Evaluating Impact of Wearing Masks in Face Recognition Using Deep Learning Algorithms

Mustafa Atay
Department of Computer Science
Winston-Salem State University
Winston-Salem, NC USA
ataymu@wssu.edu

Megh Poudyel
Department of Computer Science
Winston-Salem State University
Winston-Salem, NC USA
mpoudyel119@rams.wssu.edu

***Abstract*— Automated and contactless face recognition is a widely used machine learning technology for identifying people which has been applied in scenarios like secure login to electronic devices, automated border control, community surveillance, tracking school attendance. The use of face masks has become essential due to the global spread of COVID-19, raising concerns about the performance of recognition systems. Conventional face recognition technologies were primarily designed to work with unmasked faces, and the widespread use of masked face images significantly degrades their performance. To address this understudied issue, we evaluated the performance of six deep learning models, namely, VGG-16, AlexNet, GoogleNet, LeNet, ResNet-50, and FaceNet on masked and unmasked face images. We aim to find out if deep learning models struggle with masked face recognition and identify the models that mitigate the impact of masked face images. We track, and report miss rates for both masked and unmasked images, along with performance metrics like accuracy and F1 scores in this paper.**

Keywords—masked face recognition, deep learning, ocular biometrics, synthesized mask, Covid-19 pandemic

I. INTRODUCTION

Conventional facial recognition systems performed well with unmasked faces up until COVID-19 pandemic. However, as the pandemic led to mandatory mask-wearing in many countries, these systems started to struggle in identifying masked face images. Occluded faces posed challenges for facial recognition solutions, while addressing occlusion invariance became a growing research concern.

While various studies have been conducted on occluded facial recognition solutions, none have conducted a comparative study using masked, unmasked, and half-masked training datasets along with multiple Deep Learning (DL) models to thoroughly explore their strengths and weaknesses. In this study, we trained and tested six DL models (VGG-16, AlexNet, GoogleNet, LeNet, ResNet-50, and FaceNet) with unmasked, masked, and half-masked face images. Our goal was to identify the high-performing models in each case and report the best and poor performers.

We conducted experiments using images taken in a controlled environment from facial image databases with sufficient number of images per subject for training and testing. We tracked and reported miss rates for masked and unmasked images, as well as performance metrics such as accuracy, precision, recall, and F1 score.

There are various facial image databases that contain images captured in a controlled environment, uncontrolled environment, or both. In this initial study, we choose to work with images taken in a controlled environment. So, we explore and use those databases that contain all images captured in a controlled environment and have sufficient number of facial images per subject for training and testing.

II. RELATED WORK

Several models have been developed and widely used for unmasked face recognition [10, 16, 17, 19-23]. However, there has been limited progress in the field of masked face recognition [1, 4-6, 8,15, 18].

Dharanesh et al. [5] proposed a solution for recognizing faces with masks using a dynamic ensemble of deep learning models. Their experimental results suggested that their solution achieved comparable performance to conventional deep learning face recognition systems without masks. They experimented with DL models using only one database.

Damer et al. [1] performed an exploratory analysis of face recognition systems considering the effect of masks on recognition performance. They studied two non-commercial models (ArcFace and SphereFace) and one commercial off-the-shelf (COTS) model (MegaMatcher 11.2 SDK). The effect of masks was significant, especially on the genuine score's distribution. They emphasized the need to re-evaluate face recognition solutions for proper performance when dealing with masked faces. They experimented with three DL models using their specifically collected database.

Montero et al. [6] proposed an approach modifying the ArcFace model to create a Multi-Task ArcFace model, which showed higher accuracy in recognizing masked faces without compromising accuracy on non-masked datasets. They experimented with their proposed Multi-Task ArcFace and the original ArcFace DL models.

Anwar et al. [4] addressed a methodology to augment current facial datasets with masked faces using the open-source tool MaskTheFace. Their approach achieved low false-positive rates and high accuracy in recognizing masked faces without

1946-0759/23/$31.00 ©2023 IEEE
DOI 10.1109/ICMLA58977.2023.00217

requiring a new dataset. They experimented with the FaceNet model.

