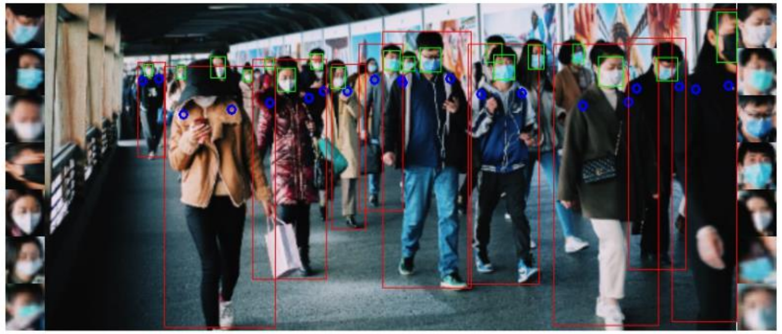
Deep Covariance Feature and CNN-based End-to-End Masked Face Recognition

Masum Shah Junayed, Arezoo Sadeghzadeh, Md Baharul Islam

Bahcesehir University, Istanbul, Turkey
American University of Malta, Bormla, Malta

15-18th December 2021

IEEE International Conference on Automatic Face and Gesture Recognition 2021, Jodhpur, India (Virtual Event)



The International Workshop on Face and Gesture Analysis for COVID-19 (FG4COVID19)

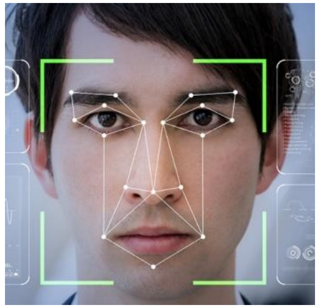
BAU BAHÇEŞEHİR ÜNİVERSİTESİ

Funded by TÜBİTAK

AUM AMERICAN UNIVERSITY OF MALTA

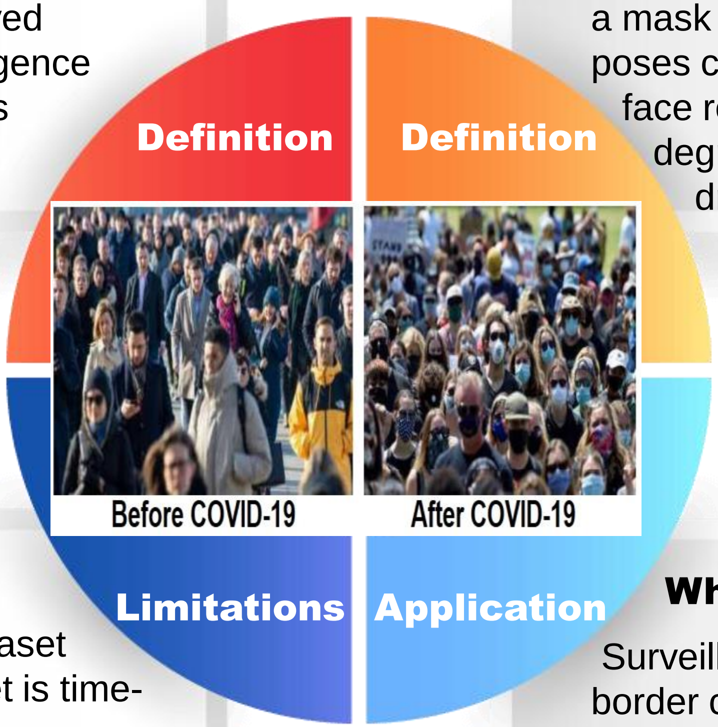
Contents

Introduction



Why face recognition?

It has been the topic of interest in the field of artificial intelligence and computer vision which achieved more attention with the emergence of COVID-19 as a contactless identity verification system.



Why masked face recognition?

During COVID-19, people are forced to wear facial masks in public places. As a mask covers 50-80% of a face, it poses challenges for the conventional face recognition systems and degrades their performance dramatically.

Existing Limitations?

