Improving recognition accuracy for facial expressions using scattering wavelet

Mehdi Davari[1,*], Aryan Harooni [2], Afrooz Nasr[3], Kimia Savoji[4], Masoumeh Soleimani[5]

[1]Department of Management, Dolatabad Branch, Islamic Azad University, Isfahan, Iran
[2]Department of Computer Science, Kent State University, Kent, OH, USA
[3]Department of Computer Engineering, Khorasgan Branch, Islamic Azad University, Isfahan, Iran
[4]School of Computing, Clemson University, SC, USA
[5]Department of Industrial Engineering, Clemson University, SC, USA

Abstract

One of the most evident and meaningful feedback about people's emotions is through facial expressions. Facial expression recognition is helpful in social networks, marketing, and intelligent education systems. The use of Deep Learning based methods in facial expression identification is widespread, but challenges such as computational complexity and low recognition rate plague these methods. Scatter Wavelet is a type of Deep Learning that extracts features from Gabor filters in a structure similar to convolutional neural networks. This paper presents a new facial expression recognition method based on wavelet scattering that identifies six states: anger, disgust, fear, happiness, sadness, and surprise. The proposed method is simulated using the JAFFE and *CK+* databases. The recognition rate of the proposed method is 99.7%, which indicates the superiority of the proposed method in recognizing facial expressions.

1. Introduction

Sensory communication is communication between humans that is much older than verbal communication. This communication is so deep and rooted that it is difficult to hide and deceive. He knows what is happening in his environment [1]. These reactions are clearly shown in the voice, the way the body moves, and the face. A person may not express his feelings honestly, but his facial expression, voice, and body movements show his feelings with little error. Detection of these states has many applications in psychological sciences, social network data processing, marketing systems, and intelligent education systems. One of the most apparent and meaningful feedback about people's emotions is facial expressions. Studies on the identification of emotions from faces date back to the end of the 20th century [2–5]. After that, extensive studies and research have been conducted to respond to the challenges and problems of the systems. In these studies, several classifications have been made for facial expressions, among which angry, happiness, sad, surprised, standard, and surprised can be named[6] and to improve the recognition rate of these states, machine learning methods and techniques such as local binary patterns, directional gradient histograms [7] and support vector machine and k-nearest neighbor [8] were used. Some of these methods have not been successful, so studying facial expression recognition is still considered one of the most exciting topics in machine learning and pattern recognition. In 2012 [9], Niz et al. proposed a new facial emotion recognition method based on transient light flux and geometric features. In this work, the three-dimensional features of each image are extracted using photogrammetric methods. The transient features are then obtained using optical flow-based motion between successive images. Classification is performed using a support vector machine and artificial neural networks. In 2014 [10], Loconsole et al. first determined 54 facial features using

*Corresponding author. Email: mahdi61380@gmail.com

the Active Shape Model (ASM). Then, the positions of the landmarks are used to extract two types of features, namely linear and centrifugal features. To create the eccentricity features, the eccentricity of the ellipses is calculated. The linear distance between feature pairs is calculated to determine linear features. In 2014 [11], Saeed et al. determined eight face points in the first image and detected the points manually. To address the shortcomings of point localization, they used a training point distribution model (PDM) to project viewpoints onto viewpoint subspaces. Experiments were conducted using the $CK+$ database. Similarly, Guimir et al. 2015 [12] used Elastic Bunch Graph Matching (EBGM) to rank face features in the first frame and then track them for other frames in time. AdaBoost is used with an infinite machine learning classifier that includes multiple classes to select the differential geometric features. The results are 95.5% and 97.80% after tests with the Multimedia Understanding Group (MUD) and Cohn-Kanade ($CK+$) databases, respectively. In 2016, De Souza et al. presented a simple but effective method for mode detection. This method combines a convolutional neural network and an extraordinary preprocessing step. This research found that convolutional neural networks are primarily used for big data because facial expression recognition has a limited data range. To solve this problem, therefore, an extreme preprocessing stage is used to prepare images in large quantities and numbers in this method. The data from three databases, $CK+$, JAFFE, and Bu_3DFE, are used to evaluate the proposed method, and the results of this research have a recognition rate of 96.76%. In 2016 [13], Ghimire et al. presented a new facial expression recognition method combining local region-specific features and a Support Vector Machine. These local region-specific features are often geometric and produce different facial expressions than others. One of the distinguishing innovations of this research is that in most face recognition research [14], a person's face is divided into local regions. However, the entire face region was used for identification in this research.

The geometric features are also calculated from specific points, such as the corners of the lips, the areas around the eyes, and the eyebrows, where the facial expressions are visible with these points [15]. In this study, after extracting the geometric features of the face using a Support Vector Machine classification, the feature vectors were classified into six classes, and the results have shown the success of this study in recognizing some facial expressions. The $Coohan-canade$ database was used to evaluate this research. However, the main weakness of this method is the low recognition rate in some situations, such as anger and fear, which needed better recognition rates. Alizadeh et al. (2017). [16] used a deep learning method for facial expression recognition. In this method, the values of

the image pixels and the values obtained from applying the directed gradient histogram method are used for the inputs of the wrinkled neural networks.

