Face Mask Wearing Detection using Independent Component Analysis and Naïve Bayes Classifier

Ayman M Mansour
Computer and Communication Department,
Tafila Technical University,
'IS, Tafila, 66110,
Jordan

Received: August 7, 2021. Revised: March 10, 2022. Accepted: April 5, 2022. Published: May 4, 2022.

Abstract—In this paper, a new method was developed to detect the face mask wearing conditions using both Independent Component Analysis as a feature extractor and naïve Bayes as a classifies. The method was tested using real face images. The dataset was used for both training and testing. The MATLAB is used as programming software. The achieved accuracy in the testing 92.67%. The method was also tested using real live face pictures. The results were excellent, and the accuracy was 87.29%.

Keywords-Face Mask, Classification, Detection, Independent Component Analysis, Naïve Bayes Classifier

I. INTRODUCTION

After the spread of Corona disease and its classification as a pandemic. Several studies have emerged showing the rapid spread of the disease. This called for the necessity of social distancing and wearing a mask. Many individuals have committed to applying protocols to resist the spread of the epidemic. But there is a percentage of people who do not wear the mask or wear the mask in the wrong way. Many of them may show commitment when entering the institutions, stores, universities, schools or commercial malls, but as soon as they pass the initial checkpoint, these people remove the mask or wear it in the wrong way, exposing the nose and mouth.

This called for the existence of a monitoring system that detects the correct wearing of the face mask and sends an alert if the matter is not achieved. The alarm may be in a central monitoring unit linked to a group of cameras deployed in several places in the place which needs monitoring. Many work have been developed recently after COVID-19 [1-6]. The research is still going on in this field to get the best results and the highest accuracy.

In this paper we introduce the use of Independent component analysis (ICA) as a feature extractor in face mask detection. ICA have been applied and proven successful in many recent applications such in [7-13].

Independent Component Analysis is a statistical method that has been used in many applications. It is main task to get independent component factors that describes the structure of data [14-17]. Then, the data can be described as a linear combination of these independent components [16]. Theses independent components are statistically independent of each other. It has been used in Blind Source Separation, Multi-input Multi output systems, wireless communication, biomedical applications, and signal processing. The ICA is distinguished from other methods in its ability to deal with non- gaussian and statistically independent components of data [17].

The developed approach is based on creating Independent Components (IC's) bank which is called as basis functions of face masks images. The resulted bank is used in both training and testing stages for feature extraction. Then Naïve bayes classifier is used to classify the images either as with mask or no mask. Naïve bayes classifier is considered one of the best classifiers and has been used in many recent applications as in [17-27]. It has proven highly effective in the classification process and with low complexity compared to other methods.

The of the paper is organized such that: section 2 describe methodology and system Model and shows the model implementation. Section 3 discusses and analyzes the experimental procedures and the achieved result. Finally, the conclusion section is presented.

II. METHODOLOGY AND SYSTEM MODEL

In order to develop a system that can handle the classification process how people do wear the face mask correctly or not, a group of photos of people who wore the face mask were used in several ways, including correct and incorrect, and some of them without a face mask. A label has been added manually to each face picture to specify the class of mask wearing as " Mask "or " No Mask" in order to be used in the experimental part.