Ejaz et al. [8] applied Principal Component Analysis (PCA), a successful tool for non-masked face recognition, to the masked face recognition problem. Their comparative study using the ORL face database showed poor recognition rates for masked face images compared to non-masked faces. They used only the unsupervised PCA model and the ORL database.

While we evaluate masked face recognition system with 6 DL algorithms, we utilize 3 different types of training datasets using ORL and GTFD databases separately in our study. None of the above studies in the literature conducted a thorough evaluation study with multiple databases using large number of DL models, and including multiple scenarios like training with unmasked versus testing with unmasked (UM/UM), training with unmasked versus testing with masked (UM/M), training with half-masked (that is, 50% images are unmasked and remaining 50% images are masked) versus testing with unmasked (HM/UM), training with half-masked versus testing with masked (HM/M), training with masked versus testing with unmasked (M/UM) and training with masked versus testing with masked (M/M) . We believe that our study gives a broader view of masked face recognition compared to the existing work in the literature to the best of our knowledge.

III. Preliminaries

We first provide a brief overview of commonly used DL models in face recognition tasks. We then introduce the 6 CNN algorithms (VGG-16, AlexNet, GoogleNet, LeNet, ResNet-50, and FaceNet) used in our experiments, along with the ORL (Our Database of Faces) and GTFD (Georgia Tech Face Database) image databases. To address masked face recognition, different DL models have been employed based on specific needs and requirements:

Fig. 1. Sample images from GTFD

1. Convolutional Neural Networks (CNN): CNNs are widely used for image recognition tasks, including face recognition. They excel at feature extraction, a crucial aspect of masked face recognition. Models like VGGNet, ResNet, InceptionNet, AlexNet, FaceNet, LeNet, and GoogleNet have been adapted and fine-tuned for image recognition tasks. We choose to work with CNN models in our study due to their strength at feature extraction.
2. Siamese Network: This network involves one or more identical networks and is used for one-s hot or few-shot learning. It compares two images to determine if they belong to the same person. This network is useful when labeled masked face data is limited.
3. Generative Adversarial Networks (GANs): GANs are deep neural networks that can generate new data resembling a given set of training data. In face recognition, they have been used to synthesize masked face images for data augmentation and model training, effectively improving masked face recognition tasks.
4. One-Shot Learning Model: This Machine Learning based algorithm evaluates the similarity and differences between two images. In face recognition, it aims to recognize a subject's face with only one example, utilizing metric learning and feature embedding techniques to identify subjects even with limited data.
5. Ensemble Models: Ensemble models combine multiple learning algorithms to achieve better performance than individual algorithms alone. They enhance accuracy and robustness by combining different models trained on various datasets.

A. CNN Models Used in Our Experiments

Classification algorithms are used to classify objects of various types. They help to classify objects into similar or dissimilar groups. These algorithms also play an integral role in facial recognition. They help to categorize the images and determine their relationship to each other. Our study uses a total of 6 different CNN classification algorithms for experimentation. They are VGG-16, AlexNet, GoogleNet, LeNet, ResNet-50, and FaceNet.. These are Convolutional Neural Networks (CNN) models which are broadly used for image recognition tasks, including face recognition. Feature extraction from images is a crucial part in masked face recognition, on which CNNs models perform good.

B. Databases

Two publicly available facial image databases, namely, ORL (Our Database of Faces) and GTFD (Georgia Tech Face Database) are selected for our study. Since our study is focused on performing experiments in a controlled environment, ORL and GTFD are the adequate choices made for our study as they are generated in controlled environments. In addition, these two databases have an equal number of images per subject. Both of these databases lack masked images. We used “MaskTheFace” software and generated masked face images out of the unmasked images from these databases. ORL face database consists of 10 different images of 41 subjects with a total of 410 images. The size of each image is 92x112 with 8-bit grey levels in PGM format. [8][13]. GTFD database consists of a total of 750 images of 50 different subjects each having 15 images in various conditions and poses. This database is commonly used for research and development work in the domain of computer vision. This database is used by most researchers to study facial recognition algorithms and techniques [13]. Figures 1 shows sample images from the GTFD database.

