KSÜ Mühendislik Bilimleri Dergisi, 26(Özel Sayı), 2023 KSU J Eng Sci, 26(Special Issue), 2023
9th Uluslararası IFS Çağdaş Matematik ve Mühendislik Konferansı Özel Sayısı
9th International IFS Contemporary Mathematics and Engineering Conference Special Issue
Araştırma Makalesi Research Article

Kahramanmaras Sutcu Imam University
Journal of Engineering Sciences

Geliş Tarihi : 09.08.2023 Received Date : 09.08.2023
Kabul Tarihi : 25.09.2023 Accepted Date : 25.09.2023

MASKED AND UNMASKED FACE RECOGNITION ON UNCONSTRAINED FACIAL IMAGES USING HAND-CRAFTED METHODS

KISITLANMAMIŞ YÜZ GÖRÜNTÜLERİNDE EL YAPIMI YÖNTEMLERLE MASKELİ VE MASKESİZ YÜZ TANIMA

Ali TORBATİ[1] (ORCID: 0009-0005-5908-5840)
Önsen TOYGAR[1*] (ORCID: 0000-0001-7402-9058)

[1] Eastern Mediterranean University, Computer Engineering Department, Famagusta, North Cyprus, via Mersin 10, Türkiye

*Sorumlu Yazar / Corresponding Author: Önsen Toygar, onsen.toygar@emu.edu.tr

ABSTRACT

In this study, the face recognition task is applied on masked and unmasked faces using hand-crafted methods. Due to COVID-19 and masks, facial identification from unconstrained images became a hot topic. To avoid COVID-19, most people use masks outside. In many cases, typical facial recognition technology is useless. The majority of contemporary advanced face recognition methods are based on deep learning, which primarily relies on a huge number of training examples, however, masked face recognition may be investigated using hand-crafted approaches at a lower computing cost than using deep learning systems. A low-cost system is intended to be constructed for recognizing masked faces and compares its performance to that of face recognition systems that do not use masks. The proposed method fuses hand-crafted methods using feature-level fusion strategy. This study compares the performance of masked and unmasked face recognition systems. The experiments are undertaken on two publicly accessible datasets for masked face recognition: Masked Labeled Faces in the Wild (MLFW) and Cross-Age Labeled Faces in the Wild (CALFW). The best accuracy is achieved as 94.8% on MLFW dataset. The rest of the results on different train and test sets from CALFW and MLFW datasets are encouraging compared to the state-of-the-art models.

Keywords: Masked face recognition, unmasked face recognition, hand-crafted methods.

ÖZET

Bu çalışmada, maskeli ve maskesiz yüzlerde el yapımı yöntemler kullanılarak yüz tanıma görevi uygulanmıştır. COVID-19 ve maskeler nedeniyle, kısıtlanmamış görüntülerden yüz tanıma önemli bir konu haline gelmiştir. COVID-19'dan kaçınmak için çoğu insan dışarıda maske kullanmaktadır. Birçok durumda, tipik yüz tanıma teknolojisi işe yaramaz. Çoğu çağdaş ileri yüz tanıma yöntemi derin öğrenmeye dayanır ve büyük ölçüde birçok eğitim örneğine dayanır, ancak maske takılmış yüz tanıma, derin öğrenme sistemlerini kullanmaktan daha düşük bir hesaplama maliyeti ile el yapımı yaklaşımlar kullanılarak araştırılabilir. Maske takılmış yüzleri tanımak için düşük maliyetli bir sistem oluşturmak ve maske kullanmayan yüz tanıma sistemlerinin performansını karşılaştırma amaçlanmıştır. Önerilen yöntem, öznitelik düzeyi kaynaşım stratejisi kullanarak el yapımı yöntemleri birleştirir. Bu çalışma, maske takılmış ve takılmamış yüz tanıma sistemlerinin performansını karşılaştırmaktadır. Deneyler, maskeli yüz tanıma için erişime açık iki veri kümesi Masked Labeled Faces in the Wild (MLFW) ve Cross-Age Labeled Faces in the Wild (CALFW) üzerinde gerçekleştirilmiştir. En iyi doğruluk oranı, MLFW veri kümesinde %94,8 olarak elde edilmiştir. CALFW ve MLFW veri kümelerinden farklı eğitim ve test kümeleri kullanılarak elde edilen diğer sonuçlar, mevcut en iyi modellere göre cesaret vericidir.

Anahtar Kelimeler: Maskeli yüz tanıma, maskesiz yüz tanıma, el yapımı yöntemler.

