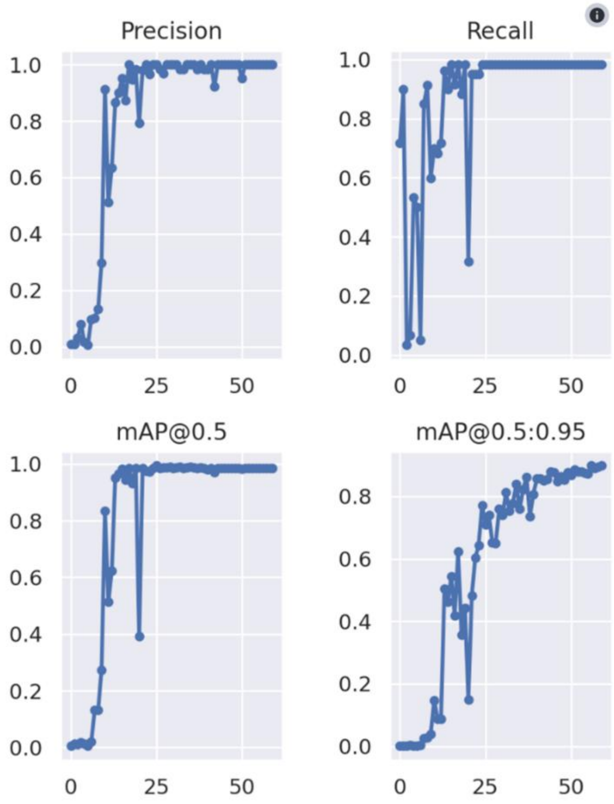

Figure 16:Precision, Recall, mAP@0.5, mAP@0.5:0.95

The results in figures14, 15 and 16 show that attempts have been made to make a system which could identify a person even if he was wearing a mask, or his face was partially covered, maybe synthetically or by some object placed in front of his face. It can be seen that the intended results have being achieved.

V. CONCLUSION

This paper goes into great length about face recognition techniques. Several approaches to face detection, feature extraction, and face recognition have been compared. The paper concludes with a discussion of the Occluded Face Identification problem and a case study of one of the occluded face detection techniques.

REFERENCES:

[1]. R. Louban. Image Processing of Edge and Surface Defects Theoretical Basis of Adaptive Algorithms with Numerous Practical Applications, volume 123, chapter Edge Detection, pages 9–29. Springer Berlin Heidelberg, 2009.

[2]. M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images:A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1):34–58, January 2002.

[3]. S. K. Singh, D. S. Chauhan, M. Vatsa, and R. Singh. A robust skin color-based face detection algorithm. Tamkang Journal of Science and Engineering, 6(4):227–234, 2003

[4]. M. C. Nechyba, L. Brandy, and H. Schneiderman. Lecture Notes in Computer Science, volume 4625/2009, chapter PittPatt Face Detection and Tracking for the CLEAR 2007 Evaluation, pages 126–137. Springer Berlin / Heidelberg, 2009.

[5]. K. Wan-zeng and Z. Shan-an. Multi-face detection based on downsampling and modified subtractive clustering for color images. Journal of Zhejiang University SCIENCE A, 8(1):72–78, January 2007.

[6]. L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A - Optics, Image Science and Vision, 4(3):519–524, March 1987.

[7]. M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(1):103–108, 1990.