Fig. 2. Sample masked GTFD images

C. Synthesizing Masked Face Images

We used open-source software, MaskTheFace to augment faces from the ORL and GTFD databases with masks. MaskedTheFace is a computer-vision based software which is used to synthesize masked face images. It uses a dlib based face-landmark detector to recognize face tilt. It has six mask templates to use from. Based on the face tilt, mask template is chosen from the library of masks. We can select several masks for generating masked faces. We can use this software to convert most face images to masked-face images. A single image or the whole directory of images can be converted to masked faces [14]. Figures 2 shows samples of masked faces from GTFD database.

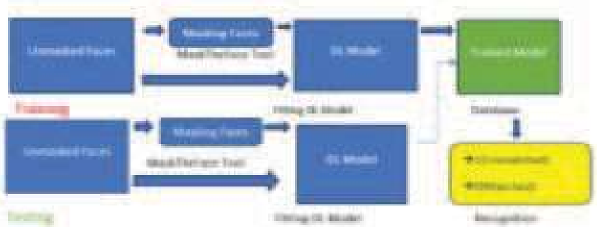

Fig. 3. Experimental Design

IV. METHODOLOGY

In our study, we conduct experiments with masked, half-masked and unmasked training datasets to investigate the impact of masked face recognition. We perform experiments with 6 selected DL models to be able to compare effectiveness of DL techniques in mitigating impact of wearing masks in facial recognition. We repeat all our experiments with datasets formed out of two face image databases. As both image databases are only composed of unmasked images, we utilize the open-source software MaskTheFace [14] to generate masked counterparts of unmasked images.

We use two databases, namely, ORL (Our Database of Faces) and GTFD (Georgia Tech Face Database), and conduct experiments with 6 DL algorithms.

We performed six experiments using the 6 DL models for each of the ORL and GTFD databases. We conducted 12 experiments with both ORL and GTFD databases. The execution time for each of the 12 experiments with DL algorithms conducted in our study was between 5 minutes to 45 minutes on average. Figure 3 illustrates the experimental design in our study.

Experiment 1: *Training with unmasked and testing with unmasked images.* We performed these experiments by training 6 DL models using both ORL and GTFD databases to observe the performance of unmasked face recognition with DL models when the system is completely trained with unmasked images. In ORL, we used 8 unmasked images for training, 1 unmasked image for validating for each subject and then tested each of the DL models with 41 unmasked images, 1 for each one of 41 individuals. Similarly, In GTFD, we used 12 unmasked images for training, 2 unmasked images for validating for each 50 subjects and then tested each of the DL models with 50 unmasked images, 1 for each one of 50 individuals.

Experiment 2: *Training with unmasked and testing with masked images.* We performed these experiments by training 6 DL models for both ORL and GTFD databases to observe the performance of masked face recognition with DL models when the system is completely trained with unmasked images. We followed the same procedures as in Experiment 1 except we tested the models with masked images.

Experiment 3: *Training with masked and testing with unmasked images.* We performed these experiments by training 6 DL models for both ORL and GTFD databases to observe the performance of unmasked face recognition with DL models when the system is completely trained with masked images. In ORL, we used 8 masked images for training, 1 masked image for validating and then tested each of the DL models with 41 unmasked images, 1 for each one of 41 individuals. Similarly, In GTFD, we used 12 masked images for training, 2 masked images for validating for each subject and then tested each of the DL models with 50 unmasked images, 1 for each one of 50 individuals.

Experiment 4: *Training with masked and testing with masked images.* We performed these experiments by training 6 DL models for both ORL and GTFD databases to observe the performance of masked face recognition with DL models when the system is completely trained with masked images. We followed the same procedures as in Experiment 3 except we tested the models with masked images.