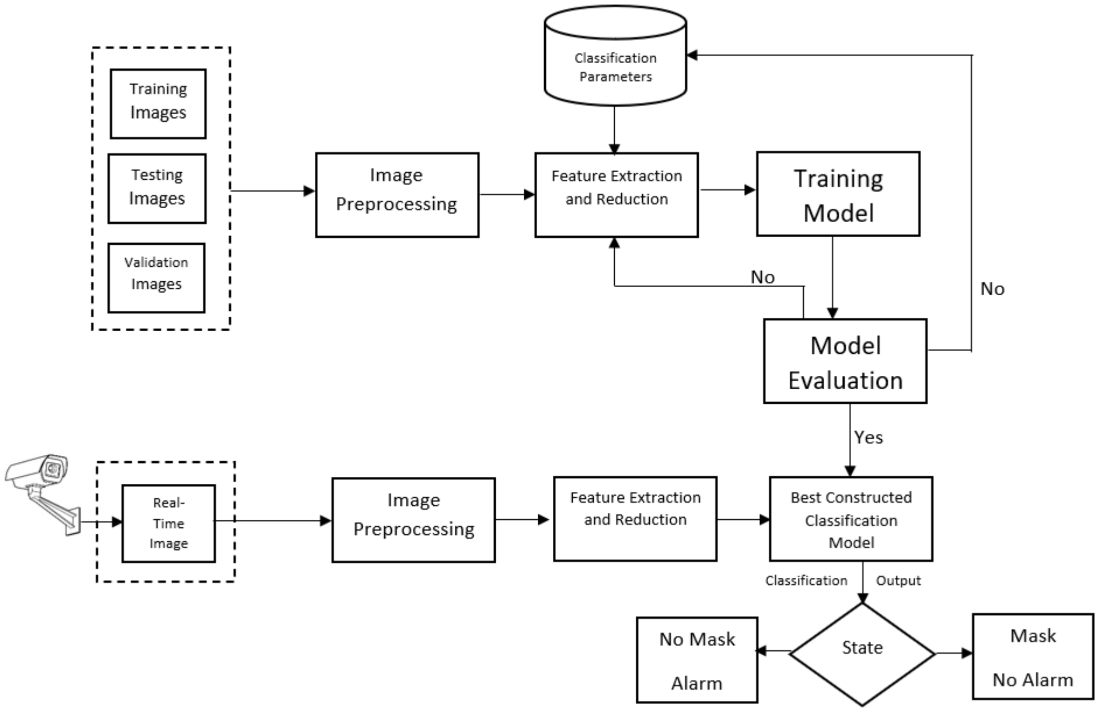

Figure 1. The developed system block diagram.

The collected pictures are divided into three groups, a group for training, a group for testing, and a group for verification. Image processing is performed as an initial stage to extract the features and reduce the dimensions of these features after determining a set of influencing factors.

The resulting features are then used to train the system to obtain the basis image bank, which is the backbone of the classification system. The ability of the proposed system to classify efficiently and with high accuracy is ensured. If the required accuracy is not obtained, the features and influence factors are referred to and factors are re-selected until the best ready-to-use model is reached.

The system is then used with the presence of a direct live camera, the image is directly processed, and the features are extracted from it. The image is distinguished by the condition that the person wears the face mask. If the person wears the mask correctly, he is considered to be wearing the face mask. If he is not wearing it correctly or not wearing it completely, he is considered not wearing the face mask. An audible and a visual warning is issued. The process is shown in Figure 1.

Set of images is used to produce a bank of independent components of the face mask images. Another set of images is used for training for this task. The more images used, the more comprehensive the independent basis bank will be.

The basis image bank necessary is used to produce the classification model through several stages, starting with removing the background of the image and keeping only the face. This stage is the Extraction of the Region Of Interest (ROI). The ROI contains the unique features of the face; therefore, extracting this region led to increase in the detection accuracy. Then whitening is done to the generated matrix. Then, the PCA calculation is done to get Principle Components (PCs). The dimension is reduced by choosing the most important basis Principle Components. To produce a bank of Independent Components (ICs) for images by processing the PCs by ICA Model to estimate the ICs is shown in Figure 2.

Initially, a set of training images is used to train the image classifier to be able to distinguish images effectively and with high accuracy. The process begins with image processing, removing backgrounds, and keeping only the face. After that, the images are filtered with a bank of independent basis components to produce a set of images with the same size as the original images.

The standard deviation for each image is calculated. These values are considered as features of the image to be classified. The final stage is to obtain the optimal model used to classify images using naïve base classifier.

Bayes Classification model outperform many other algorithms and it can easily process large set of data. The model is constructed by analyzing the statistics previously found in training data in order to get the Probabilities of classes given a set of features. The process is shown in Figure 3.

