Exploring the impact of image filters in facial detection through different classification algorithms

Raj Kishor Bisht
School of Computing
Graphic Era Hill University
Dehradun, India
bishtrk@gmail.com

Daivik Moahn
Dept. of Mathematics
Graphic Era Hill University
Dehradun, India
daivikmohan8@gmail.com

Aadershi Mohan
Dept. of Mathematics
Graphic Era Hill University
Dehradun, India
aashimohan21@gmail.com

Akshay Kumar Joshi
Dept. of Mathematics
Graphic Era Hill Univeristy
Dehradun, India
akshayjoshi2905@gmail.com

Yogesh Lohumi
Dept. of Computer Science and Engineering
Graphic Era Deemed to be Univeristy
Dehradun, India
yogesh95lohumi@gmail.com

***Abstract*— In the present work, we have analyzed the effect of applying different filters in face detection through different classification methods. For the purpose, we have created different datasets: first dataset consists of 1600 different unfiltered images of eight students with 200 images of each student and second dataset consists of 8000 filtered images using a mix of filters with 1000 images of each student. Then we have created five additional datasets, each dataset created by applying a single filter. Five filters namely blur, cut-out, mosaic, noise and shear are used for the present study. Six different algorithms CNN, Random Forest, SVM, Gaussian Naïve Bayes, KNN, and Logistic Regression were applied and results of unfiltered image dataset and filtered image dataset using a mix of filters have been compared. We found a little improvement in the accuracy in case of KNN and CNN and a little decline in accuracy of other algorithms applied to filtered dataset created using a mix of filters. The significance of difference between the two results is checked through paired t-test. On the basis of paired t-test, we conclude that filters in images, in general, do not have much effect in face detection through different classification algorithms. The results of unfiltered dataset and individual filtered datasets show that there is a significant difference in results between the unfiltered dataset and blur filtered dataset, others being identical. This indicates that the filter 'blur' has some effect in face detection and other has no significant effect.**

Keywords—face detection, filters, CNN, KNN, SVM, Random Forest, Naïve Bayes

I. Introduction

Face detection is the interaction of machines with visuals [1]. Face detection plays an important role in identifying human faces for various purposes like student attendance system [2], [3], [4], criminal investigation [5], security [6], [7] etc. Face detection is a computer vision technology that comprises the identification of human faces within digital images or video frames. It is the process of identifying individuals by their faces and it can be considered as a form of biometric authentication. Other forms of biometric authentications include recognition through fingerprints, voice, and iris etc. It is used in identifying peoples in images, videos, and real time. Most of us are already familiar with face recognition as we use it in our daily life through our smartphones. Face Recognition doesn't need a database of photos to check an individual's identity, it simply captures facial features of a person and recognizes that person as the sole owner of the device, while limiting access to others. Face Recognition uses Artificial Intelligence (AI) and different types of algorithms to identify and verify an individual by doing feature extraction. The process involves capturing an image, extracting features from the image, and comparing them with another image for identification. Face Recognition has very widespread of applications which includes tracking individual in public places, border, restricted areas etc. Due to the various challenges associated with automatic face detection and its wide applications, the area of face detection has a lot of potential for research.

II. Literature Review

Many researchers have worked on face recognition in various field, identifying faces in photos, videos or real-life face detection. A review of face detection methods available in the literatures is available in [8], [9], [10]. One of the earlier works using neural network for face detection is [11]. In face detection, features are identified, and similarity is measured. Euclidean distance between eyes, nose structure, distances between lips are some of the features that can be used for face detection. Facenet and facial landmark are the algorithms used for face detection[12]. In their work, the authors applied reverse engineering approach for identifying critical features for familiar face recognition[13]. Converting a face into segments like nose, hair, skin, eyes, mouth, background etc. and then applying classifications algorithms may be fruitful for face detection [14]. Skin color segmentation is one of the useful tasks in face recognition. In their work, the authors combined facial feature with skin color segmentation for face detection [15].

With the emergence of Covid-19, people started wearing masks covering majority of the area of their face. This created new challenges for face detection[16], [17] . In their work, the authors tried to answer the question of face recognition with masking [18] Some datasets are also available for research in the direction of face detection with masks[19]. Many methods have been proposed so far for face mask detection. Some of the methods are as follows: convolution neural network-based in [20], deep learning-based in[21] [22], a new face mask detector YOLOV3 in [23] and DeepMaskNet model in [24]. Automatic student attendance is one of the applications of face recognition which is quite demanding system as with the growing nature of technology, attendance of students in classes should be automatic for time saving, accuracy and discipline. A face recognition based automatic student attendance system using face biometric is proposed in [25]. Online students' attendance with face masks is proposed in [26].