The neural network has been implemented. This study contains two convolutional layers. To prevent overfitting of this network, methods such as stack removal and normalization were used for the settings of the second layer [17, 18]. After the convolution phase and neural network, all images are classified. This research used the cross-validation method for cross-validation and determining the optimal parameters. However, the results of this research could have been better for iterations less than five and were less than 90% in most cases. However, good results were obtained for iterations more than 5 in recognizing facial expressions, but the time and computational complexity increased significantly. Qayyum et al. presented a new method for facial expression recognition in 2017 [19]. This research introduced feature extraction as one of facial expression recognition's most critical and significant phases. This descriptor significantly improves the recognition rate of facial expressions by extracting distinctive features in both the spectral and spatial domains. This study claims that a combination of horizontal and vertical wavelet subbands contains complete information about facial muscle movements. To reduce the dimensions of the feature vectors, a discrete cosine transform is applied to each of the mentioned subbands. Furthermore, in the last step, the selected features are transmitted to a neural network (feed-forward) trained with the backpropagation algorithm [20]. In this research, two databases, jaffa and CK+dataset, were used. The recognition rate of facial expressions in these two studies was 98.83% and 96.61%, respectively. Sariyanidi et al. proposed a method based on a data-driven feature extraction structure in 2017 [21]. It represents the structure of facial changes as a linear combination of local basis functions whose coefficients are proportional to the brightness intensity of the pixels. This research showed that a linear model with Gabor fuzzy shift could calculate the linear functions based on the framework [22, 23]. The proposed method uses video images as training and test data. The proposed method obtained more than 95% results, which cannot be considered very strong compared to the results of other methods. Mvnir et al. researched facial expression recognition comprehensively in 2018 [24]. This research combined the challenges of facial expressions with challenges such as changing the face's light and changing the face's angles, which is considered very comprehensive and complete research.

The proposed method is used for the first time to combine Fast Fovrier and Contrast Limited Adaptir HE to solve the luminance imbalance. Then the fused binary code (MBPC) is generated for the pixel. Two

bits per neighborhood are used to generate a 16-bit code per pixel. This code fuses the local features to improve the effectiveness of the face recognition system. In conclusion, the MBPC method makes changes along the edges and prominent patterns in the eyes, eyebrows, mouth, and edges around the face. The proposed method was performed on the wild Static Facial Expression (SFEW) database, and the proposed method was performed with 96.5% accuracy for facial expression recognition. However, the proposed method has one strength and two weaknesses [25]. First, the proposed method has several. It has studied the challenge, which has an acceptable scope. However, first, this method is very computationally expensive, and second, the recognition rate of facial expressions could have been better.

According to what has been mentioned in the literature and the background of the research, this research area has yet to reach its ideals. In recent years, in order to reach the desired ideals, more Deep Learning methods have been used than other methods. More has been used [26], but these methods are associated with two problems: computational complexity and recognition rate. Therefore, the need for a facial expression recognition system with fewer errors and less complexity is becoming more apparent. In this paper, the paper has tried to present a method based on the answer to the mentioned challenge. One of the methods for extracting features from images, which is considered a type of Deep Learning but does not have the time and computational complexity of neural networks, is the Scattering Wavelet [27].

What distinguishes this algorithm from other feature extraction methods is that this algorithm has a very high capability and accuracy in distinguishing similar images belonging to different classes, even if the image is subject to different types of damage. This descriptor generates feature vectors for severe problems such as noise or optical changes so that the classification error is minimal. This paper aims to investigate the performance of the proposed method in comparison with some other machine-learning methods using two databases of facial expressions. In the rest of this paper, it is implemented in this way. Section 2 describes the scattering wavelet, Section 3 interprets the method proposed for this article, and Section 4 discusses the results of this investigation. In the last part of this paper, a broad conclusion from this article has been drawn.

2. Scattering Wavelet

The coefficients, Fourier transform, and the subset of this transform are reasonably stable against transmission. However, this transform is very unstable against deformation, especially at high signals, and any deformation of the signal causes large changes at significant signals[28]. For this reason, they have been used more often because of their stability to deformation. A wavelet is a local wave as opposed to a sine wave. Furthermore, that makes it stable for changes. However, since the wavelet transform computes convolutions with wavelet filters, the wavelet transform is unstable to changes [29, 30]. To this end, a set of wavelet filters is needed to produce a descriptor with stable features against deformation, transmission, scaling, direction, and dilation [31–33]. Two-dimensional directional wavelets are obtained by scaling and rotating a single-pass filter ψ. Assume that G is a group of limited and discrete rotational angles in R^2. Consequently, the set of multi-scale filters for each scale $j \in Z$ and rotation $r \in G$ is obtained with the following relationship:

$$\psi_{2^j r}(u) = 2^{2j} \psi(2^j r^{-1} u)$$

that if the Fourier transform $\psi(w)$ is the frequency center η, therefore $\psi_{2^j r}(w) = 2^{2j} \psi(2^j r^{-1} w)$ has a support at the center $2^j r \eta$ with a bandwidth proportional to 2^j.