Experiment 5: *Training with half-masked and testing with unmasked images.* In this experiment, we aim to observe the performance of unmasked face recognition with DL models when the system is trained with half-masked images. Out of 10 images of each 41 individual in ORL database, 1 image is set aside for testing and 2 images for validating, while from the remaining 7 images, 3 images are masked with MaskTheFace software and combined with 3 unmasked images from the remaining images making a total of 6 images. Thus, we generated this new dataset having 6 images and named it as half-masked image dataset. This dataset contains 50% masked images combined with 50% unmasked images. This dataset containing 3 masked and 3 unmasked images of each 41 subjects was used to train 6 DL models. Then, we tested each of the DLmodels with 41 unmasked images, 1 for each one of 41 individuals. Similarly, out of 15 images of each 50 individuals in GTFD database, 1 image is set aside for testing, 2 images are set aside for validating, while from the remaining 12 images, 6 images are masked with MaskTheFace software and combined with 6 unmasked images to make a total of 12 images, that is, 50% masked images combined with 50% unmasked images. The dataset containing 6 masked and 6 unmasked images of each one of 50 subjects was used to train 6 DL models. Then, we tested each of the models with 50 unmasked images, one for each one of 50 individuals.

Experiment 6: *Training with half-masked and testing with masked images.* In this experiment, we aim to observe the performance of masked face recognition with DL models when the system is trained with half-masked images. We followed the same procedures as in Experiment 5 except we tested the models with masked images.

V. EXPERIMENTAL RESULTS

We used 3 training datasets for the DL models which are Trainin_UM, Training_HM, and Training_M. Then, we generated 3 validation datasets, namely, Validating_UM, Validating_HM, and Validating_M for validating our models. To test our models, we generated 2 testing datasets, namely,

Testing_UM and Testing_M. We used two face image databases, which are ORL (Our Database of Faces), and GTFD (Georgia Tech Face Database). We created those training, validating, and testing datasets for both databases.

We chose 6 deep learning (DL) algorithms used in face recognition studies which are VGG16, AlexNet, GoogleNet, LeNet, FaceNet, and ResNet50. We trained, validated, and tested these selected DL models. The experiments were performed in PyCharm environment. The results are recorded and reported in the following section.

A. Experimentation with ORL Database

In this experimentation, we choose 6 deep learning algorithms and conduct experiments using training , validation and testing datasets prepared from ORL database. Through these experiments, we aim to test and compare the performance of the 6 DL algorithms by using metric such as accuracy. We train these algorithms using unmasked, half-masked and masked datasets separately, and test them using unmasked and masked datasets prepared from the ORL database. We also track and report miss rates for masked and unmasked images besides reporting performance metrics such as accuracy.

TABLE 1. OVERALL AVERAGE MISS RATES OF 6 DL MODELS FOR ORL DATABASE

Overall Miss Rates for All Datasets	UM	HM	M	Average
Average Unmasked Miss Rate	2.0%	7%	4%	4%
Average Masked Miss Rate	28.0%	9%	4%	14%
Averages	15.0%	8.0%	4.0%	9.0%

As shown in Table 2, the DL models trained with unmasked face images and tested with masked images, AlexNet model is found to have highest accuracy of 88% whereas GoogleNet is found to have the lowest of 51%.

The models trained with half-masked and tested with masked images, FaceNet has the highest accuracy of 100% and ResNet50 has the lowest of 76%. The models trained with masked images and tested with masked images, both AlexNet and FaceNet have the highest accuracy of 100% and GoogleNet has lowest score of 93%.

In Table 2, we observe that the highest average performance is 99% when dataset is trained with unmasked images and tested with unmasked images. This is also understandable because these DL models are tuned to work with unmasked face images. We observe that the lowest average performance is 72% when the system is trained with unmasked faces and tested with masked faces. This shows that models trained with unmasked faces are not suitable for testing with masked faces.