ToCite: TORBATİ, A., & TOYGAR, Ö., (2023). MASKED AND UNMASKED FACE RECOGNITION ON UNCONSTRAINED FACIAL IMAGES USING HAND-CRAFTED METHODS. *Kahramanmaraş Sütçü İmam University, Journal of Engineering Sciences*, 26(Özel Sayı), 1133-1139.

INTRODUCTION

The study of computer vision involves developing algorithms and models that can allow computers to interpret, analyze, and understand visual data from the world around them. It involves techniques for analyzing, processing, and understanding images and video data, as well as techniques for recognizing patterns, objects, and features in this data. There are some key differences in the way that computers and humans process images. One major difference is that computers can process images much more quickly and accurately than humans, making them well-suited for tasks such as analyzing large datasets of images or detecting patterns and features in images.

Computers may use a technique called face recognition to identify and authenticate people in images or streaming videos. The ability of computers is what makes it possible to analyze and understand visual data, and it involves developing algorithms and models that can extract and analyze features from images of faces. When it comes to recognizing faces, one of the major obstacles is being able to accurately identify and distinguish one face from another, even when there are variations in lighting, pose, expression, and other factors. To achieve this, face recognition systems typically use a combination of machine learning algorithms and image processing techniques to analyze and compare the unique features of different faces.

Machine learning's ultimate aim is to make it so computers can teach themselves new skills by analyzing existing data and make inferences or choices without being explicitly programmed. It's predicated on the premise that computers can analyze data, spot trends, and decide what to do next.

In this approach, three different hand-crafted methods have been used in order to do the face recognition task, namely Principal Component Analysis (PCA) (Rani et al., 2022), Local Binary Patterns (LBP) (Kulkarni et al., 2017) and Histogram of Oriented Gradients (HOG) (Ahamed et al., 2018). All the experimental works are conducted on state-of-the-art datasets, namely Cross-Age Labeled Faces in the Wild (CALFW) (Zheng et al., 2017) and Masked Labeled Faces in the Wild (MLFW) (Wang et al., 2022). Sample images from both datasets are shown in Figure 1. We propose a feature-level fusion approach using LBP and HOG for the recognition of masked and unmasked facial images.

In the rest of the paper, PCA, LBP and HOG feature extraction methods are explained in the next section. After that, the proposed method using feature-level fusion strategy is discussed. Experimental results and comparison with the state-of-the-art models are given in the next section and finaly, the conclusions are stated in the last section of this paper.

Figure 1. Sample Images from CALFW (top) and MLFW (bottom) Datasets

FEATURE EXTRACTION AND RECOGNITION

In general, feature extraction is the reduction of the complexity of the input images while preserving the most important information for a specific task. There are many feature extraction techniques such as:

• Hand-crafted feature extraction methods that use one's own experience and knowledge to determine which traits are important for a certain job. Examples include Principal Component Analysis, Histograms of Oriented Gradients, Speeded Up Robust Features (SURF), Linear Discriminant Analysis (LDA), Local Binary Patterns, Scale-Invariant Feature Transform (SIFT) and Gabor Filter.

• Deep Learning-based feature extraction approaches that directly discover features from the data using deep neural networks. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are used often in this circumstance.

• Hybrid feature extraction methods that are about combining traditional feature extraction with that performed by deep neural networks.

Hand-crafted feature extraction methods used in this study are Principal Component Analysis, Local Binary Patterns and Histogram of Oriented Gradients. PCA is a widely used method for feature extraction over a several disciplines, including but not limited to pattern recognition, computer vision, and image processing. The original data are transformed using a linear method into a reduced-dimensional subspace, in which the directions with the highest variance in the data are retained, and the others are discarded (Rani et al., 2022).

In the areas of pattern recognition, computer vision, and image processing, LBP is implemented as a feature extraction approach. It is particularly useful for texture analysis and image classification tasks. LBP's underlying premise is generating a binary pattern for each pixel in a picture based on the surrounding pixels' intensity levels. The LBP operator contrasts a central pixel's intensity value with the intensity values of the pixels around it, and the output is a binary code that represents the local texture around that pixel (Kulkarni et al., 2017).

The feature extraction technique known as HOG is widely utilized in computer vision and pattern recognition, especially for object detection tasks. The primary premise of HOG is to compute the distribution of gradient orientations in an image, which captures the shape and structure of an object (Ahamed et al., 2018).

PROPOSED METHOD

Feature extraction, normalization, standardization, matching and classification are the main components of the proposed method used in this study.