A wavelet transform x filters the signal using a family of wavelets $x * \psi_\lambda(u)$ This filter is computed as a filter bank that contains rotating and dilated wavelets that have no unusual properties. The use of the filter bank is stable to changes such as rotation and dilation as long as the scale and rotation wavelet filters cover the entire frequency range. that this research has used Morlet wavelet filter banks and the relationship $\psi(u) = C_1(e^{iu} - C_2)e^{-|u^2|/(2\sigma^2)}$. has been used to obtain these filters. In conclusion, the use of this filter bank provides stability against rotation and dilation. But the important point is that a wavelet transform is obtained by transmission and therefore is not stable against transmission. In order to generate a feature vector that is stable to change, this vector must be generated in a nonlinear structure (under the operator).

For example, suppose R is a linear or nonlinear operator computed by transfer and $R(L_c x) = L_c R x$. Therefore, the integral $\int Rx(u)du$ is stable against the transfer, but if this operator is obtained by convolution of the wavelet filter with the signal, we should say that it is reasonably stable against the transfer because $\int \psi_\lambda(u)du = 0$ and this condition leads to $\int x * \psi_\lambda(u)du = 0$.

If consider $Rx = M(x * \psi_\lambda)$ where M is a linear operator, the integral can be a small value. But if the operator is a non-linear operator (pointwise). Therefore, the resulting integral becomes zero and is before the transition [34]. One of the most useful nonlinear operators suitable for this problem is the modulus operator. So, in summary, the most stable coefficients against transfers are computed by repeating modulus operators and wavelet transforms.

Assume that $U[\lambda]x = |x * \psi_\lambda|$ and each sequence $p = (\lambda_1, \lambda_2, ..., \lambda_m)$ generates a path that produces a nonlinear product: $U[p]x = U[\lambda_1]U[\lambda_2]...U[\lambda_m]x = |\langle x, \psi_{\lambda_1}\rangle| \cdot |\langle x, \psi_{\lambda_2}\rangle|...|\langle x, \psi_{\lambda_m}\rangle|$.

This research used Morlet wavelet filter banks, and the relationship was used to obtain these filters. Finally, using this bank filter provides stability against rotation and stretching. However, the critical point is that a wavelet transform is obtained by transmission and therefore is unstable against transmission. This vector must be generated in a nonlinear structure (under the operator) to generate a feature vector that is stable to change. For example, suppose R is a linear or nonlinear operator computed by transmission. Therefore, the integral is stable against transmission. However, this operator is obtained by convolution of the wavelet filter with the signal. In that case, it should be said that it is stable against transmission because it is, and it leads to this condition. If one considers it a linear operator, the integral can be a small value. However, the operator is nonlinear (point-wise).

Therefore, the resulting integral becomes zero and is before the transition [35]. The moduli operator is one of the most useful nonlinear operators suitable for this problem. In summary, repeating modulus operators and wavelet transforms compute the most stable coefficients against transfers. Assume that each sequence generates a path that produces a nonlinear product: a path scattering transform is defined as an integral normalized by the Dirac response. Each scattering coefficient is stable against the signal transmission x. Nevertheless, what is very clear is the remarkable similarity of this transform to the Fourier transform modulus. That this transform is also stable against transmission. However, the scattering is a Lipschitz continuity and has high stability and resistance to deformation, but this condition does not exist in the Fourier transform. For the classification, performing the calculation locally for descriptors is often better[36, 37].

This is stable for transitions smaller than a predefined scale 2^j, while spatial variability is preserved for scales more significant. This condition is achieved by localizing the scattering integral by scaling the spatial window $\phi_{2J}\phi(u - v)dv$. The following relation defines a neighborhood scattering transformation.

$$S_J[p]x(u) = U[p]x * \phi_{2J}(u) = \int U[p]\phi_{2J}(u - v)du$$

Now, if p is a set with length M, $S_j[p]x(u)$ is called scattering coefficients in depth m and scale 2^j. These coefficients are calculated in the specified layer m of a channel network. For significant coefficients, several layers of a convolutional mesh are required so that even the most minor information is recovered. Figure 1 is
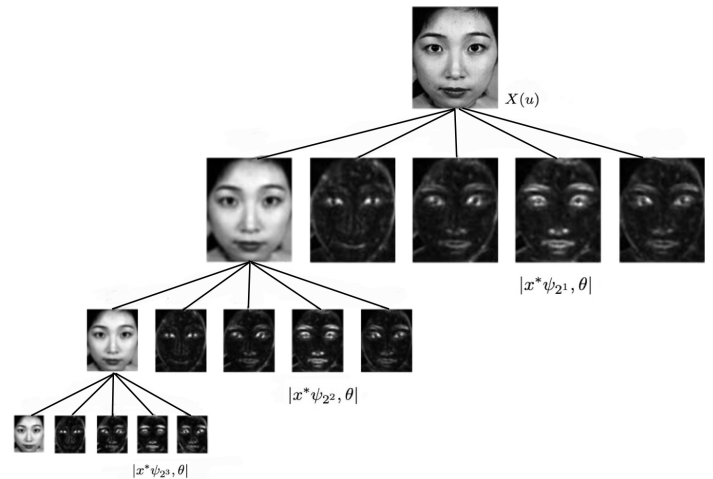