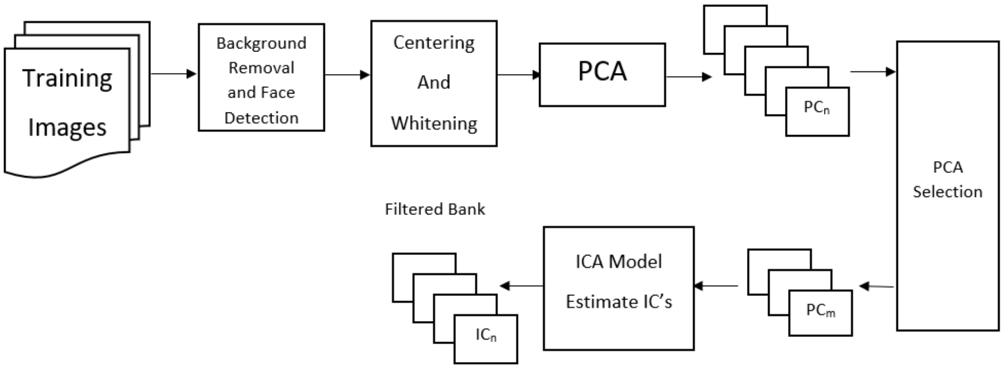

Figure 2. Production of ICs bank from training images

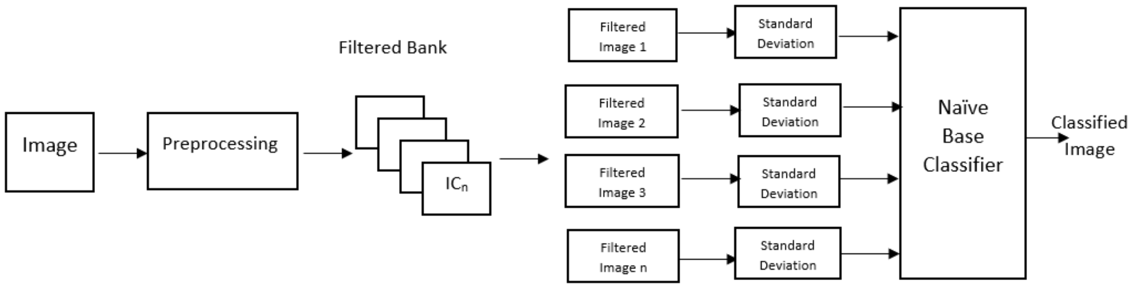

Figure 3. Classification Process using ICs bank ad Naïve bayes classifier

The centering is done by subtracting the expected value which is the average of all points from each point as shown in the following equation

$$E[D] = \overline{D} = \frac{1}{M}\sum_{i=1}^{M} D_i$$

$$D = D - \overline{D}$$

Then, the whitening is achieved using computing SVD of the centered data. The SVD produces a unity covariance matric of D. The SVD is computed using $D=V\ K^{-1/2}\ V^{T}\ D$
The whitening reduces the number of parameters to be estimated. Instead of having to estimate the n, it will be reduced to s n(n−1)/2 degrees of freedom.
Then, the eigen values are processed to ignore those that are too small. This is done be Principal Component Analysis
This is achieved by finding the orthogonal K of the equation
D=K S
Where S are considered independent
To estimate Independent Component by ICA algorithm after setting the number of j of independent component and choosing initialization vector K randomly, the K is updated using

$$K = E[Dg(K,x)] - E[g'(K,x)]K$$

Where, K should be orthogonalized as

$$K_j = K_j - \sum_{i=1}^{j-1} (K_j^T K_i\ )K_i$$

And then normalized as shown,

$$K = \frac{K+}{||K||}$$

Where K and K_{n-1} should be in the same direction and the process is repeated to get the j independent component. The functions g(t) and g'(t) are defined as

$$g(t) = \frac{dG(t)}{dt}$$

$$g'(t) = \frac{d^2G(t)}{dt^2}$$

Where, g is a nonlinear function such as kurtosis. Two non- quadratic functions are used which are,