[8]. M. Turk and A. Pentland. Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1):71–86, 1991
[9]. H. A. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. IEEE trans. Pattern Analysis and Machine Intelligence, 20(1):23–38, January 1998.
[10]. S. Bentin, T. Allison, A. Puce, E. Perez, and G. McCarthy. Electrophysiological studies of face perception in humans. Journal of Cognitive Neuroscience, 8(6):551–565, 1996.
[11]. R. Diamond and S. Carey. Why faces are and are not special. an effect of expertise. Journal of Experimental Psychology: General, 115(2):107– 117, 1986.
[12]. J. S. Bruner and R. Tagiuri. The perception of people. Handbook of Social Psychology, 2(17), 1954.
[13]. A. Jain, R. Duin, and J. Mao. Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1):4–37, January 2000.
[14]. S. Tulyakov. Review of classifier combination methods. Studies in Computational Intelligence (SCI), 90:361–386, 2008.
[15]. T.M. Cover and J.M. Van Campenhout, On the Possible Orderings in the Measurement Selection Problem, IEEE Trans. Systems, Man, and Cybernetics, vol. 7, no. 9, pp. 657-661, Sept. 1977.
[16]. G. Castellano, A.M. Fanelli, and M. Pelillo, An Iterative Pruning Algorithm for Feedforward Neural Networks, IEEE Trans. Neural Networks, vol. 8, no. 3, pp. 519-531, 1997
[17]. T.M. Cover, The Best Two Independent Measurements are not the Two Best, IEEE Trans. Systems, Man, and Cybernetics, vol. 4, pp. 116-117, 1974.
[18]. F. Ferri, P. Pudil, M. Hatef, and J. Kittler, Comparative Study of Techniques for Large Scale Feature Selection, Pattern Recognition in Practice IV, E. Gelsema and L. Kanal, eds., pp. 403-413, 1994.
[19]. K. Fukunaga, Introduction to Statistical Pattern Recognition. second ed., New York: Academic Press, 1990.
[20]. A.K. Jain and D. Zongker, Feature Selection: Evaluation, Application, and Small Sample Performance, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 2, pp. 153-158, Feb. 1997.
[21]. J. Mao, K. Mohiuddin, and A.K. Jain, Parsimonious Network Design and Feature Selection through Node Pruning, Proc. 12th Int'l Conf. Pattern on Recognition, pp. 622-624, Oct. 1994.
[22]. P. Pudil, J. Novovicova, and J. Kittler, Floating Search Methods in Feature Selection, Pattern Recognition Letters, vol. 15, no. 11, pp. 1,119-1,125, 1994.
[23]. P. Somol, P. Pudil, J. Novovicova, and P. Paclik, Adaptive Floating Search Methods in Feature Selection, Pattern Recognition Letters, vol. 20, nos. 11, 12, 13, pp. 1,157-1,163, 1999.
[24]. G.V. Trunk, A Problem of Dimensionality: A Simple Example, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 1, no. 3, pp. 306-307, July 1979.
[25]. Cai, D., He, X., & Han, J. (2007). Semi-supervised discriminant Analysis. 2007 IEEE 11th International Conference on Computer Vision. doi:10.1109/iccv.2007.4408856
[26]. Kinser, J. M. (2018). Principle component analysis. Image Operators, 111-126. doi:10.1201/9780429451188-8
[27]. Linear discriminant analysis. (n.d.). Statistical Pattern Recognition, 123-168. doi:10.1002/0470854774.ch4
[28]. Esparza, F. C., Tinoco, S. J., & Pabon, R. C. (2016). Face deformation system based on active shape models asm. 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA). doi:10.1109/stsiva.2016.7743330
[29]. T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham,Active Shape Models-Their Training and Application, Computer Vision and Image Understanding,Volume 61, Issue 1, 1995, Pages 38-59,
[30]. Template-based approaches. (1990). Readings in Speech Recognition, 113-114. doi:10.1016/b978-0-08-051584-7.50013-9
[31]. Template matching as testing. (n.d.). Template Matching Techniques in Computer Vision, 43-71. doi: 10.1002/9780470744055.ch3
[32]. Davies, E. (2012). Statistical pattern recognition. Computer and Machine Vision, 672-717. doi:10.1016/b978-0-12-386908-1.00024-0
[33]. M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1):34–58, January 2002.
[34]. Baltrusaitis, T., Zadeh, A., Lim, Y. C., &Morency, L. (2018). OpenFace 2.0: Facial behavior analysis toolkit. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). doi:10.1109/fg.2018.00019

[35]. Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. Journal of Applied Science and Technology Trends,1(2), 56-70. doi:10.38094/jastt1224

[36]. Mohammed, M. G., &Melhum, A. I. (2020). Implementation of hog feature extraction with tuned parameters for human face detection. International Journal of Machine Learning and Computing,10(5), 654-661. doi:10.18178/ijmlc.2020.10.5.987

[37]. Ekenel H.K., Stiefelhagen R. (2009) Why Is Facial Occlusion a Challenging Problem? In: Tistarelli M., Nixon M.S. (eds) Advances in Biometrics. ICB 2009. Lecture Notes in Computer Science, vol 5558. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-01793-3_31

[38]. Rui Min, Abdenour Hadid, Jean-Luc Dugelay, "Efficient Detection of Occlusion prior to Robust Face Recognition", The Scientific World Journal, vol. 2014, Article ID 519158, 10 pages, 2014. https://doi.org/10.1155/2014/519158

[39]. Tao Zhang, Jingjing Li, Wenjing Jia, Jun Sun, Huihua Yang, Fast and robust occluded face detection in ATM surveillance, Pattern Recognition Letters, Volume 107, 2018, Pages 33-40, ISSN 0167-8655, https://doi.org/10.1016/j.patrec.2017.09.011.

[40]. Z. Guo, W. Zhou, L. Xiao, X. Hu, Z. Zhang and Z. Hong, "Occlusion Face Detection Technology Based on Facial Physiology," 2018 14th International Conference on Computational Intelligence and Security (CIS), 2018, pp. 106-109, doi: 10.1109/CIS2018.2018.00031.

[41]. L. Song, D. Gong, Z. Li, C. Liu and W. Liu, "Occlusion Robust Face Recognition Based on Mask Learning with Pairwise Differential Siamese Network," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 773-782, doi: 10.1109/ICCV.2019.00086.

[42]. Zhifeng Zhang, Xiaohui Ji, Xiao Cui, and Junxia Ma. 2020. A Survey on Occluded Face recognition. In 2020 The 9th International Conference on Networks, Communication and Computing (ICNCC 2020). Association for Computing Machinery, New York, NY, USA, 40–49. DOI: https://doi.org/10.1145/3447654.3447661

[43]. Kumar, A., Kumar, M. & Kaur, A. Face detection in still images under occlusion and non-uniform illumination. Multimed Tools Appl **80,** 14565–14590 (2021). https://doi.org/10.1007/s11042-020-10457-9