Figure 1. An example of a scattering wavelet

an example of a scattering wavelet. This figure visually represents the structure or characteristics of a scattering wavelet. It illustrates the transformation process, how the input signal is processed through different layers, and how the coefficients are obtained.

3. Proposed Method

The facial expression recognition system proposed in this research is based on the scattering wavelet descriptor.

Like many other facial expression recognition methods, this method consists of three main parts: feature extraction, feature selection, and state detection.

In feature extraction, the scattering wavelet algorithm was used in this study, principal component analysis was used in feature selection, and the K-nearest-neighbor classifier was used in identification. Figure 2 is the block diagram of the proposed method. This block diagram visually represents the flow of information or processes in a system. In the context of the study, the proposed method involves three key components: scattering wavelet algorithm for feature extraction, principal component analysis (PCA) for feature selection, and the K-nearest-neighbor (KNN) classifier for identification.

It is evident that our proposed method consistently achieves higher recognition accuracy compared to the ZM method across various datasets. Notably, during certain epochs, our method exhibits superior performance, indicating its robustness and effectiveness in capturing facial expressions accurately. The consistency observed in both training and testing phases underscores the reliability and generalizability of our proposed approach. This comparison provides a comprehensive understanding of the improved recognition
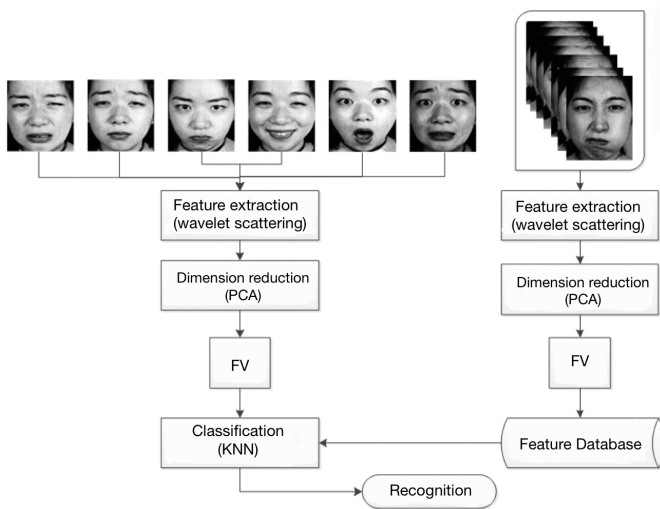
Figure 2. An example of a scattering wavelet

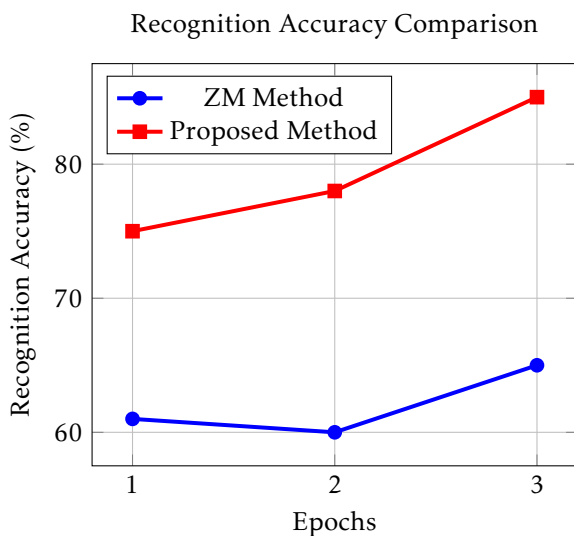accuracy offered by our method over the existing ZM method, can be seen in Figure 3.

Figure 3. Recognition Accuracy Comparison between Proposed Method and ZM Method

3.1. Feature Extraction

The feature extraction part is one of the most difficult parts of the research related to identity recognition, so in the first part, the feature extraction is performed using a set of Gabor wavelet filters, and depending on the size of the wavelet cascade images, the required product is generated. In the second step, the scattering of the iris image is calculated using the generated wavelet cascade (scattering transformation). This transformation calculates the features of each

depth of the image separately. In the last step, the feature vector of each image is generated by concatenating the vectors obtained from each depth.