G(t) =log(cosh(t))

$$G(t) = -e^{-\frac{t}{2}}$$

The standard deviation is calculated for each filtered image. The values of the standard deviation of each filtered image is calculated using the following equation:

$$Standard\ Deviation = \frac{1}{W-1}\sum_{i=1}^{W}(FI - E[FI])^2$$

Where ,

$$E[FI] = \overline{FI} = \frac{1}{W}\sum_{i=1}^{W} FI_i$$

The naïve bayes classifier is based on Bayesian classifier where the detected class out of M classes. This method assumes that the properties used to build the model are separate from each other, so changing the values of one property does not affect another property. It is one of the very fast algorithms when doing classification and therefore it is used in real time classification of data. The algorithm is based on selecting the output with maximum probability for a given set of inputs for all possible values of the class

$$C_M = \arg\max P(class|data)$$

Where the data d has n variables d= $d_1, d_2, d_3, \ldots, d_n$

$$= \frac{P(data|class)P(class)}{P\ (data)}$$

Where P(.) is also called class probability and P(|) is called conditional probability. Dropping the denominator, as the denominator remains constant for a given input, we can remove that term:

$$C_M = \arg\max P(data|class)\ P\ (class)$$

III. SIMULATION AND RESULTS

In this section, the result of the experimental part is presented to ensure the effectiveness of the develop method. The collected images are taken suing Samsung A50 camera and iPhone 8 camera. The size of the images is scaled to 600 x 600 pixels. MATLAB was used as programming software. The average processing time was 0.87 second per image using AMD Ryzen 7 5800 U processor.

To get the independent component bank, a set of face images is extracted from the collected dataset. The images will be processed, normalized, and made to have zero mean. This ensures that no face image overpower the basis set. Example of these images is shown in Figure 4. In this experiment 150 images were used to produce ICs bank.

Figure 4. Training images used to produce ICs bank.

The used ICA algorithm is based on FastICA algorithm developed by Hyvainen ([14], [15]). The FastICA algorithm has several desirable structures compared to existing ICA methods. Coverages is cubic (or at least quadratic) compared to standard ICA algorithms based on gradient methods, where the interaction is only linear. This leads to a very rapid convergence. Lack of step size parameters in FastICA makes the algorithm easier to use. Using FastICA, the individual components can be measured individually, equivalent to making a projection. This reduces the calculation load of the method especially when some of the independent components need to be measured, not all. FastICA is uniform, distributed, computerized, and requires minimal memory space. This faster ICA algorithm has been shown to be 10 to 100 times faster than other ICA algorithm

These image windows are used as input to the FastICA to extract the independent components bank. PCA is used as another stage after FastICA to get more dimension reduction by selecting only the eign-vectors of the largest L eign-values of the independent components bank matrix. The reduced dimension by PCA resulting 17 basis functions. The performance of PCA was compared with Chi-squared reduction as shown in Figure 5. The used performance measure is F1 measurement.

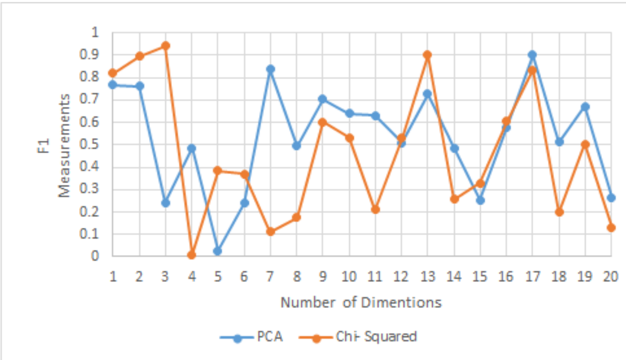

Figure 5. F1 Measurements for PCA and Chi-squared for different dimensions

Example of ICA basis bank which contains 42 ICS is shown in Figure 6.

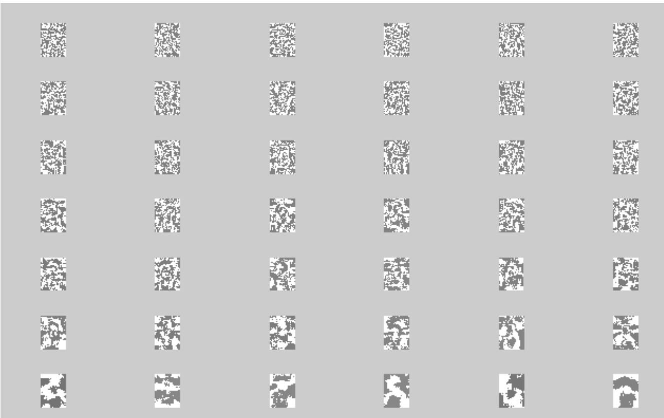

Figure 6. Example of ICs bank resulted from 150 image

Then, the developed method was applied to another 300 images to train naïve bayes classifier. Some of the pictures contain people wearing the mask correctly and