TABLE 2. THE EXPERIMENTAL RESULTS SHOWING ACCURACY OF 6 DL MODELS WITH ORL DATABASE

Accuracy Table	Experiments						AVERAGES
DL Algorithm	UM/UM	UM/M	HM/UM	HM/M	M/UM	M/M	
VGG16	100%	82%	95%	99%	98%	98%	[illegible]
AlexNet	95%	88%	100%	99%	96%	100%	[illegible]
GoogleNet	98%	51%	91%	98%	98%	93%	[illegible]
LeNet	100%	68%	95%	97%	98%	98%	[illegible]
FaceNet	100%	86%	100%	100%	98%	100%	[illegible]
ResNet50	98%	63%	78%	76%	98%	98%	[illegible]
AVERAGES	99%	72%	94%	91%	97%	97%	[illegible]

In Table 2, we see that for testing masked face images, the average accuracy of DL models decreases when trained with unmasked images, but the accuracy increases when the DL models are trained with masked faces, and the best performance is observed when the models are trained with masked images.

As shown in Table 2, LeNet, FaceNet and VGG16 are found to outperform other models in identifying unmasked facial images for unmasked training datasets. FaceNet and AlexNet both outperform other models for identifying unmasked face images when trained with half-masked face images. Except AlexNet, all 5 tested models perfom best for identifying masked images when trained with masked images.

In Table 1, we notice that while testing masked face images, models trained with masked face images have the lowest average miss rate of 4%, and the models trained with unmasked face images have the highest average miss rates of 28%. But, while testing unmasked face images, we observe that models trained with half-masked face images have the highest average miss rate of 7%, and the models trained with unmasked face images have lowest average miss rate of 2%.

B. Experimentation with GTFD Database

In this experimentation, we choose 6 deep learning algorithms and conduct experiments using training , validation and testing datasets prepared from GTFD database. Through these experiments, we aim to test and compare the performance of the 6 DL algorithms by using metrics such as accuracy. We train these algorithms using unmasked, half-masked and masked datasets separately, and test them using unmasked and masked datasets prepared from the GTFD database. We also track and report miss rates besides reporting performance metrics such as accuracy.

TABLE 3. THE EXPERIMENTAL RESULTS SHOWING ACCURACY OF 6 DL MODELS WITH GTFD DATABASE

Accuracy Table	Experiments						AVERAGES
DL Algorithm	UM/UM	UM/M	HM/UM	HM/M	M/UM	M/M	
VGG16	92%	56%	90%	80%	74%	86%	80%
AlexNet	96%	74%	88%	94%	96%	96%	91%
GoogleNet	94%	54%	98%	92%	90%	94%	85%
LeNet	98%	52%	98%	86%	88%	82%	84%
FaceNet	100%	82%	100%	98%	100%	100%	97%
ResNet50	96%	64%	86%	88%	90%	90%	86%
AVERAGES	96%	64%	93%	90%	88%	91%	87%

As shown in Table 3, the DL models trained with unmasked face images and tested with masked images, FaceNet model is found to have highest accuracy of 82% whereas LeNet is found to have the lowest of 52%.

When the models are trained with half-masked and tested with masked images, FaceNet has the highest accuracy of 98% and VGG16 has the lowest of 80%. When the models are rained with masked images and tested with masked images, FaceNet has the highest accuracy of 100% and ResNet50 has lowest score of 90%.

In Table 3, we see that the highest average performance is 96% when the model is trained with unmasked images and tested with unmasked images.

We observe that the lowest average performance is 64% when the system is trained with unmasked faces and tested with masked faces. This shows that models trained with unmasked faces are not performing well with masked faces.

In Table 3, we see that the average accuracy of DL models decreases for testing masked face images when trained with unmasked images. The accuracy is found to be increased for testing masked face images when the DL model is trained with masked faces.

TABLE 4. OVERALL AVERAGE MISS RATES OF 6 DL MODELS FOR GTFD DATABASE

DL Algorithms Overall Miss Rates for All Datasets	UM	HM	M	Average
Average Unmasked Miss Rate	4.0%	7%	12%	8%
Average Masked Miss Rate	36.0%	10%	7%	18%
Averages	20.0%	8.5%	9.5%	12.7%

As shown in Table 3, FaceNet is found to outperform other models in identifying unmasked facial images for all 3 types of training datasets. FaceNet outperforms other models for identifying masked face images when trained with unmasked face images. FaceNet outperforms other models for identifying masked images when trained with masked or half-masked images.