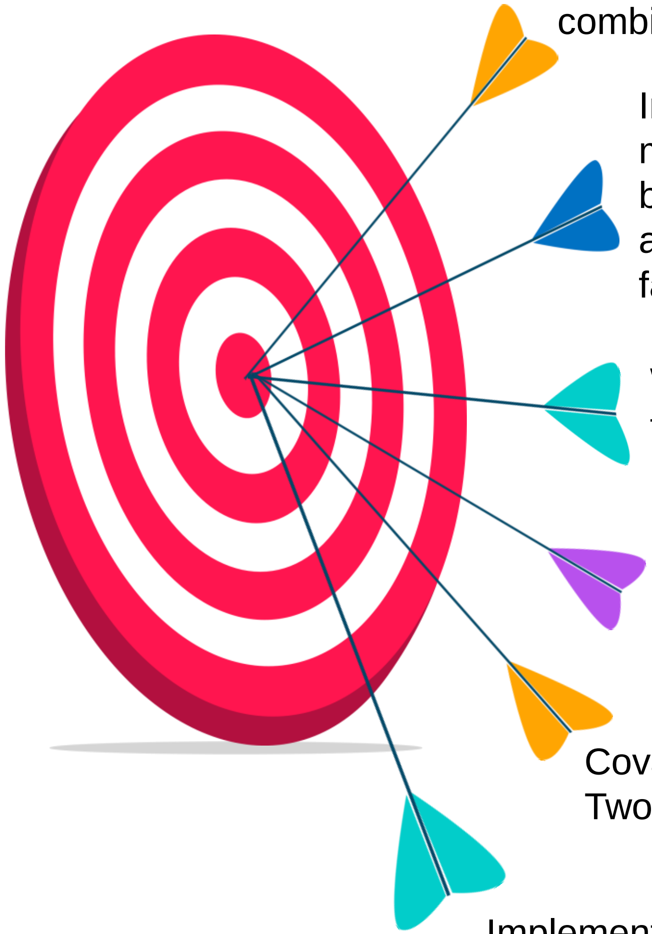
They rely on only masked dataset while collecting such a dataset is time- and energy-consuming.
They need face reconstruction which adds more computation burden.
Their accuracy is still unsatisfactory.

Which applications?

Surveillance control, facial attendance, border control gates, entrance into/exit from public communities, facial security check at the airports and train stations, facial authentication in smartphones.

Introduction

The Main Goal and Contributions of the Paper



An automated Masked Face Recognition (MFR) system is proposed based on the combination of a mask occlusion discarding technique and a deep learning model.

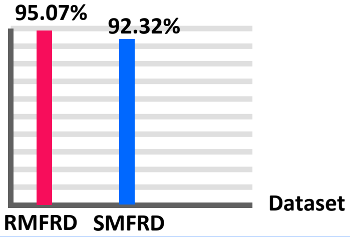
Instead of making an effort to reconstruct the occluded areas of the face by mask, we make full use of only occlusion-free areas in the faces. So, it can be applied for both masked and non-masked face recognition scenarios and overcome the challenge of collecting or synthesizing many masked faces for training.

We benefit from the shading augmentation and smoothing filter to improve the robustness of the system against illumination variations and noise, respectively.

Proposing a lightweight novel CNN-based feature extractor analyzed with two different optimizers and compared with two pre-trained models.

Covariance-based deep features are extracted followed by two layers of Bitmap and Eigenvalue. Two different classifiers are investigated for classification.

Implementing and evaluating on RMFRD and SMFRD datasets, our proposed system outperforms the state-of-the-art methods.



Literature Review

Controlled Environments

The conventional face recognition approaches perform successfully and accurately.

Improvement

The Performance of the Conventional Face Recognition Systems

Degradation

Uncontrolled Environments

These challenging environments, including illumination variations, pose variations, facial expressions, and occlusions extremely degrade their performance.