These feature vectors are differentiated enough to be placed in the correct classes by the distance criterion of most classifiers. However, one of the main problems in deep learning research is the generation of feature vectors with immense dimensions, and this research is subject to this problem. The feature vectors extracted from face images with the mentioned dimensions have more than 100,000 features. The critical point is that, firstly, the classification of these dimensions of feature vectors is complicated and time-consuming for each classification. By calculating the variance of the feature vectors, this algorithm not only selects the most considerable variances and the first components that contain the most accurate information about the feature vector. It also reduces the dependency between the data, and this reduction of the dependency between data leads to a more robust differentiation of the feature vectors. The generated feature vector contains new values that are quantitatively much smaller than the original feature vector and qualitatively more different from the feature vector extracted by the descriptor.

4. Tests

In this dedicated section, the efficacy of the proposed method is subjected to a comprehensive evaluation, employing the specifically introduced database within the confines of this study. The methodology outlined in this study is inherently global in its approach, prompting a judicious comparison with Zernike moments—a well-regarded and established method renowned for its effectiveness in facial expression identification.

Table 1 serves as an insightful repository, encapsulating the intrinsic details of both the proposed and comparison methods. Not only does it delineate the parameters unique to the proposed method, but it also furnishes a comprehensive account of the parameters associated with Zernike moments, enriching the comparative analysis.

It is imperative to note that the stipulated parameters, meticulously presented in Table 1, remain constant throughout the entirety of the experimental phases. This deliberate decision ensures a consistent and fair comparison across the different stages of the evaluation. By maintaining parameter constancy, the robustness and repeatability of the experimental results are fortified, allowing for a nuanced assessment of the proposed method's performance against the benchmark set by Zernike moments. This strategic approach underscores the methodological rigor employed in this evaluation, thereby enhancing the reliability and validity of the ensuing findings.

Method	Specification
ZM	The selected order of Zernike moments is 10. The values of r and σ are empirically set to 10 and 2 due to the best performance of systems using these parameters.
Proposed Method	•The orientation parameter in the scattering wavelet is set to 2. • The scale parameter of this algorithm is set to 2. • The average distances between the samples of a training class and a test class were chosen as the criterion for calculating the distance between the two classes. • The parameters for the average Gaussian noise are zero, and the standard deviation of the Gaussian noise is 0.01. • The parameters for the density of the salted pepper noise were set to 0.01.

Table 1. Method Specifications

Figure 4. An example of images from the JAFFE database in 6 poses

Figure 5. An example of images from the CK+ database in 6 poses

In the remainder of this section, the database of images used is first described, and then the results of the tests and the comparison between the proposed methods are explained.

4.1. Databases

Images from both JAFFE [38] and $CK+$ [39, 40] databases were used to evaluate the proposed method. This paper used 180 images from the JAFFE database, and 981 images from $CK+$ database to train and evaluate the proposed method. These images have been used in the training and testing phases of the proposed method, with a two-thirds ratio for training and one-third for testing. These two databases are classified into six emotional categories: anger, disgust, fear, happiness, sadness, and surprise. Figure 4 and 5 illustrates a sample of images from the two databases under consideration in the six specified positions, anger, disgust, fear, happiness, sadness and suprise. The first row is from $CK+$, and the next row is from JAFFE.

4.2. Hardware and software platforms for performing tests

The simulation of the proposed method was meticulously conducted within the MATLAB R2017a programming environment, operating seamlessly under the Windows 10 operating system. The hardware platform selected for the simulation of this research embodies a robust Intel® Core™ i5-8500 system, equipped with 8 GB RAM as the primary memory, further fortified with an additional 16 GB reserve RAM. The storage architecture is facilitated by an SSD hard disc drive, ensuring swift data access and seamless execution of computational tasks.

4.3. Lab Results

In this section, it is very important to go into the intricacies of implementing both the proposed methods and the methods under evaluation. The process involved a careful selection of training and testing images, and several critical points should be considered in explaining these details. The experimental setup entailed a random selection of images for both training and testing phases across all experiments. Specifically, two-thirds of the images sourced from each database were allocated for training purposes, while the remaining one-third constituted the testing set [41]. This balanced partitioning strategy ensures a comprehensive evaluation of the models' performance on distinct datasets.

To enhance the robustness and reliability of the results, the experiments were repeated ten times, each time with randomly sampled test and training sets. This iterative approach not only accounts for the inherent variability in the dataset but also facilitates a more comprehensive understanding of the methods' stability and consistency.

The identification accuracy recorded in the results table is the result of the average of 10 runs of these tests, taking into account the values of identification accuracy from 10 replicates of each test and calculating the confidence level of 95% of the confidence interval

range of all tests. Table 2 shows the results of the the ZM method and the results of the proposed method, both with the k-nearest neighbor classification and the CK+ database. Comparing the results of the proposed method and the ZM method with the knn classifier and the CK+ database show that the proposed method has a much higher detection rate than the comparison method, which has a detection rate of 100% in most cases.

However, the results of Table 3 are very far from the predictions made for this combination, and the combination of the proposed descriptor with the support vector machine classifier could not obtain a favorable detection rate in the face of image damage [42, 43]. These results show that although facial expression recognition in the unchallenged mode has a very high recognition rate, in the condition that the image pixel values are destroyed under conditions such as noise, the recognition rate decreases significantly, so the use of a support vector machine classifier can not obtain good results. The proposed method was implemented on the JAFFAE database for further evaluation, and the results were presented in Table 4.

