International Journal of Mathematics and
Computer Science, **17**(2022), no. 3, 1405–1417

Emotions Students' Faces Recognition using Hybrid Deep Learning and Discrete Chebyshev Wavelet Transformations

Shaymaa Ayad Mohammed, Asma Abdulelah Abdulrahman, Fouad Shaker Tahir

Branch of Mathematics and Computer Applications
Department of Applied Sciences
University of Technology Iraq
Baghdad, Iraq

email: as.20.24@grad.uotechnology.edu.iq,
Asma.A.Abdulrahman@uotechnology.edu.iq, 11015@ uotechnology.edu.iq

(Received March 1, 2022, Accepted April 27, 2022)

Abstract

This research reflects students' emotions while learning from instructors to predict students' conditions or degree of contentment. The methodology utilized in this research involves six students, depending on the number of emotional sensations being studied. Then deep learning is used to analyze face recognition. The convolutional neural network is trained using Alex Net and Google Net. The emotions on the pupils' faces were: natural, happy, angry, afraid, sad, and surprised. This collection of emotions could be related to the students' familiarity with the lecture. The convolutional neural network was improved in MATLAB. The quality of the images is enhanced using a new filter derived from Discrete Chebyshev Wavelet Transformations (DCHWT) convoluted with neural network called Chebyshev Wavelet convolutional neural network (CHWCNN). ChWCNN improved the quality of the results by accomplishing better accuracy. In this report, six students were tested during lecture time and the images were analyzed. The results are obtained through two successive approaches.

Key words and phrases: Discrete Chebyshev Wavelet Transformations (DCHWT), Convolutional Neural Network (CNN), Chebyshev Wavelet convolutional neural network (CHWCNN).
AMS Subject Classifications: 68R01.
ISSN 1814-0432, 2022, http://ijmcs.future-in-tech.net

The first approach is to test the Alex and Google Net before and after employing DCHWT. The results have shown that using DCHWT improved the image from 90% and 72% to 95% and 91% for Alex and Google Net, respectively. The results suggested that Google Net under DCHWT improved the image by 28.4% compared to 5.8% for Alex Net. The results showed that the feature of the happy face acquired the highest percentage of about 50%, while the natural, angry, and sad faces were equally evaluated at 17% each. The features of afraid and surprise were not detected at all. The study could contribute to the learning process by showing those who were satisfied versus those who have different emotions. The study could be very valuable in assessing the quality of the lectures or presenters by showing the face feeling recognition.

1 Introduction

Emotion is one of the most essential characteristics of humans that allows them to express their feelings [1]. Computers and robots cannot process face recognition to a certain level of acceptance unless deep learning is used to mimic human recognition ability [2]. As a result, it is expected that computers and robots will process emotion and naturally interact with human users. Recently, Human-Computer Interaction (HCI) research efforts have been focused on ways to enable computers to understand human emotions [3]. Although few when compared to the efforts being made could mean that some researchers are attempting to realise man-machine interfaces due their capabilities of facial expression recognition analysis [4]. Physiological signal analysis is another possible approach for emotion recognition because emotion on facial expression or speech can be suppressed relatively quickly, and emotional status is inherently reflected in nervous system activity [5]. Traditional tools for studying human emotional states are based on the recording and statistical analysis of physiological signals from central and autonomic nervous systems. Several approaches to determining the relationship between emotional changes and facial signals have been reported by various researchers [6]. One of the challenges of categorizing emotion is that the difficulty of distinction between emotion categories. Everyone expresses their emotions differently: judging or modelling emotions is a difficult task. Researchers frequently employ two distinct methods [7]. For example, one approach could be happy, sad, surprised, angry, fearful, disgusted, etc. Another approach is to categorise emotions using multiple dimensions or scales. Instead of selecting discrete labels, observers can express their reactions to each stimulus on various continuous scales. For instance, pleasant-unpleasant, attraction-rejection, simple-complicated, and others cannot be differentiated well. Because of the ease of protocol design and signal processing techniques, many researchers have attempted to use two-dimensional modelling of emotions [8]. On the other hand, this work aims to realise the basic emotions in the discrete mode to develop an intelligent emotion recognition system. Furthermore, assessing discrete mode emotions helps create appropriate responses in man-machine systems such as robots. Developing a human emotion recognition system is challenging in protocol design and efficient machine learning algorithms for deriving discrete emotions from the multiple complex emotions encountered by humans [9]. Meanwhile, with the help of deep learning, computers have been providing solutions to significant challenges in facilitating the detection of facial emotions, which includes many fields of education, engineering, and sciences [10], [11], and [12]. Deep learning with wavelets has been used in a variety of applications. The Alex net Architecture neural network has been trained to detect faces using discrete wavelet transforms (DWT) [13], and [14]. Furthermore, the MATLAB program's standard orthogonal and discrete wavelet were used [15], [16], and [17]. In this study, students' facial expressions were examined using a new and efficient proposed technique that combined Discrete Chebyshev Wavelet Transformations (DCHWT) with mathematical methods and a convolutional neural network (ChWCNN). This method has produced more precise and accurate results. The newly proposed technique effectively identified the face and read its emotions based on these findings.