Literature Review

Facial Occlusions
e.g. eyeglasses, hats, hair, masks

Current Approaches

Image Reconstruction Approaches

Limitations:
- Overcomplete dictionary is required
- Computationally complex
- Low generalization

Occlusion Discarding Approaches

Process:
1) Occlusion detection
2) Using occlusion-free areas

Highest Accuracy:
95%

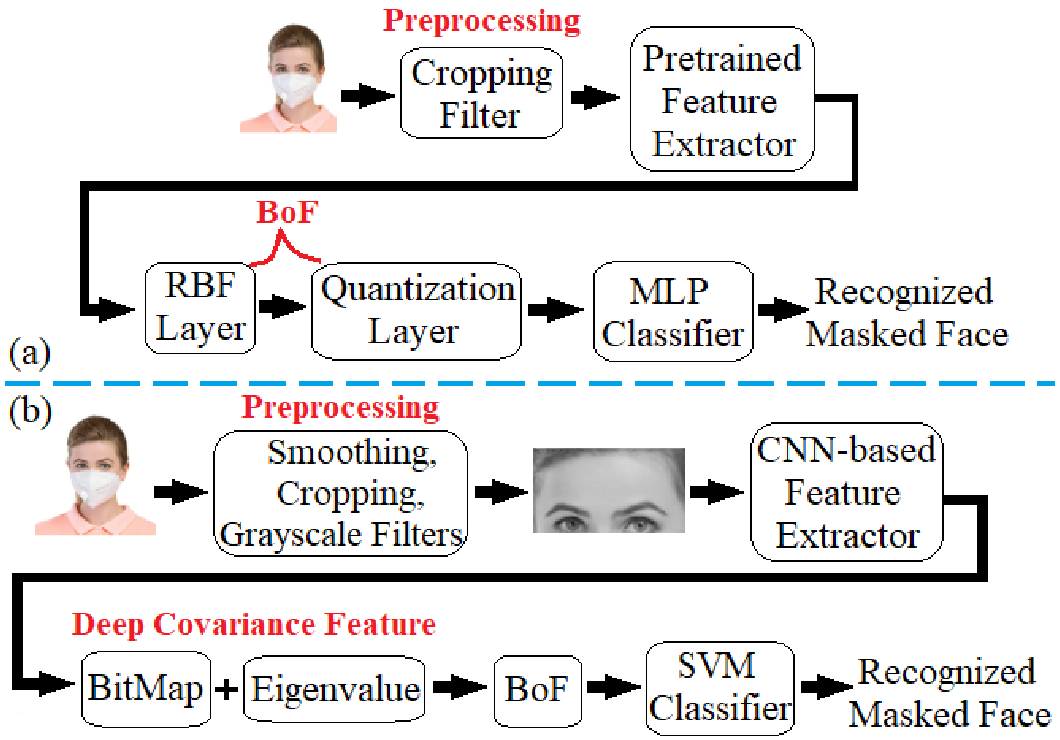
Limitations:
No baseline and detailed algorithm with unclear experimental and evaluation setup