others not wearing it or even wearing it incorrectly. The pictures varied for people of different ages, with different appearance, such as the presence or absence of a beard.
In addition, face pictures of different genders were taken. 10- fold cross validation is used in this experiment. It is based on selecting a set of 10 records. 8 of the records are using for training, one for validation and the last for testing. The process is repeated 10 rounds with different sets for validation and testing as moving window. The 8 parts are merged to be used as a training set in each round.

Table 1. Accuracy for 2-Fold, 4-Fold, 8-Fold and the 10-Fold cross-validation.

K-Fold	Accuracy (%)
2-Fold	79.08%
4-Fold	83.40%
8-Fold	91.01%
10-Fold	92.67%

The confusion matrix of the 300 images was constructed based on testing data. The classes of confusion matrix are (Mask, No Mask). The confusion matrix has the form shown in Table 2,

Table 2. Confusion matrix.

Detected Classification	Real Classification	
	Mask	No Mask
Mask	143	15
No Mask	7	135

True Positive (TP) value is 150 person in actual wear mask and the model also is predicted these 143 Patients do wear mask . True Negative (TN) is 150 person in actual do not wear mask correctly or do not wear it ,and the model also has predicted that 135 person do not wear mask . False Positive (FP) is 15 persons in actual do not wear mask , whereas model says that these 15 person wear mask. False Negative (FN) is 7 person in actual do wear mask , but model predicts that these 7 person don't wear mask. The performance measurements result is shown in Table 3.

Table 3. Performance measurements results.

Measure	Equation	Result
Sensitivity	TPR = TP / (TP + FN)	0.9533
Specificity	SPC = TN / (FP + TN)	0.9000
Precision	PPV = TP / (TP + FP)	0.9051
Negative Predictive Value	NPV = TN / (TN + FN)	0.9507
False Positive Rate	FPR = FP / (FP + TN)	0.1000
False Discovery Rate	FDR = FP / (FP + TP)	0.0949
False Negative Rate	FNR = FN / (FN + TP)	0.0467
Accuracy	ACC = (TP + TN) / (P + N)	0.9267
F1 Score	F1 = 2TP / (2TP + FP + FN)	0.9286

Figure 7 illustrates the ROC space which is a plot of the ratio of true positives versus the ratio of false positives. This is done when some parameter of the classifier (for example threshold) is varied,

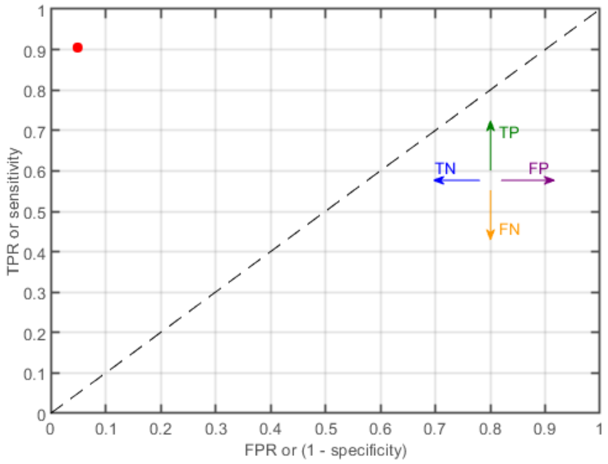

Figure 7. ROC space

The results obtained from the developed ICA-Naïve Bayes Classifier are compared by the results achieved using CNN, SVM, Decision Tree and using Neural Network. The accuracy of these methods is shown in Table 4 as in [28]. Using the developed methodology has a large classification improvement is observed compared to other methods in accuracy.

Table 4. Accuracy rates (%) using different classifiers and different image.

Method	Accuracy (%)
ICA-Naïve Bayes	95.7%
CNN	91.1%
Neural Network	86.02%
Decision Tree	83.35%
SVM	89.4%

The time analysis of the developed method for both training and testing stages compared with recent methods is shown in Figure 8. The outcomes show that the ICA-Naïve Bayes technique has provided effective time improvement comparing to other techniques.