The LBP and HOG approaches are combined to create the feature-level Fusion strategy employed in this research. Initially, HOG and LBP features are retrieved from each image using this handcrafted approach. As we extract different type of features using our methods, normalization is applied to both feature vectors. It helps our data points to be in the same domain so it would be possible to concatenate and compare them with each other. After normalization, concatenation is applied to both normalized feature vectors in order to get a concatenated fused feature vector. Finally, the fusion vector related to each image in the training set is used in the matching process with the test images. Block diagram of the proposed method is shown in Figure 2.

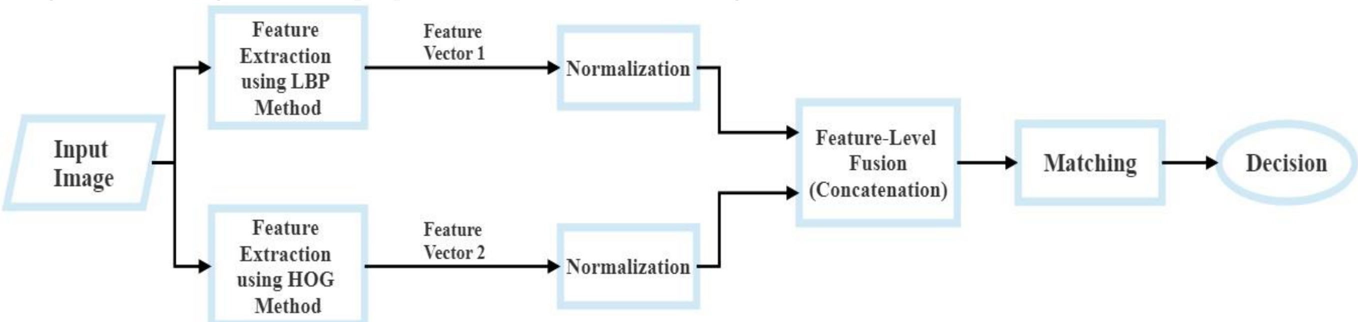

Figure 2. Block Diagram of the Proposed Method

EXPERIMENTAL RESULTS

The experiments are implemented using MATLAB 2022a by a computer with Intel i7 processor (10th gen.) and 16 GB of RAM. The experiments are typically organized into three sections to further illustrate the performance. The first section involves training and testing with masked images of faces. The second section involves training with masked face images and testing with the unmasked ones. The third experiments use unmasked and masked face images for training and evaluation respectively.

Experiments using Different Datasets

The outcomes of the tests, which are carried out on several datasets, are as follows. Every feature extractor is applied on each database in three different experimental setups. Features from each and every train set are retrieved, and a handcrafted classifier is used to match those features. In the first experimental setup, all of the images are from MLFW dataset with 500 images used for train and 500 images used for testing. The total number of images that is used for train and test in the second and the third experimental setups are equal to the numbers used in the first experiments. However, for the second experimental setup, train images are chosen from MLFW dataset, test images are chosen from CALFW dataset and in the 3rd experimental setup it is vice versa. Information about datasets used in this study for three experimental setups are shown in Table 1.

Table 1. Information about Datasets Used in Experimental Setups

Experimental Setup	Train Set	Train Image per Person	Total number of Train Images	Test Set	Test Image per Person	Total number of Test Images	Total number of Individuals	Total number of Images
I	MLFW	1	500	MLFW	1	500	500	1000
II	MLFW	1	500	CALFW	1	500	500	1000
III	CALFW	1	500	MLFW	1	500	500	1000

Experimental Setup I

In the first experimental setup, all feature extraction methods are employed for the masked face recognition after being trained on images of masked faces. Moreover, images for this experimental setup have been chosen from the MLFW dataset. The final results have been measured in both cases so that we can train the program with the first image and test with the second one, or vice versa. Table 2 illustrates the observations.

Table 2. Recognition Rates Using the Experimental Setup I on MLFW Dataset

Feature Extractor	Train Image (MLFW)	Test Image (MLFW)	Accuracy
PCA	1st	2nd	70.2 %
LBP	1st	2nd	93.2 %
HOG	1st	2nd	93.2 %
Proposed Method	1st	2nd	94.2 %
PCA	2nd	1st	68.8 %
LBP	2nd	1st	93.6 %
HOG	2nd	1st	92.8 %
Proposed Method	2nd	1st	94.8 %

It is apparent that in this study, the results from the Proposed, LBP, and HOG methods are all above 90% and very similar to each other, however, the PCA method does not perform well as these three methods. The proposed method has the highest recognition rate in both scenarios, with 1% or more improvement over the other methods.

Experimental Setup II

In the second part of the experiments, all feature extraction methods are used to recognize unmasked faces after being trained on images of faces with masks. Masked face images are chosen from the MLFW dataset in this scenario, while unmasked face images are chosen from the CALFW dataset. Also, images are split evenly so that half are utilized for learning and the other half for testing. Table 3 displays the findings.