Deep Learning Approaches

Highest Accuracy:
91.3%

Limitations:
Not satisfactory accuracy for the real-world applications

Proposed Method



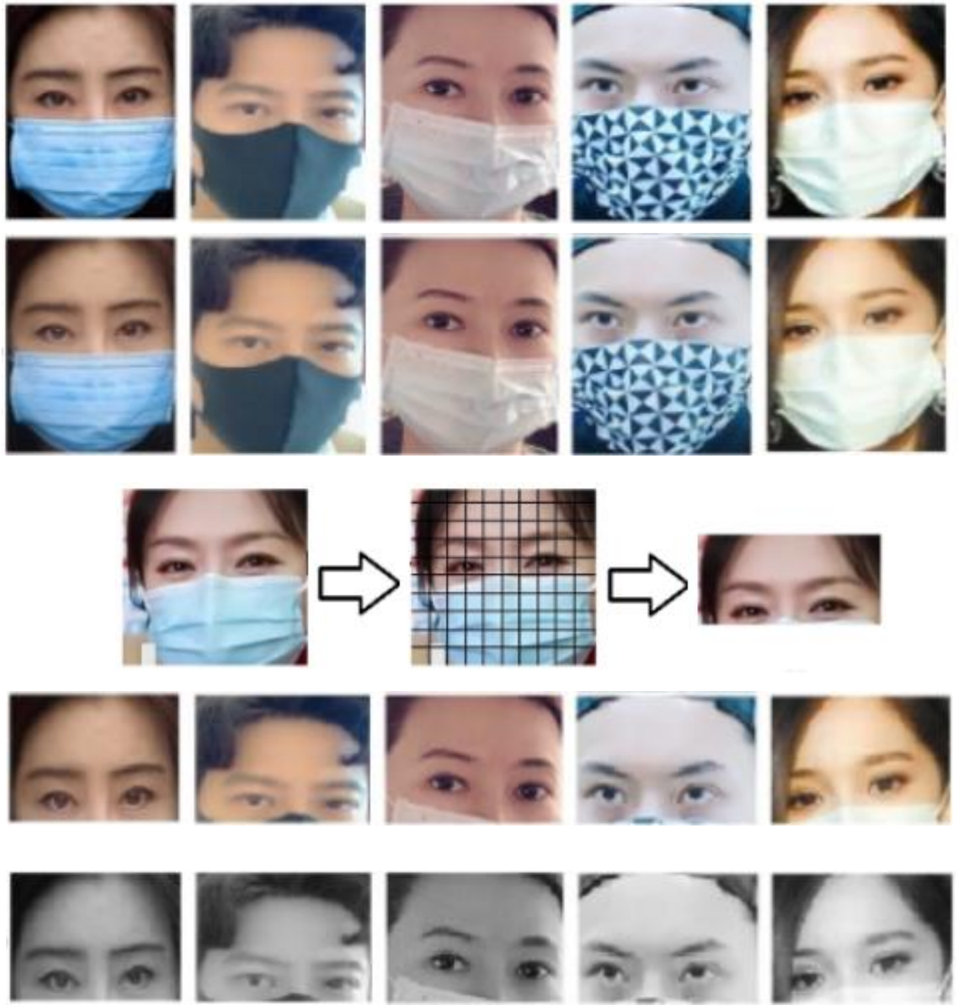
The central concept of the proposed model. Here, (a) and (b) represent the pipelines of the existing model proposed by Hariri et al and our proposed method, respectively, to clarify the main differences between these methods.

- W. Hariri. Efficient masked face recognition method during the covid-19 pandemic. arXiv preprint arXiv:2105.03026, 2021.