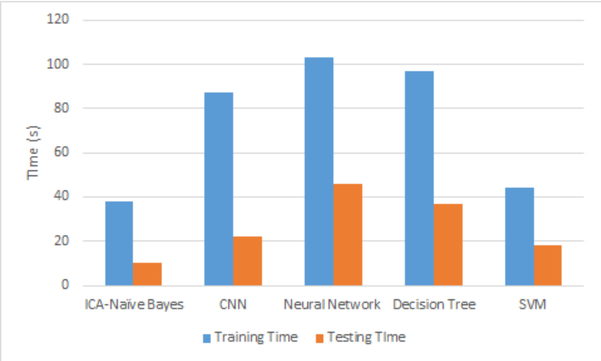

Figure 8. Time Analysis of ICA-Naïve Method compared to other methods

A graphical user interface was developed as shown in Figure 9. It has been tested using 100 new cases not previously seen and all the images are correctly classified.

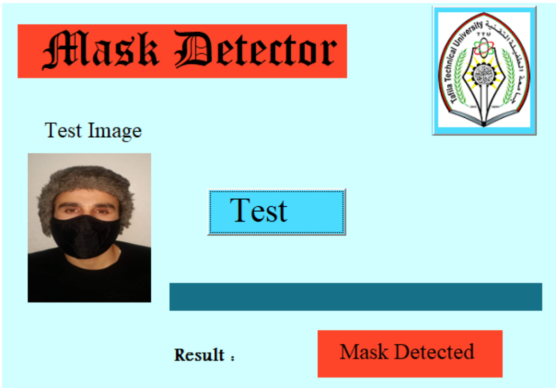

Figure 9. Graphical User Interface of the developed detection system

The developed algorithm was also tested with distorted images created from 300 original images. The generated images were also 300 images with balanced labels, where 150 images were used with the mask wore correctly and 150 images without the mask or with the mask wore incorrectly. Figure 10 shows an example of such mages.

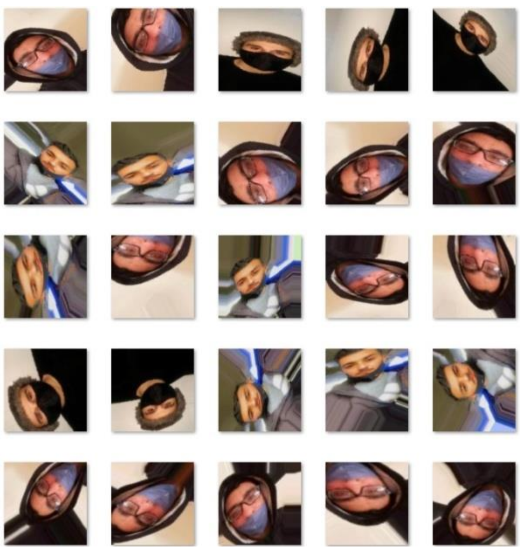

Figure 10. Distorted images used to test the system.

Table 5 summarized the confusion matrix of this new set and the resulted performance is shown in Table 6.

Table 5. Confusion matrix of distorted images.

Detected Classification	**Real Classification**		
		Mask	No Mask
	Mask	133	22
	No Mask	16	128

Table 6. Performance measurements results of distorted images.

Measurement	**Value**	**Measurement**	**Value**
Sensitivity	0.8926	False Positive Rate	0.1467
Specificity	0.8533	False Discovery Rate	0.1419
Precision	0.8581	False Negative Rate	0.1074
Negative Predictive Value	0.8889	F1 Score	0.8750
Accuracy	0.8729		

The accuracy rate of distorted images is considered acceptable compared to previous results of other methods shown in Table 4. The accuracy of the developed system using distorted images is 87.29 % when existing method such as ICA-Naïve Bayes, CNN, Neural Network , Decision Tree, SVM achieves 91.1% ,86.02%, 83.35%, 89.4% respectively using undistorted images.