Table 3. Recognition Rates Using the Experimental Setup II on MLFW and CALFW Datasets

Feature Extractor	Train Image (MLFW)	Test Image (CALFW)	Accuracy
PCA	Masked	Unmasked	45.6 %
LBP	Masked	Unmasked	57.6 %
HOG	Masked	Unmasked	57.6 %
Proposed Method	Masked	Unmasked	64.8 %

It can be seen that the proposed method has the highest recognition rate with over 7% improvement compared to the accuracy of the other feature extraction methods. Following the proposed method, LBP and HOG methods have the same recognition rate, and finally, PCA method is at the bottom with a recognition rate under 50%.

Experimental Setup III

In the third phase of the experiments, all feature extraction algorithms are utilized to distinguish masked faces following training on images of unmasked faces. In this situation, face images with masks are selected from the MLFW dataset, whereas face images without masks are selected from the CALFW dataset. Table 4 provides a summary of the results.

Table 4. Recognition Rates Using the Experimental Setup III on CALFW and MLFW Datasets

Feature Extractor	Train Image (CALFW)	Test Image (MLFW)	Accuracy
PCA	Unmasked	Masked	30.4 %
LBP	Unmasked	Masked	58.0 %
HOG	Unmasked	Masked	58.0 %
Proposed Method	Unmasked	Masked	58.6 %

In the third experimental setup, when algorithms initially trained by unmasked images and then evaluated on masked ones, the proposed method, along with LBP and HOG methods, produce the best results with a recognition rate of near 60%.

Comparison with State-of-the-Art Models

At the conclusion of these experimental activities, the proposed method results relevant to this study are compared with various state-of-the-art models. These open sourced state-of-the-art deep face recognition methods are as follows: (1) ResNet50 model trained on a private Asia face dataset (Wang et al., 2021) with ArcFace (Deng et al., 2019), (2) ResNet50 model trained on CASIAWebFace database (Yi et al., 2014) with ArcFace (Deng et al., 2019), (3) ResNet50 model trained on VGGFace2 database (Cao et al., 2018) with ArcFace (Deng et al., 2019), (4) ResNet100 model trained on MS1MV2 database (Guo et al., 2016) refined by insightface with ArcFace (Deng et al., 2019), (5) ResNet100 model trained on MS1MV2 database (Guo et al., 2016) with CurricularFace (Huang et al., 2020), (6) ResNet100 model trained on MS1MV2 database (Guo et al., 2016) with SFace (Zhong et al., 2021).

All of these state-of-the-art methods are compared to the proposed method in this study in terms of accuracy using Cross-age LFW (CALFW) and Masked LFW (MLFW) datasets. It is important to note that, for the proposed method, CALFW and MLFW datasets are used in different cases in order to train the model, while for those deep learning methods, different benchmarks using various loss functions were used for this purpose. The comparison results are shown in Table 5.

In the masked face recognition task, the proposed method in this approach has better performance compared to the deep learning methods presented in Table 5. Overall, hand-crafted methods are easier to implement while they need less powerful hardware next to taking less amount of time for doing the different tasks in comparison with the deep learning methods that need much more powerful hardware and also take more time to be implemented and do different tasks. It is clear that deep learning methods are the latest technology that is employed for doing different tasks but it does not mean that it is always better than the traditional or older methods as shown in this study. It is critical to note that MLFW dataset was used as a training set in this study, as opposed to other datasets which have high computational cost and computation time for training deep learning models.

Table 5. Comparison of the Proposed Method with the State-of-the-Art Models

Reference	Method	Train Set	Accuracy on CALFW	Accuracy on MLFW
Wang et al. (2021)	ResNet50 (ArcFace)	Private-Asia	91.12 %	74.85 %
Yi et al. (2014)	ResNet50 (CosFace)	Casia-WebFace	92.43 %	82.87 %
Cao et al. (2018)	ResNet50 (ArcFace)	VGGFace2	93.72 %	85.02 %
Guo et al. (2016)	ResNet100 (ArcFace)	MS1MV2	95.83 %	90.13 %
Guo et al. (2016)	ResNet100 (Curricularface)	MS1MV2	95.97 %	90.60 %
Guo et al. (2016)	ResNet100 (SFace)	MS1MV2	95.83 %	90.63 %
Proposed Method	Feature-level Fusion of HOG and LBP	MLFW	64.80 %	94.80 %
		CALFW	52.80 %	58.60 %