There are many other areas where face detection is quite useful. Monitoring videos and face detection in videos is also a challenging task. The authors in their work, proposed a video oriented intelligent face detection algorithm utilizing deep learning by cascading different features [27]. Security of vehicles in parking with face detection is discussed in [28]. There are some open-source libraries used for face detection. OpenCV in one of such libraries used by researchers for face detection [29], [30], [31]. Filters also play a crucial role in modifying a face. The effect of different filters is studied in [32].

The organization of the paper is as follows: Section I deals with introduction to the topic; Section II provides literature review related to the topic. In Section III, we describe the methodology of the proposed work and Section IV shows the experimental results. Finally, we conclude the study in Section V.

III. Methodology

To compile the dataset, first, we have taken the images of eight students using OpenCV as shown in Fig. 1. Using these images, a dataset of 1600 different images has been created, with 200 images of each student. These are coloured images of size 128 × 128. We call this dataset as unfiltered (original) dataset. Since, images of eight different students are there; thus, we have eight classes for classification. We have also created another dataset, called filtered dataset, which consist of the filtered images of the students. We have applied the filters like shear, grayscale, cut-out, hue, saturation, brightness, exposure, blur, noise, mosaic, old, with specs and few more. Some other filters involve rotation of images at various angles including 30, 45, 60, 120, 135, -30, - 60, -45, -135 etc. The number of images per student is increased after applying different filters. We have selected 1000 images per student randomly and a filtered dataset of 8000 images has been created. A sample of filtered dataset is shown in Fig. 2. Then we have created five more dataset to compare the effect of individual filter on the unfiltered or original dataset. These datasets have been created by applying the filters on the original dataset keeping the size of each dataset equal. Here we have worked on five main filters blur, cut-out, mosaic, noise and shear filters. The sample images from each of these datasets are shown in Fig.3.

Fig. 1. Image of students whose data has been collected.

Now after compiling the datasets, we have applied six algorithms CNN, Random Forest, SVM, Gaussian Naïve Bayes, KNN, and Logistic Regression on both the datasets. For the execution of CNN, we have used 3 convolutional layers, 2 dense layers and a flatten layer followed by the 'ReLU' activation function as shown in Fig. 4. We have used 'Adam' Optimizer, and 'Categorical-cross entropy' as loss function. We considered 80 percent training data and 20 percent testing data and applied stratified cross validation technique so that we can have equal portions of images for both training and testing followed by a five-convolution layer of which four layers work on 'relu' which can be defined as:

$$F(x) = \max(0, x) \tag{1}$$

where x being any positive value and the last layer works on 'softmax' which is defined as:

$$\text{softmax}(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}} \tag{2}$$

where Z is the input vector, N is the number of classes in a multi-class classifier, i represents input and j represents output. To compile this model, we have used the 'Adam' optimizer and categorical cross-entropy loss is defined by the formula:

$$\text{CE} = -\sum_{i}^{c} t_i \, \log(f(s)_i) \tag{3}$$

where t_i is the truth label and $f(s)_i$ is the Softmax probability for i^{th} class. The results of the application of this model are described in Section IV.

Fig. 2. Sample of Filtered Images

We applied GridSearchCV to find the parameters resulting in the best accuracy for Random Forest, SVM, Gaussian Naïve Bayes, KNN, and Logistic Regression. Finally, we have applied the paired t-test to check whether the results of unfiltered and filtered datasets are significantly different or not. The null and alternative hypotheses under study are H_0: The accuracy results for different algorithms for both unfiltered and filtered datasets are identical and H_1: The accuracy results of different algorithms for both original and filtered datasets are not identical. The paired t test is defined as:

$$t = \frac{\sum d}{\sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n-1}}} \quad (4)$$

where, d is the difference of paired value, n is the number of samples. We have considered $\alpha = 0.05$ as the level of significance.