IV. CONCLUSION

In this paper, we have proposed a face mask application. We have shown that the basic functions of ICA are able to capture the natural properties of face masks images. The proposed system demonstrates a high level of accuracy in obtaining test results. Another advantage of the proposed method is that the vector size of the feature is very small, and this results in a reduction in memory size and faster detection. The achieved excellent classification rate (92.67%.) which was obtained in the test ensures that the use of ICA basis functions is well suited for the detection of a face mask wearing conditions. So the proposed algorithm with naïve bayes classifier can be considered as reliable.

ACKNOWLEDGMENT

The authors wish to express their thanks for the support of Tafila Technical University.

References

[1] Z. Yu, X. Li, P. Wang and G. Zhao, "TransRPPG: Remote Photoplethysmography Transformer for 3D Mask Face Presentation Attack Detection," in IEEE Signal Processing Letters, vol. 28, pp. 1290-1294, 2021.

[2] X. Fan, M. Jiang and H. Yan, "A Deep Learning Based Light-Weight Face Mask Detector With Residual Context Attention and Gaussian Heatmap to Fight Against COVID-19," in IEEE Access, vol. 9, pp. 96964-96974, 2021.

[3] Y. Chen et al., "Face Mask Assistant: Detection of Face Mask Service Stage Based on Mobile Phone,"

in IEEE Sensors Journal, vol. 21, no. 9, pp. 11084-11093, 1 May1, 2021.

[4] B. Wang, Y. Zhao and C. L. P. Chen, "Hybrid Transfer Learning and Broad Learning System for Wearing Mask Detection in the COVID-19 Era," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-12, 2021, Art no. 5009612.

[5] J. Zhang, F. Han, Y. Chun and W. Chen, "A Novel Detection Framework About Conditions of Wearing Face Mask for Helping Control the Spread of COVID-19," in IEEE Access, vol. 9, pp. 42975-42984, 2021.

[6] D. Gonzalez Dondo, J. A. Redolfi, R. G. Araguás and D. Garcia, "Application of Deep-Learning Methods to Real Time Face Mask Detection," in IEEE Latin America Transactions, vol. 19, no. 6, pp. 994-1001, June 2021.

[7] Yang and N. Tavassolian, "An Independent Component Analysis Approach to Motion Noise Cancelation of Cardio-Mechanical Signals," in IEEE Transactions on Biomedical Engineering, vol. 66, no. 3, pp. 784-793, March 2019.

[8] Ayman M Mansour, Jafar Abukhait, Imad Zyout, “Speed Sign Recognition using Independent Component Analysis,” International Journal of Electrical, Electronics and Computer Systems (IJEECS), Vol. 17, no.01, pp. 832-838, November 2013.

[9] J. Tierney, J. Baker, D. Brown, D. Wilkes and B. Byram, "Independent Component-Based Spatiotemporal Clutter Filtering for Slow Flow Ultrasound," in IEEE Transactions on Medical Imaging, vol. 39, no. 5, pp. 1472-1482, May 2020.

[10] Z. Luo, C. Li and L. Zhu, "Full-Duplex Cognitive Radio Using Guided Independent Component Analysis and Cumulant Criterion," in IEEE Access, vol. 7, pp. 27065-27074, 2019.

[11] X. Shi, H. Yang, Z. Xu, X. Zhang and M. R. Farahani, "An Independent Component Analysis Classification for Complex Power Quality Disturbances With Sparse Auto Encoder Features," in IEEE Access, vol. 7, pp. 20961-20966, 2019.

[12] J. Ahmad, A. Akula, R. Mulaveesala and H. K. Sardana, "Probability of Detecting the Deep Defects in Steel Sample Using Frequency Modulated Independent Component Thermography," in IEEE Sensors Journal, vol. 21, no. 10, pp. 11244-11252, 15 May15, 2021.

[13] X. Wu, B. Zhou, Z. Lv and C. Zhang, "To Explore the Potentials of Independent Component Analysis in Brain-Computer Interface of Motor Imagery," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 3, pp. 775-787, March 2020.

[14] Aapo Hyvainen, Erkki Oja, Juha Karhunen. “Independent Component Analysis”, (1st ed.). New York: John wiley &sons, 2001.

[15] Aapo Hyvainen, and Erkki Oja. , “Independent component analysis: Algorithms and applications,” Neural Networks, vol. 13, pp. 411–430, 2000.