CONCLUSION

In this approach, an in-depth examination of various analyses of facial recognition algorithms are undertaken by measuring the efficiency of four methods for extracting features, namely Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), Principal Component Analysis (PCA) and a custom-designed proposed method using the feature-level fusion of LBP and HOG methods. Furthermore, a handcrafted classifier is applied to assess the precision of the recognition procedure. The experiments are conducted using two benchmark datasets namely, the Cross-Age LFW (CALFW) and Masked Labeled in the Wild (MLFW). These datasets are selected because they are widely employed in the area of facial recognition and offer a trustworthy assessment of the performance of the systems under investigation. The findings are analyzed and compared to determine the effectiveness of each method. The study also aims to identify any potential limitations and areas for further improvement in the field of face recognition.

This study indicates that when it comes to recognizing masked faces using a training set of masked face images, the proposed method outperformed the other feature extraction methods, while HOG and LBP methods have nearly similar performance to the proposed method. In the case of recognizing masked faces using a training set of unmasked face images, the proposed method has the highest performance, followed by Local Binary Patterns and Histogram of Oriented Gradients methods, which have comparable accuracy levels. When it comes to recognizing unmasked faces using a training set of masked faces, the proposed method again is the most accurate one, and the performance of the Local Binary Patterns and Histogram of Oriented Gradients methods are equivalent.

It is important to note that, machine learning and deep learning methodologies are constantly evolving. Regarding future work, it would be beneficial to use more benchmark unmasked and masked face image datasets that include a variety of illumination angles and rotation variations, as well as incorporating other feature extraction methods based on deep learning techniques.

REFERENCES

Ahamed, H., Alam, I. and Islam, M. M. (2018), HOG-CNN Based Real Time Face Recognition, 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), 2018, pp. 1-4, doi: 10.1109/ICAEEE.2018.8642989.

Cao, Q., Shen, L., Xie, W., Parkhi, O. M. and Zisserman, A. (2018), VGGFace2: A Dataset for Recognising Faces across Pose and Age, 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018, pp. 67-74, doi: 10.1109/FG.2018.00020.

Deng, J., Guo, J., Xue, N. and Zafeiriou, S. (2019), ArcFace: Additive Angular Margin Loss for Deep Face Recognition, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 4685-4694, doi: 10.1109/CVPR.2019.00482.

Guo, Y., Zhang, L., Hu, Y., He, X., and Gao, J. (2016), MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition, Sep. 17, 2016. https://link.springer.com/chapter/10.1007/978-3-319-46487-9-6

Huang, Y. et al. (2020), CurricularFace: Adaptive Curriculum Learning Loss for Deep Face Recognition, 2020 IEEE/CVF Conference on Computer Vision andPattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 5900-5909, doi: 10.1109/CVPR42600.2020.00594.

Kulkarni, O. S., Deokar, S. M., Chaudhari, A. K., Patankar, S. S. and Kulkarni, J. V. (2017), Real Time Face Recognition Using LBP Features, 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), 2017, pp. 1-5, doi: 10.1109/ICCUBEA.2017.8463886.

Rani, G. E., Suresh, S. M, M. P., Abhiram, M., Surya, K. J. and Kumar, B. Y. A. N. (2022), Face Recognition Using Principal Component Analysis, 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2022, pp. 932-936, doi: 10.1109/ICACITE53722.2022.9823434.

Wang, C., Fang, H., Zhong, Y., and Deng, W. (2022), MLFW: A Database for Face Recognition on Masked Faces, In: et al. Biometric Recognition. CCBR 2022. Lecture Notes in Computer Science, vol 13628. Springer, Cham. 2022. doi: 10.1007/978-3-031-20233-9-18.

Wang, Q., Zhang, P., Xiong, H., and Zhao, J. (2021), Face.evoLVe: A High-Performance Face Recognition Library, arXiv.org, Jul. 19, 2021. https://arxiv.org/abs/2107.08621v4

Yi, D., Lei, Z., Liao, S., and Li, S. Z. (2014), Learning Face Representation from Scratch, arXiv.org, Nov. 28, 2014. https://arxiv.org/abs/1411.7923v1

Zheng, T., Deng, W., and Hu, J. (2017), Cross-age LFW: A database for studying cross-age face recognition in unconstrained environments, CoRR, vol. abs/1708.08197, 2017.

Zhong, Y., Deng, W., Hu, J., Zhao, D., Li, X. and Wen, D. (2021), SFace: Sigmoid- Constrained Hypersphere Loss for Robust Face Recognition, in IEEE Transactions on Image Processing, vol. 30, pp. 2587-2598, 2021, doi: 10.1109/TIP.2020.3048632.