2 Facial Emotion Recognition (FER)

The main task of FER is to map various facial expressions to their corresponding emotional states. The traditional FER is divided into two steps: feature extraction and emotion recognition followed by image preprocessing, including face detection, cropping, resizing, and normalisation. Face detection eliminates background and non-face areas before cropping the face area. The most important task in a traditional FER system is feature extraction from the processed image. Existing methods employ well-known techniques such as discrete wavelet transform (DWT), linear discriminant analysis, and ChWCNN [18]. Finally, the extracted features are used to classify emotions, which is typically done using convolutional neural networks (CNN) and other

machine learning techniques.
The convolutional neural networks (CNNs) attract interest in FER due to their built-in feature extraction mechanism from images [19]. A few works using CNN to solve FER problems have been reported [20], and [21]. However, existing FER methods only considered the CNN with a few layers, even though its deeper model has been shown to perform better at other image-processing tasks [22]. The facts underlying this approach could be the difficulties associated with FER. To begin, emotion recognition necessitates a moderately high-resolution image, which necessitates the computation of high-dimensional data. Furthermore, due to the vanishing gradient problem, increasing the number of layers does not increase accuracy after a certain level [23]. Improving accuracy requires various modifications, resulting in introducing techniques to the deep CNN architecture [24]. Training such a deep model necessitates a large amount of data and a high level of computational power.

2.1 Machine Learning-Based FER Approaches

Xiao-Xu and Wei [25] and Zhao et al. [26] enhanced the facial image's wavelet energy feature (WEF) using K-nearest neighbour neural network (KNN). Feng et al. [27] extracted local binary pattern (LBP) histograms from different small regions of the image, combined them into a single feature histogram, and then classified emotion using a linear programming (LP) technique. Zhi and Ruan (2007) used 2D discriminant locality preserving projections to generate facial feature vectors. Furthermore, different methods used the support vector machine (SVM) to classify emotion from extracted feature values. Shih et al. [28] investigated various feature representations and DWT with 2D-linear discriminant analysis (LDA). Alshami et al. [29] used SVM to investigate two feature descriptors, facial landmarks descriptor, and centre of gravity descriptor. Liew and Yairi [30] conducted a comparative study on SVM and several other methods for classification on features extracted using various methods.

2.2 Convolutional Neural Network (CNN)

The input-layer structure, multiple convolutional-pooling hidden layers, and an output layer, the CNN structure creates the best suitable model for the image domain [19]. Convolution is a two-function mathematical operation that can result in a third modified function. CNN also has a small-sized ver-

sion that can search through images for functional patterns using convolution. As the pooling is a non-linear down-sampling technique, the corresponding pooling layer merges non-overlapping areas from one layer into a single value in the following layer. The generic architecture of a standard CNN with two convolutional-pooling layers is depicted in Figure 1.
The first pooling operation yields the first subsampled feature maps (SFMs), which takes the process into the second convolutional-pooling layer operations. Flattening the 2nd SFMs values, the fully connected layer (i.e., dense layer) performs the final reasoning where the neurons are connected to all activations in the previous layer. The final layer, also known as the loss layer, specifies how training penalizes actual and predicted output deviations. A CNN architecture of this type is widely used for pattern recognition from small-sized input images, such as handwritten numeral recognition [18]. CNN Overview, Deep CNN Models, and Transfer Learning (TL). This research includes pre-trained DCNN models and the TL technique. Several pretrained DCNN models are investigated to find the best one for FER.