The results of these tables highlight the effectiveness of the proposed method in identifying facial expressions with different challenges compared to the Zernike moment-based method. However, this superiority should be discussed from two perspectives. First, on what factors does the stability of the proposed method against image damage depend, and second, where does the superiority of the proposed method come from compared to the Zernike moments-based method? To answer these two questions, it must first explain the diffusion wavelet. A series of Morlet high-pass filters (Gabor-Violet) and a low-pass filter extract diffusion wavelet features from the image.

These filters are applied at different scales and to the image and outputs of each filter, as shown in Figure 1, with a structure similar to the convolutional neural network. There is a penalty. These vectors have such a distinguishing feature that their classification error rate is minimal even in severe damage, such as noise, saving the research target from any damage [44]. Utilizing scattering wavelet applications in robotics enhances recognition accuracy for facial expressions, bolstering human-robot interaction. This technology harnesses advanced signal processing to discern subtle nuances in facial features, improving the robot's ability to interpret and respond to human emotions effectively, fostering more seamless and empathetic communication [45].

5. Conclusion

Methods grounded in Deep Learning have undeniably showcased their effectiveness in the realm of facial expression identification. However, despite their prowess, these approaches grapple with challenges, notably in terms of computational complexity and achieving consistently high recognition rates. In light of these concerns, this paper introduces a novel approach for facial expression recognition founded on wavelet scattering. The proposed method is meticulously designed and rigorously evaluated through simulations, focusing on the recognition of diverse facial expressions such as anger, disgust, fear, happiness, surprise, and sadness. Notably, under normal conditions, the method achieves an outstanding identification rate of 100%. Furthermore, the robustness of the proposed method is tested in the presence of challenges, including Gaussian noise and salt-pepper noise. In the presence of Gaussian noise, the identification rate remains commendably high at 91%, showcasing the resilience and reliability of the method in real-world scenarios. Even in the more challenging salt-pepper noise mode, the method maintains a substantial identification rate of 85%, underscoring its superiority and practical applicability in less-than-ideal conditions. These results affirm the efficacy of the proposed wavelet scattering-based method, offering a promising avenue for enhanced facial expression recognition, particularly in the presence of environmental noise.

References

[1] Samta Jain Goyal, Arvind K Upadhyay, RS Jadon, and Rajeev Goyal. Real-life facial expression recognition systems: a review. *Smart computing and informatics*, 77(1):311–331, 2018.

[2] Zhihong Zeng, Maja Pantic, Glenn I Roisman, and Thomas S Huang. A survey of affect recognition methods: audio, visual and spontaneous expressions. In *Proceedings of the 9th international conference on Multimodal interfaces*, pages 126–133, 2007.

[3] Keith Anderson and Peter W McOwan. A real-time automated system for the recognition of human facial expressions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 36(1):96–105, 2006.

[4] Beat Fasel and Juergen Luettin. Automatic facial expression analysis: a survey. *Pattern recognition*, 36(1):259–275, 2003.

[5] Ira Cohen, Nicu Sebe, Ashutosh Garg, Lawrence S Chen, and Thomas S Huang. Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and image understanding*, 91(1-2):160–187, 2003.

[6] Paul Ekman. Facial expression and emotion. *Facta universitatis-series: Philosophy and Sociology*, 1(3):195–212, 1996.

[7] Paul Ekman and Wallace V Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124, 1971.

[8] Xiaobai Li, Xiaopeng Hong, Antti Moilanen, Xiaohua Huang, Tomas Pfister, Guoying Zhao, and Matti Pietikäinen. Reading hidden emotions: spontaneous

Table 2. ZM method vs proposed method (New) results with the Knn classifier and the CK+ database

Noise	Anger		Disgust		Fear		Happiness		Sadness		Suprise		Overall	
	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New
Normal	95.5%	100%	94.5%	100%	94.7%	98.2%	94.6%	100%	95.2%	100%	95.2%	100%	94.9%±0.4	99.7%±0.6
Gaussian	93.1%	100%	92.4%	100%	91.5%	98.3%	93.4%	100%	92.1%	100%	92.4%	100%	92.7%±0.6	99.7%±0.6
Salt-Pepper	88.2%	98.5%	90.9%	95.2%	91.4%	100%	90.5%	100%	90.8%	100%	89.6%	97%	90.2%±0.7	98.2%±0.8

Table 3. ZM method vs proposed method (New) Results with SVM Classifier on CK+ Database

Noise	Anger		Disgust		Fear		Happiness		Sadness		Suprise		Overall	
	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New
Normal	98.1%	100%	98.7%	100%	99.1%	100%	98%	100%	97.2%	100%	97.5%	100%	98.9%±0.3	100%
Gaussian	21.6%	20.1%	18.3%	19.5%	25.1%	22.4%	31/1%	28.1%	25.3%	17.5%	38.3%	40.1%	28%±1.6	27%±0.9
Salt-Pepper	21.8%	14.8%	19.2%	11%	24%	17.8%	29.2%	22.6%	24%	13.7%	35.7%	35.6%	21.8%±1.7	22.35%±1.2