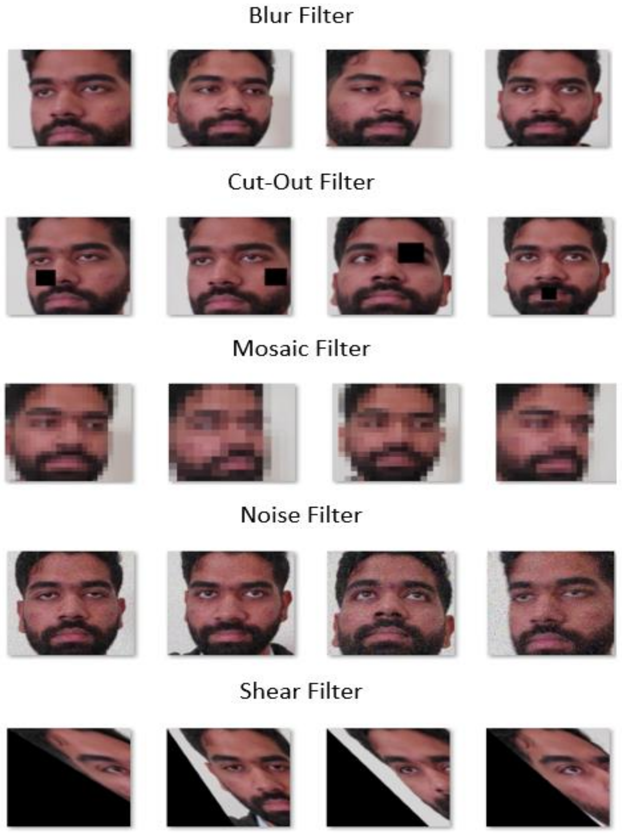

Fig. 3. Sample of Individual Filtered Images

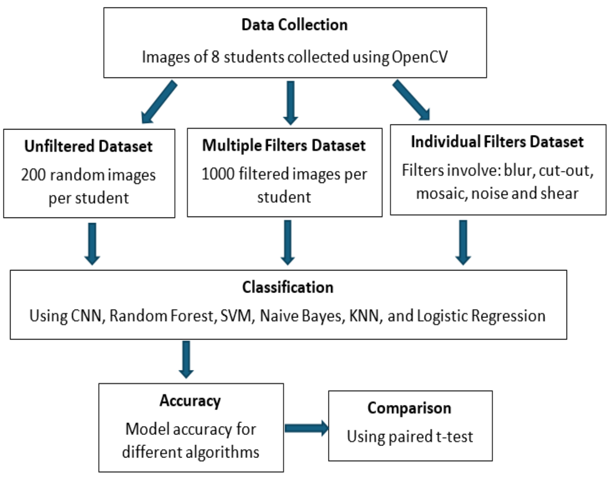

Fig. 4. Flowchart of the proposed work

IV. Experimental Results

The accuracies of six different algorithms for both unfiltered and filtered datasets are shown in Table I.

TABLE I. ACCURACY OF DIFFERENT ALGORITHMS

Algorithm	Unfiltered dataset	Filtered dataset
CNN	0.9938	0.9981
Random Forest	1.00	0.9993
SVM	1.00	0.9993
Naïve Bayes	0.9406	0.9118
KNN	0.9969	0.9987
Logistic Regression	1.00	0.9987

From Table I, we can observe that there is a little improvement in the accuracy in case of KNN, and CNN and a little decline in accuracies in case of other algorithms. But, since the difference margin is not very significant so to check if both the accuracies are identical or not, we applied paired t-test. We have taken significance level $\alpha = 0.05$ for inference. On applying paired t-test, we obtained the value of t-statistic = 0.84875 and p-value = 0.43475. Since, the p-value (the observed level of significance) is greater than the significance level i.e. $\alpha = 0.05$, we accept the null hypothesis. In other words, we conclude that the two accuracy results of different algorithms for both unfiltered and filtered datasets are identical.

TABLE II. ALGORITHM ACCURACY FOR INDIVIDUAL FILTER DATASETS

Algorithm	Un-filtered dataset	Blur	Cut-out	Mosaic	Noise	Shear
CNN	0.99	0.99	0.99	0.99	0.99	0.96
Random Forest	1.00	1.00	1.00	1.00	0.97	0.97
SVM	1.00	0.99	0.99	0.99	0.99	0.98
Naïve Bayes	0.94	0.91	0.89	0.95	0.94	0.65
KNN	0.99	0.99	0.99	0.99	0.99	0.81
Logistic Regression	1.00	0.99	0.99	0.99	0.99	0.97

From Table II, we can observe the accuracies for different algorithms applied on the five filtered dataset. Later on applying the paired t-test between unfiltered dataset and filtered datasets separately we get the p values as shown in Table III. We have taken significance level $\alpha = 0.05$ for inference. When the p-value (the observed level of significance) is greater than the significance level i.e. $\alpha = 0.05$, we accept the null hypothesis. So, we can clearly see that in case of blur filter the p-value is less than 0.05 indicating that there is a significant difference between the results of original dataset and blur dataset, that is, the results of two datasets are not identical. Apart from that, all other filter datasets involving mosaic, shear, noise and cut-out filters have greater p-value than 0.05 thus for these algorithms indicating that the accuracy scores are identical to the original dataset.