In Table 4, we notice that while testing masked face images, models trained with masked face images have the lowest average miss rate of 7%, and the models trained with unmasked face images have the highest average miss rates of 36%. But, while testing unmasked face images, we observe that models trained with masked face images have the highest average miss rate of 12%, and the models trained with unmasked face images have the lowest average miss rate of 4%.

C. Evaluating Experimental Results

FaceNet stands out as the top performer in identifying unmasked facial images, followed by LeNet, regardless of the type of training dataset. FaceNet also outperforms other models in identifying masked face images when trained with unmasked face images for both ORL and GTFD databases. FaceNet consistently excels in identifying masked images when trained with masked or half-masked images for both databases. On the other hand, AlexNet, ResNet50, GoogleNet, and VGG16 exhibit weaker performance in specific scenarios.

If a DL recognition system needs to recognize both masked and unmasked images, the recommended configuration is to train with half-masked or masked face images and use the FaceNet classification model, as it achieves an average accuracy of at least 98% with both ORL and GTFD databases. Similarly, if the DL system needs to recognize only unmasked face images, training with unmasked face images and using FaceNet as the DL model yields 100% accuracy with both ORL and GTFD databases. On the other hand, if the focus is solely on recognizing masked face images, training with masked faces and employing FaceNet as the classification model yields 100% accuracy with both databases.

From Tables 3 and 5, it is observed that training a DL model with more masked images improves masked face recognition metrics while degrading unmasked face recognition performance. Optimal performance scores for both unmasked and masked face recognition are achieved when training the models with half-masked datasets. Masked face recognition reaches its highest accuracy when models are trained with all masked images, while unmasked face recognition achieves its highest accuracy when models are trained with all unmasked images.

VI. CONCLUSIONS AND FUTURE WORK

Our observations indicate that face recognition models intended primarily for unmasked faces show a decline in performance when recognizing masked face images. DL models do not perform in masked face recognition as much good as they do in unmasked face recognition in general. We observed a trend of increasing accuracy in identifying masked face images for DL models when trained with more masked face images, while a decrease is observed in identifying unmasked images.

We found that FaceNet performs exceptionally well with both unmasked and masked facial images in the DL domain. The choice of the appropriate model and training configuration depends on the specific requirements of the recognition system, whether it is for masked, unmasked, or both types of facial images.

We purposely design our experiments to work with images captured in a controlled environment like a studio setting to focus only on examining impact of masks avoiding interference of outside factors like various backgrounds, illuminations, and occlusions. Larger image databases like LFW are mostly composed of images collected from web with various environmental settings, sizes, resolutions, and illuminations which do not fit to our experimental design to explore impact of mask wearing in face recognition while avoiding other factors as much as possible. There is limited number of facial image databases which are captured in a controlled environment and composed of lesser number of images as in our selected ORL and GTFD databases.

Due to the limitation in the availability of masked face databases, we chose to use synthetic masked faces in our study. The ORL and GTFD databases were selected as they provided a controlled environment for conducting our experiments. Future work includes exploring the differences between real and synthesized masked images in face recognition, conducting experiments using images captured in both controlled and uncontrolled environments and with larger number of subjects for both real and synthetic images.

VII. ACKNOWLEDGEMENT

This research is funded by NSF Award #1900087. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF.

REFERENCES

[1] N. Damer, J. H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner and A. Kuijper, "The Effect of Wearing a Mask on Face Recognition Performance: An Exploratory Study," 2020 International Conference of the Biometrics Special Interest Group (BIOSIG), 2020, pp. 1-6.

[2] M. Opitz, G. Waltner, G. Poier, H. Possegger, and H. Bischof, “Grid loss: Detecting occluded faces,” in Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III, ser. Lecture Notes in Computer Science, vol. 9907. Springer, 2016, pp. 386–402.