Table 4. ZM method vs. proposed Method(New) Results with KNN Classifier on JAFFAE Database

Noise	Anger		Disgust		Fear		Happiness		Sadness		Suprise		Overall	
	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New	ZM	New
Normal	100%	100%	90%	90%	80%	100%	90%	100%	90%	100%	100%	100%	91.66%	98.33%
Gaussian	60%	90%	60%	80%	50%	90%	40%	100%	50%	90%	60%	100%	60%	91.66%
Salt-Pepper	50%	80%	60%	80%	60%	80%	50%	90%	50%	90%	60%	90%	65%	85%

micro-expression spotting and recognition. *arXiv preprint arXiv:1511.00423*, 2(6):7, 2015.

[9] Yandan Wang, John See, Yee-Hui Oh, Raphael C-W Phan, Yogachandran Rahulamathavan, Huo-Chong Ling, Su-Wei Tan, and Xujie Li. Effective recognition of facial micro-expressions with video motion magnification. *Multimedia Tools and Applications*, 76:21665–21690, 2017.

[10] R Niese, A Al-Hamadi, A Farag, H Neumann, and B Michaelis. Facial expression recognition based on geometric and optical flow features in colour image sequences. *IET computer vision*, 6(2):79–89, 2012.

[11] Claudio Loconsole, Catarina Runa Miranda, Gustavo Augusto, Antonio Frisoli, and Verónica Orvalho. Real-time emotion recognition novel method for geometrical facial features extraction. In *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, volume 1, pages 378–385. IEEE, 2014.

[12] Anwar Saeed, Ayoub Al-Hamadi, Robert Niese, and Moftah Elzobi. Frame-based facial expression recognition using geometrical features. *Advances in human-computer interaction*, 2014:4–4, 2014.

[13] Deepak Ghimire, Joonwhoan Lee, Ze-Nian Li, Sunghwan Jeong, Sang Hyun Park, and Hyo Sub Choi. Recognition of facial expressions based on tracking and selection of discriminative geometric features. *International journal of multimedia and Ubiquitous Engineering*, 10(3):35–44, 2015.

[14] Amjad Rehman Khan, Majid Harouni, Sepideh Sharifi, Saeed Ali Bahaj, and Tanzila Saba. Face detection in close-up shot video events using video mining. *Journal of Advances in Information Technology*, 14(2), 2023.

[15] Ata Jahangir Moshayedi, Sithembiso Khaya Sambo, and Amin Kolahdooz. Design and development of cost-effective exergames for activity incrementation. In *2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pages 133–137. IEEE, 2022.

[16] Deepak Ghimire, Sunghwan Jeong, Joonwhoan Lee, and San Hyun Park. Facial expression recognition based on local region specific features and support vector machines. *Multimedia Tools and Applications*, 76:7803–7821, 2017.

[17] Mehdi Marani, Morteza Soltani, Mina Bahadori, Masoumeh Soleimani, and Atajahangir Moshayedi. The role of biometric in banking: A review. *EAI Endorsed Transactions on AI and Robotics*, 2(1), 2023.

[18] Ge Xu, Amir Sohail Khan, Ata Jahangir Moshayedi, Xiaohong Zhang, and Yang Shuxin. The object detection, perspective and obstacles in robotic: A review. *EAI Endorsed Transactions on AI and Robotics*, 1(1), 2022.

[19] Shima Alizadeh and Azar Fazel. Convolutional neural networks for facial expression recognition. *arXiv preprint arXiv:1704.06756*, 2017.

[20] Mina Bahadori, Masoumeh Soleimani, Morteza Soltani, and Mehdi Davari. Application of iot in drug supply chain management (scm) during covid-19 outbreak. In *Advanced IoT Technologies and Applications in the Industry 4.0 Digital Economy*, pages 224–237. CRC Press.

[21] Huma Qayyum, Muhammad Majid, Syed Muhammad Anwar, Bilal Khan, et al. Facial expression recognition using stationary wavelet transform features. *Mathematical Problems in Engineering*, 2017, 2017.

[22] Ali Khansefid, Seyed Mahmoudreza Yadollahi, Gerhard Müller, and Francesca Taddei. Induced earthquake hazard by geothermal power plants: Statistical evaluation and probabilistic modeling. *International Journal of Disaster Risk Science*, 13(5):758–777, 2022.

[23] Ata Jahangir Moshayedi and DC Gharpure. Development of position monitoring system for studying performance of wind tracking algorithms. In *ROBOTIK 2012; 7th German Conference on Robotics*, pages 1–4. VDE, 2012.

[24] Evangelos Sariyanidi, Hatice Gunes, and Andrea Cavallaro. Learning bases of activity for facial expression recognition. *IEEE Transactions on Image Processing*, 26(4):1965–1978, 2017.