TABLE III. ACCURACY OF DIFFERENT ALGORITHMS

Filters	t-statistic	p-value
Blur	2.714	0.042
Cut-out	1.312	0.246
Mosaic	-0.168	0.875
Noise	2.043	0.096
Shear	2.048	0.095

V. DISCUSSION AND CONCLUSION

The present work is conducted to check the effect of applying different filters in face detection through different classification methods. For the purpose, we have created different datasets: one unfiltered consists of the original images and second filtered dataset consisting of the images after applying mix of filters and other five filtered image datasets using individual filters. Six different algorithm CNN, Random Forest, SVM, Gaussian Naïve Bayes, KNN, and Logistic Regression were applied, and results of unfiltered and filtered image datasets have been compared. We found a little improvement in the accuracy in case of KNN and CNN and a little decline in the accuracy of other algorithms for filtered dataset. Further, based on paired t-test, we conclude that the two results are not significantly different, that is, a mix of filters in images, in general, do not have much effect in face detection through different classification algorithms. However, the filter 'blur' has a significant effect in classification if applied individually.

REFERENCES

[1] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: Past, present and future," *Computer Vision and Image Understanding*, vol. 138, 2015, doi: 10.1016/j.cviu.2015.03.015.

[2] Prof. S. Ahire, R. Jadhao, V. Gawai, and P. Giri, "Automated Student Attendance Management System Using Face Recognition," *International Journal of Advanced Research in Science, Communication and Technology*, 2021, doi: 10.48175/ijarsct-2019.

[3] A. Khan, I. khan, J. Golandaj, A. George, and S. Bhandari, "A Survey on Automated Student Attendance Management System Using Face Recognition," *International Journal of Research Publications*, vol. 122, no. 1, 2023, doi: 10.47119/ijrp1001221420234626.

[4] U. Patel and P. Swaminarayan, "Development of a Student Attendance Management System Using RFID and Face Recognition," *International Journal of Advance Research in*, vol. 2, no. 8, 2014.

[5] L. F. Huang *et al.*, "Criminal Investigation with the help of Face Recognition," *International Journal of Engineering Research & Technology (IJERT)*, vol. 2, no. 6, 2021.

[6] N. S. Irjanto and N. Surantha, "Home Security System with Face Recognition based on Convolutional Neural Network," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, 2020, doi: 10.14569/IJACSA.2020.0111152.

[7] O. E. Putra, R. Devita, and N. Wahyudi, "Safe Security System Using Face Recognition Based on IoT," *SinkrOn*, vol. 8, no. 2, 2023, doi: 10.33395/sinkron.v8i2.12231.

[8] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," *Artif Intell Rev*, vol. 52, no. 2, 2019, doi: 10.1007/s10462-018-9650-2.

[9] M. K. Hasan, M. S. Ahsan, Abdullah-Al-Mamun, S. H. S. Newaz, and G. M. Lee, "Human face detection techniques: A comprehensive review and future research directions," *Electronics (Switzerland)*, vol. 10, no. 19, 2021, doi: 10.3390/electronics10192354.

[10] E. Hjelmås and B. K. Low, "Face detection: A survey," *Computer Vision and Image Understanding*, vol. 83, no. 3, 2001, doi: 10.1006/cviu.2001.0921.

[11] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans Pattern Anal Mach Intell*, vol. 20, no. 1, 1998, doi: 10.1109/34.655647.

[12] S. Saleem, J. Shiney, B. Priestly Shan, and V. Kumar Mishra, "Face recognition using facial features," *Mater Today Proc*, vol. 80, 2023, doi: 10.1016/j.matpr.2021.07.402.

[13] N. Abudarham, L. Shkiller, and G. Yovel, "Critical features for face recognition," *Cognition*, vol. 182, 2019, doi: 10.1016/j.cognition.2018.09.002.

[14] S. Benini, K. Khan, R. Leonardi, M. Mauro, and P. Migliorati, "Face analysis through semantic face segmentation," *Signal Process Image Commun*, vol. 74, 2019, doi: 10.1016/j.image.2019.01.005.

[15] S. Yadav and N. Nain, "A novel approach for face detection using hybrid skin color model," *J Reliab Intell Environ*, vol. 2, no. 3, 2016, doi: 10.1007/s40860-016-0024-8.