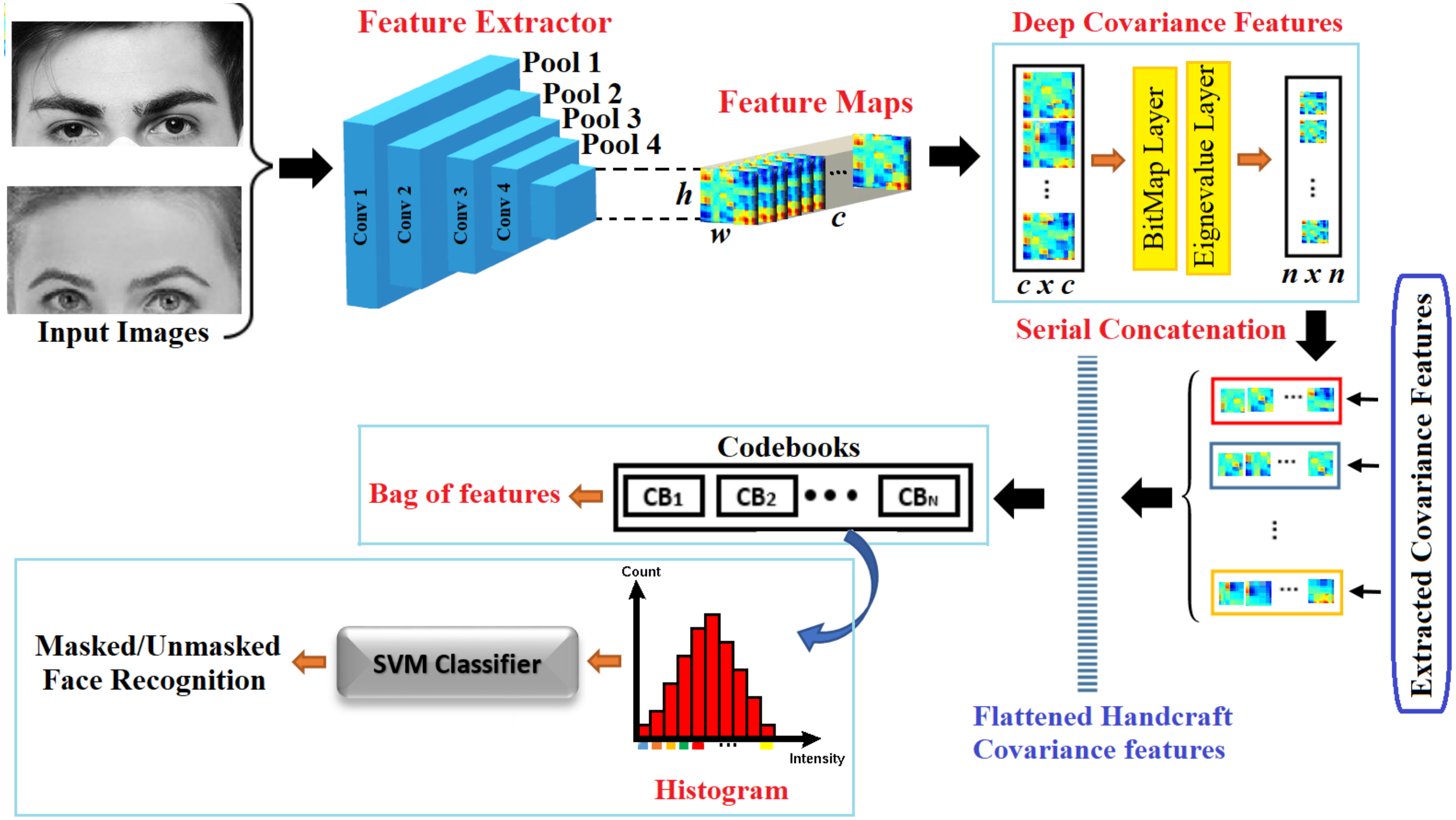
Proposed Method

Pre-processing

1. Original face images
2. Smoothing, resizing, and augmentation
3. Cropping Filter
4. Conversion to Grayscale



Proposed Method



The details of the proposed feature extractor.

Experimental Results and Comparison

Datasets

Datasets

Unmasked Faces

Ground-truth real world face images without images.

RMFRD

5000 masked faces and 90000 unmasked faces (out of which only 5000 images are used) from 525 identities

Real-World-Masked-Face-Recognition-Dataset

SMFRD

500000 simulated masked faces from 10000 identities collected from LFW and Webface

Simulated-Masked-Face-Recognition-Dataset

Experimental Results and Comparison

Model	Optimizer	Classifier	Accuracy (%)		Sensitivity (%)		Specificity (%)	
			RMFRD	SMFRD	RMFRD	SMFRD	RMFRD	SMFRD
VGG-16	SGD	MLP	91.28	87.44	89.42	87.85	92.98	87.04
		SVM	91.32	88.99	89.52	87.72	92.16	90.27
	ADAM	MLP	91.78	89.43	90.38	88.50	93.04	90.35
		SVM	91.32	89.08	91.09	88.39	91.53	89.74
AlexNet	SGD	MLP	91.24	89.43	90.91	88.39	91.53	90.43
		SVM	91.74	89.87	91.92	90.00	91.60	89.74
	ADAM	MLP	91.71	90.27	91.00	90.00	92.31	90.52
		SVM	92.20	90.67	91.92	90.83	92.44	90.52
Ours	SGD	MLP	92.63	90.99	93.75	92.45	91.74	89.66
		SVM	93.09	91.67	92.93	93.40	93.22	90.52
	ADAM	MLP	93.55	91.89	93.00	**95.10**	**94.55**	89.17
		SVM	**94.01**	**92.34**	**94.85**	93.46	93.33	**91.30**

Comparing performance of the proposed feature extractor with other pre-trained feature extractors with different optimizers on RMFRD and SMFRD datasets (no shading augmentation).

Experimental Results and Comparison

Illumination

Shading Effects (only on test set)

10% 20% 30% 40%

Test set (With Shading)						
Model	Accuracy (%)		Sensitivity (%)		Specificity (%)	
	RMFRD	SMFRD	RMFRD	SMFRD	RMFRD	SMFRD
VGG-16	64.72	50.26	63.93	50.19	64.99	50.78
AlexNet	66.38	51.02	67.47	52.56	65.84	50.81
Ours	**72.61**	**53.98**	**71.02**	**54.34**	**68.37**	**52.66**

The performance of the proposed system comparing to two other pre-trained models applying illumination changes only on test set, not training set.