[25] Usman Mehmood, Ephraim Bonah Agyekum, Salah Kamel, Hossein Shahinzadeh, and Ata Jahangir Moshayedi. Exploring the roles of renewable energy, education spending, and co2 emissions towards health spending in south asian countries. *Sustainability*, 14(6):3549, 2022.

[26] Asim Munir, Ayyaz Hussain, Sajid Ali Khan, Muhammad Nadeem, and Sadia Arshid. Illumination invariant facial expression recognition using selected merged binary patterns for real world images. *Optik*, 158:1016–1025, 2018.

[27] A Jahangir Moshayedi, Ge Xu, Liefa Liao, and Amin Kolahdooz. Gentle survey on mir industrial service robots: Review & design. *J. Mod. Process. Manuf. Prod*, 10(1):31–50, 2021.

[28] Nianyin Zeng, Hong Zhang, Baoye Song, Weibo Liu, Yurong Li, and Abdullah M Dobaie. Facial expression recognition via learning deep sparse autoencoders. *Neurocomputing*, 273:643–649, 2018.

[29] Amjad Rehman, Majid Harouni, Maedeh Omidiravesh, Suliman Mohamed Fati, and Saeed Ali Bahaj. Finger vein authentication based on wavelet scattering networks. *Computers, Materials & Continua*, 72(2), 2022.

[30] Amjad Rehman, Majid Harouni, Negar Haghani Solati Karchegani, Tanzila Saba, Saeed Ali Bahaj, and Sudipta Roy. Identity verification using palm print microscopic images based on median robust extended local binary pattern features and k-nearest neighbor classifier. *Microscopy Research and Technique*, 85(4):1224–1237, 2022.

[31] Stéphane Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics*, 65(10):1331–1398, 2012.

[32] Morteza Soltani, Jeffrey P Kharoufeh, and Amin Khademi. Optimal call center staffing and pricing under qos constraints. In *IIE Annual Conference. Proceedings*, pages 1–6. Institute of Industrial and Systems Engineers (IISE), 2023.

[33] Morteza Soltani, Jeffrey P Kharoufeh, and Amin Khademi. Structured replacement policies for offshore wind turbines. *Probability in the Engineering and Informational Sciences*, pages 1–32, 2023.

[34] Joan Bruna and Stéphane Mallat. Invariant scattering convolution networks. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1872–1886, 2013.

[35] Joan Bruna and Stéphane Mallat. Classification with scattering operators. In *CVPR 2011*, pages 1561–1566. IEEE, 2011.

[36] Masoumeh Soleimani, Zahra Forouzanfar, Morteza Soltani, and Majid Jafari Harandi. Imbalanced multiclass medical data classification based on learning automata and neural network. *EAI Endorsed Transactions on AI and Robotics*, 2, 2023.

[37] Ata Jahangir Moshayedi, Saeed Safara Fard, Liefa Liao, and Seyed Ali Eftekhari. Design and development of pipe inspection robot meant for resizable pipe lines. *Int. J. Robot. Control*, 2(1):25, 2019.

[38] Michael Lyons, Shigeru Akamatsu, Miyuki Kamachi, and Jiro Gyoba. Coding facial expressions with gabor wavelets. In *Proceedings Third IEEE international conference on automatic face and gesture recognition*, pages 200–205. IEEE, 1998.

[39] Patrick Lucey, Jeffrey F Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *2010 ieee computer society conference on computer vision and pattern recognition-workshops*, pages 94–101. IEEE, 2010.

[40] Takeo Kanade, Jeffrey F Cohn, and Yingli Tian. Comprehensive database for facial expression analysis. In *Proceedings fourth IEEE international conference on automatic face and gesture recognition (cat. No. PR00580)*, pages 46–53. IEEE, 2000.

[41] Mina Bahadori, Morteza Soltani, Masoumeh Soleimani, and Mahsa Bahadori. Statistical modeling in healthcare: Shaping the future of medical research and healthcare delivery. In *AI and IoT-based technologies for precision medicine*, pages 431–446. IGI Global, 2023.

[42] Dilnoza Mamieva, Akmalbek Bobomirzaevich Abdusalomov, Mukhriddin Mukhiddinov, and Taeg Keun Whangbo. Improved face detection method via learning small faces on hard images based on a deep learning approach. *Sensors*, 23(1):502, 2023.

[43] Shivalila Hangaragi, Tripty Singh, and N Neelima. Face detection and recognition using face mesh and deep neural network. *Procedia Computer Science*, 218:741–749, 2023.

[44] Ata Jahangir Moshayedi and DC Gharpure. Evaluation of bio inspired mokhtar: Odor localization system. In *2017 18th international carpathian control conference (ICCC)*, pages 527–532. IEEE, 2017.

[45] Omid Toutian Esfahani and Ata Jahangir Moshayedi. Accuracy of the positioning systems for the tracking of alzheimer's patients-a review. *International Journal of Applied Electronics in Physics & Robotics*, 2(2):10–16, 2014.