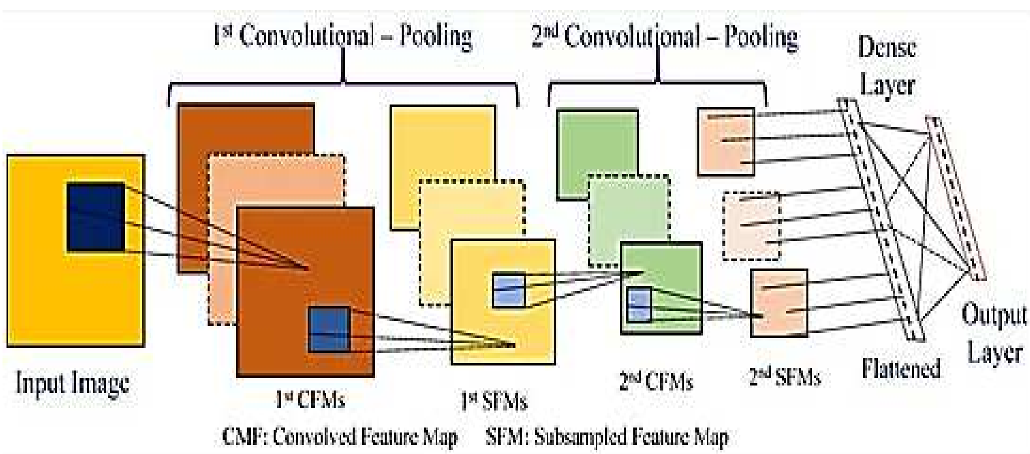

Figure 1: The generic architecture of a convolutional neural network with two convolutional-pooling layers [31].

2.3 Discrete Chebyshev Wavelet with CNN

The details of constructing Discrete Chebyshev Wavelet Transform (DCHWT), analyzing image to approximate coefficients and details coefficients, the image is divided into four parts LL, LH, HL, and HH. The selection of the part LL after compression image and de-noising image of CHWCNN. The results can be reached through the samples and then applying the theory followed by the proposed method.

3 Method of Research

The flowchart shown in Figure 2 expresses the essential steps for utilizing the new DCHWT procedure for recognizing the feelings of the face under the six famous feelings of happy, sad, neutral, afraid, surprised, and angry. DCHWT determines the face recognition using deep learning and training a convolutional neural network. This process requires proposing algorithm that inserts the image on the wavelets before being analyzed. Then, after the compression process, CHWCNN determines the face based on the feelings of the student's face. The stages of the whole process, step-by-step, are shown

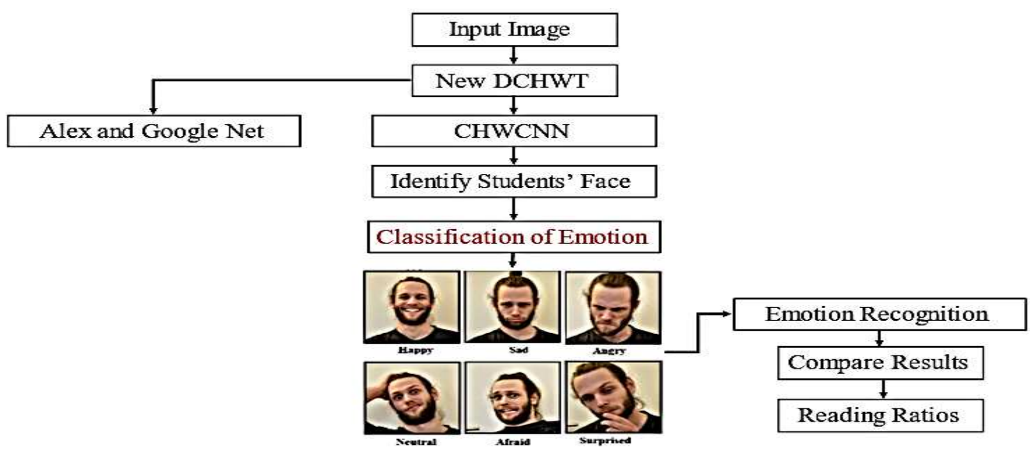

Figure 2: Steps to make the proposed system to determine the feelings of students' faces.

in Figure 3. Having a group of students in a lecture initiates the input image followed by pre-processing stages which includes resizing, enhancing, filtering, and detection. The choice of LL part is then selected followed by the last processes of compression image and the face emotion classification.