Experimental Results and Comparison

Illumination

Shading Effects (on both training and test set)

10% 20% 30% 40%

Method	Accuracy (%)		Sensitivity (%)		Specificity (%)	
	RMFRD	SMFRD	RMFRD	SMFRD	RMFRD	SMFRD
Test set (Without Shading)						
VGG-16	92.16	90.13	92.58	89.83	91.88	90.52
AlexNet	93.88	91.67	93.97	91.44	93.55	92.93
Ours	**95.07**	**92.32**	**95.21**	**92.76**	**94.79**	**92.61**
Test set (With Shading)						
VGG-16	92.75	90.76	92.87	90.54	92.69	90.76
AlexNet	94.34	92.04	94.28	92.46	94.69	91.52
Ours	**95.84**	**93.46**	**95.97**	**93.31**	**95.25**	**93.70**

The performance of the proposed system comparing to two other pre-trained models applying illumination changes on both training and test sets.

Experimental Results and Comparison

The Effects of Covariance Features

Method	Accuracy (%)		Sensitivity (%)		Specificity (%)	
	RMFRD	SMFRD	RMFRD	SMFRD	RMFRD	SMFRD
VGG-16 [23]	87.78	84.47	86.60	84.72	87.96	84.11
AlexNet [17]	88.91	86.78	89.17	86.12	88.62	85.40
Ours (without)	**90.82**	**88.38**	**90.75**	**88.96**	**90.12**	**88.05**
Ours (with)	**95.07**	**92.32**	**95.21**	**92.76**	**94.79**	**92.61**

The performance of the proposed system comparing to two other pre-trained models without and with applying deep covariance features.

Experimental Results and Comparison

Robustness to Noise

Backbone	Accuracy		Sensitivity		Specificity	
	RMFRD	SMFRD	RMFRD	SMFRD	RMFRD	SMFRD
VGG-16	90.01	87.33	90.65	87.96	89.80	87.28
AlexNet	91.48	89.56	91.82	89.23	91.21	90.34
Ours	**93.10**	**90.91**	**92.73**	**91.06**	**92.98**	**89.67**

The performance of the proposed system comparing to two other pre-trained models on noisy test images.

Experimental Results and Comparison

Comparison

Comparing the performance of our proposed method with state-of-the-art methods on RMFRD and SMFRD datasets.

Approaches	Methods	Accuracy (%)	
		RMFRD	SMFRD
Luttrel et al	Transfer Learning	85.7	83.3
Almabdy et al	CNN+SVM	87.0	86.1
Hariri et al	CNN+BoF+MLP	91.3	88.9
Ours	CNN+Covariance+BoF+SVM	**95.07**	**92.32**

- J. Luttrell, Z. Zhou, Y. Zhang, C. Zhang, P. Gong, B. Yang, and R. Li. A deep transfer learning approach to fine-tuning facial recognition models. In2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), pages 2671–2676. IEEE, 2018.
- S. Almabdy and L. Elrefaei. Deep convolutional neural network-based approaches for face recognition. Applied Sciences, 9(20):4397, 2019
- W. Hariri. Efficient masked face recognition method during the covid-19 pandemic. arXiv preprint arXiv:2105.03026, 2021.

Conclusion

\+ Comparing the results with the similar state-of-the-art approaches, our MFR system improved the recognition accuracy with minimized computational cost.

\+ Our proposed model was robust against noise and illumination changes in the data benefiting from the smoothing filter and augmentation.

\- As this method is based on the features of the eyes and forehead, covering the eyes with sunglasses can reduce these features, resulting in an extreme reduction in the system's performance.

- Applying content-aware inpainting method on the occluded eyes can be used in our future works
- Using other pre-processing techniques rather than smoothing and removing the noise while enhancing the edges such as wrinkles, scars, etc.
- Performing our system on VISABORDER dataset and compare it with other works reported in the NIST.


Thank you for
your attention!