[16] P. Comon, “Independent component analysis, A new concept?” Signal Processing, vol. 36(4), pp. 287-314, 1994.

[17] Y. Shi, X. Lu, Y. Niu and Y. Li, "Efficient Jamming Identification in Wireless Communication: Using Small Sample Data Driven Naive Bayes Classifier," in IEEE Wireless Communications Letters, vol. 10, no. 7, pp. 1375-1379, July 2021.

[18] L. Dou, X. Li, H. Ding, L. Xu and H. Xiang, "iRNA-m5C_NB: A Novel Predictor to Identify RNA 5-Methylcytosine Sites Based on the Naive Bayes Classifier," in IEEE Access, vol. 8, pp. 84906-84917, 2020.

[19] Z. Xue, J. Wei and W. Guo, "A Real-Time Naive Bayes Classifier Accelerator on FPGA," in IEEE Access, vol. 8, pp. 40755-40766, 2020.

[20] W. He, Y. He, B. Li and C. Zhang, "A Naive-Bayes-Based Fault Diagnosis Approach for Analog Circuit by Using Image-Oriented Feature Extraction and Selection Technique," in IEEE Access, vol. 8, pp. 5065-5079, 2020.

[21] M. Thammawichai, "A Double-Layered Hybrid Approach for Network Intrusion Detection System Using Combined Naive Bayes and SVM," in IEEE Access, vol. 9, pp. 138432-138450, 2021.

[22] C. K. Aridas, S. Karlos, V. G. Kanas, N. Fazakis and S. B. Kotsiantis, "Uncertainty Based Under-Sampling for Learning Naive Bayes Classifiers Under Imbalanced Data Sets," in IEEE Access, vol. 8, pp. 2122-2133, 2020.

[23] D. Li, J. Sun, H. Yang and X. Wang, "An Enhanced Naive Bayes Model for Dissolved Oxygen Forecasting in Shellfish Aquaculture," in IEEE Access, vol. 8, pp. 217917-217927, 2020.

[24] K. Nsaif et al., "FRCNN-GNB: Cascade Faster R-CNN With Gabor Filters and Naïve Bayes for Enhanced Eye Detection," in IEEE Access, vol. 9, pp. 15708-15719, 2021.

[25] P. Valdiviezo-Diaz, F. Ortega, E. Cobos and R. Lara-Cabrera, "A Collaborative Filtering Approach Based on Naïve Bayes Classifier," in IEEE Access, vol. 7, pp. 108581-108592, 2019.

[26] Ayman M. Mansour, “Texture Classification using Naïve Bayes Classifier,” International Journal of Computer Science and Network Security (IJCSNS), vol. 18, no. 1, January 2018

[27] Mohammad A. Obeidat and Ayman M. Mansour, “EEG Based Epilepsy Diagnosis System using Reconstruction Phase Space and Naïve Bayes Classifier,” WSEAS Transactions on Circuits and Systems, Volume 17, 2018.

[28] G K Jakir Hussain et al,”, The Face Mask Detection Technology for Image Analysis in the Covid-19 Surveillance System” J. Phys.: Conf. Ser. 1916 012084, 2021

Dr. Ayman M Mansour received his Ph.D. degree in Electrical Engineering from Wayne State University in 2012. Dr. Mansour received his M.Sc degree in Electrical Engineering from University of Jordan, Jordan, in 2006 and his B.Sc degree in Electrical and Electronics Engineering from University of Sharjah, UAE, in 2004. He graduated top of his class in both Bachelor and Master. Currently, Dr. Mansour is an Associate Professor in the Department of Computer and Communication Engineering, Tafila Technical University, Jordan. He is also the director of the Energy Research Center at Tafila Technical University. His areas of research include Communication Systems, Multi-agent Systems, Fuzzy Systems, Data Mining and Intelligent Systems. He conducted several researches in his area of interest. Dr. Mansour is a member of IEEE, Michigan Society of Professional Engineers, IEEE Honor Society (HKN), Society of Automotive Engineers (SAE), Tau Beta Pi Honor Society, Sigma Xi and Golden Key Honor Society.