[3] L. Song, D. Gong, Z. Li, C. Liu, and W. Liu, “Occlusion robust face recognition based on mask learning with pairwise differential siamese network,” in 2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019, 2019, pp. 773–782.

[4] A. Anwar , A. Raychowdhury, “Masked Face Recognition for Secure Authentication”, Computer Science > Computer Vision and Pattern Recognition, Aug 2020.

[5] S. Dharanesh and A. Rattani, "Post-COVID-19 Mask-Aware Face Recognition System," 2021 IEEE International Symposium on Technologies for Homeland Security (HST), 2021, pp. 1-7.

[6] D. Montero, M. Nieto, P. Leskovsky and N. Aginako, "Boosting Masked Face Recognition with Multi-Task ArcFace", Computer Science > Computer Vision and Pattern Recognition,Apr 2021.

[7] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “Arcface: Additive angular margin loss for deep face recognition,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[8] M. S. Ejaz, M. R. Islam, M. Sifatullah and A. Sarker, "Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition," *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp. 1-5, doi: 10.1109/ICASERT.2019.8934543.

[9] N. Damer, F. Boutros, M. Submilch, M. Fang, F. Kirchbuchner, and A. Kuijper, “Extended evaluation of the effect of real and simulated masks on face recognition performance”, The Institution of Engineering and Technology, May 2021.

[10] T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[11] S. Seneviratne, N. Kasthuriarachchi and S. Rasnayaka, "Multi-Dataset Benchmarks for Masked Identification using Contrastive Representation Learning," 2021 Digital Image Computing: Techniques and Applications (DICTA), 2021, pp. 01-08.

[12] S. Ge, J. Li, Q. Ye and Z. Luo, "Detecting Masked Faces in the Wild with LLE-CNNs," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 426-434.

[13] A. Maafiri and K. Chougdali, "Face Recognition using Wavelets based Feature Extraction and PCA-L1 norm," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 2019, pp. 1-4.

[14] A.Anwar and A.Raychowdhury. “What is MaskTheFace?”, MaskTheFace. https://sites.google.com/view/masktheface/home (Accessed Sep.25, 2022)

[15] M. Pudyel and M. Atay, "An Exploratory Study of Masked Face Recognition with Machine Learning Algorithms," SoutheastCon 2023, Orlando, FL, USA, 2023, pp. 877-882.

[16] K. S. Prado, “Face Recognition: Understanding LBPH Algorithm”, Towards Data Science, Nov 10, 2017.

[17] S. Ramachandran, A. V. Nadimpalli and A. Rattani, "An Experimental Evaluation on Deepfake Detection using Deep Face Recognition," 2021 International Carnahan Conference on Security Technology (ICCST),2021, pp. 1-6.

[18] N. Damer, F. Boutros, M. Submilch, M. Fang, F. Kirchbuchner and A. Kuijper,"Masked Face Recognition: Human vs. Machine”, Computer Science > Computer Vision and Pattern Recognition, March 2021.

[19] M. Atay, H. Gipson, T. Gwyn and K. Roy, "Evaluation of Gender Bias in Facial Recognition with Traditional Machine Learning Algorithms," 2021 IEEE Symposium Series on Computational Intelligence (SSCI), Orlando, FL, USA, 2021, pp. 1-7.

[20] J. Coe and M. Atay, “Evaluating Impact of Race in Facial Recognition across Machine Learning and Deep Learning Algorithms”, Computers 2021, 10(9), 113.

[21] B. W. Yohanes, R. D. Airlangga and I. Setyawan, “Real Time Face Recognition Comparison Using Fisherfaces and Local Binary Pattern”, In Proceedings of the 4th International Conference on Science and Technology (ICST), (2018), 1-5.

[22] T. Ojala, M. Pietikainen and T. Maenpaa, “Multiresolution Grayscale and Rotation Invariant Texture Classification with Local Binary Patterns”, In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, (July 2002), 971-987.

[23] Vanlalhruaia, Y. K. Singh and N. D. Singh, "Binary face image recognition using logistic regression and neural network," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 2017, pp. 3883-3888.