[16] Sudhakar P. Rupnar, Prathamesh Gujjalwar, Abhishek Aradhye, Chandroday Kalshetty, Shubham Gade, and Mr. V. M. Sale, "Face Mask Detection," *International Journal of Advanced Research in Science, Communication and Technology*, 2022, doi: 10.48175/ijarsct-4583.

[17] I. Javed *et al.*, "Face mask detection and social distance monitoring system for COVID-19 pandemic," *Multimed Tools Appl*, vol. 82, no. 9, 2023, doi: 10.1007/s11042-022-13913-w.

[18] Y. M. Saib and S. Pudaruth, "Is Face Recognition with Masks Possible?," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, 2021, doi: 10.14569/IJACSA.2021.0120706.

[19] Z. Wang, B. Huang, G. Wang, P. Yi, and K. Jiang, "Masked Face Recognition Dataset and Application," *IEEE Trans Biom Behav Identity Sci*, vol. 5, no. 2, 2023, doi: 10.1109/TBIOM.2023.3242085.

[20] M. Umer *et al.*, "Face mask detection using deep convolutional neural network and multi-stage image processing," *Image Vis Comput*, vol. 133, 2023, doi: 10.1016/j.imavis.2023.104657.

[21] A. Sharma, R. Gautam, and J. Singh, "Deep learning for face mask detection: a survey," *Multimed Tools Appl*, vol. 82, no. 22, 2023, doi: 10.1007/s11042-023-14686-6.

[22] Q. M. Ilyas and M. Ahmad, "An enhanced deep learning model for automatic face mask detection," *Intelligent Automation and Soft Computing*, vol. 31, no. 1, 2022, doi: 10.32604/IASC.2022.018042.

[23] X. Jiang, T. Gao, Z. Zhu, and Y. Zhao, "Real-time face mask detection method based on yolov3," *Electronics (Switzerland)*, vol. 10, no. 7, 2021, doi: 10.3390/electronics10070837.

[24] N. Ullah, A. Javed, M. Ali Ghazanfar, A. Alsufyani, and S. Bourouis, "A novel DeepMaskNet model for face mask detection and masked facial recognition," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, 2022, doi: 10.1016/j.jksuci.2021.12.017.

[25] D. Narendar Singh, M. Kusuma Sri, and K. Mounika, "IOT based automated attendance with face recognition system," *International Journal of Innovative Technology and*

Exploring Engineering, vol. 8, no. 6 Special Issue 4, 2019, doi: 10.35940/ijitee.F1093.0486S419.

[26] M. H. M. Kamil, N. Zaini, L. Mazalan, and A. H. Ahamad, "Online attendance system based on facial recognition with face mask detection," *Multimed Tools Appl*, vol. 82, no. 22, 2023, doi: 10.1007/s11042-023-14842-y.

[27] Z. Dong, J. Wei, X. Chen, and P. Zheng, "Face Detection in Security Monitoring Based on Artificial Intelligence Video Retrieval Technology," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2982779.

[28] N. M. Raharja, M. A. Fathansyah, and A. N. N. Chamim, "Vehicle Parking Security System with Face Recognition Detection Based on Eigenface Algorithm," *Journal of Robotics and Control (JRC)*, vol. 3, no. 1, 2022, doi: 10.18196/jrc.v3i1.12681.

[29] M. Khan, S. Chakraborty, R. Astya, and S. Khepra, "Face Detection and Recognition Using OpenCV," in *Proceedings - 2019 International Conference on Computing, Communication, and Intelligent Systems, ICCCIS 2019*, 2019. doi: 10.1109/ICCCIS48478.2019.8974493.

[30] A. Kumar, S. Chaudhary, S. Sangal, and R. Dhama, "Face Detection and Recognition using OpenCV," *Int J Comput Appl*, vol. 184, no. 11, 2022, doi: 10.5120/ijca2022922085.

[31] R. T. H. Hasan and A. B. Sallow, "Face Detection and Recognition Using OpenCV," *Journal of Soft Computing and Data Mining*, vol. 2, no. 2, 2021, doi: 10.30880/jscdm.2021.02.02.008.

[32] P. Hedman, V. Skepetzis, K. Hernandez-Diaz, J. Bigun, and F. Alonso-Fernandez, "On the effect of selfie beautification filters on face detection and recognition," *Pattern Recognit Lett*, vol. 163, 2022, doi: 10.1016/j.patrec.2022.09.018.