4 Result and Discussion

Based on the details shown in Figure 3, the pictures are passing on the normal convolutional neural network after training the network Alex Net and

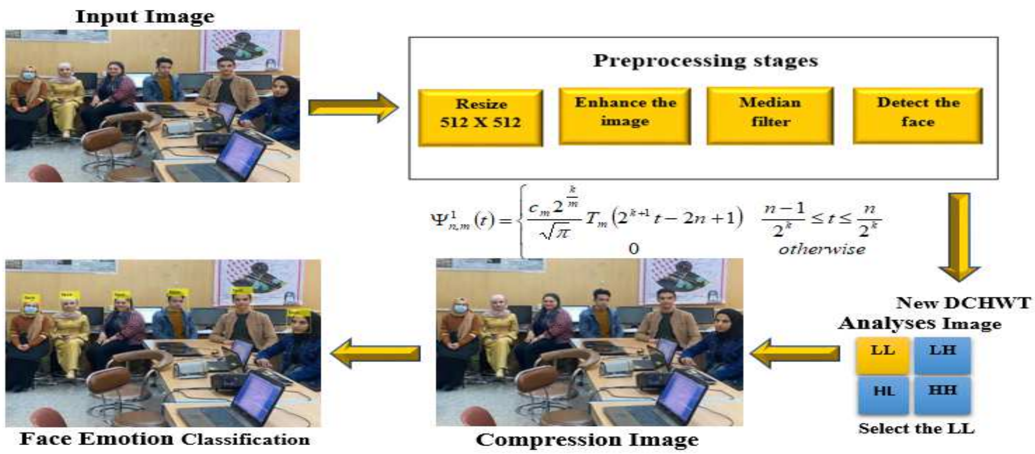

Figure 3: The stages that will take place in this work.

Google Net. The final model for the detection of the face emotion (happy, sad, neutral, surprise, angry, afraid) of the students is then evaluated. Facial emotions were investigated using a new technique called ChWCNN after photographing a group of students during a scientific lecture. ChWCNN performs some operations, including noise reduction and image compression. When the image is passed through the standard convolutional neural network, the final results have an accuracy of 90% for Alex Net and 72% for Google Net, as shown in Table 1. However, after using the wavelet, the accuracy of 95% and 91% for Alex Net and Google Net, respectively. As a result, the improvement due to utilizing the wavelet is approximately 5.8% for Alex Net and 28.4% for Google Net. The findings could persuade people to use Google Net while also improving the accuracy of Alex Net by reviewing the algorithms involved in the technique.

Table 1: Accuracy Results.

Before DCHWT	Before DCHWT	After DCHWT	After DCHWT
Accuracy Alex Net	Accuracy Google Net	Accuracy Alex Net	Accuracy Google Net
90%	72%	95%	91%

The workspace in the MATLAB program (Alex Net and Google Net) showed 251 layers and 221 transform layers with five classes and a pixel range of [-30, 30]. The most essential parameters calculated with Epoch 6 of 6 when training a convolutional neural network were 30 iterations per epoch 5, validation frequency three iterations and learning rate 0.001. Figure

4 depicts the facial recognition mechanism. The upper part of the figure depicts the MATLAB-training process's 6 x 6 Epoch matrix (Epoch 1 through Epoch 6), followed by mathematical smoothing and the variation yield in each Epoch. This outcome means that smoothing is required at this stage of face recognition to remove any noise present in the image. The number of iterations required to complete the convergence process is shown in the lower part of the figure. The lower the Epoch, the fewer iterations required, which increases as the Epoch increases. The benefit of these two processes (Epoch and corresponding iteration) demonstrates the efficacy of the new technique employed in this paper. The facial recognition projected six emotions for a

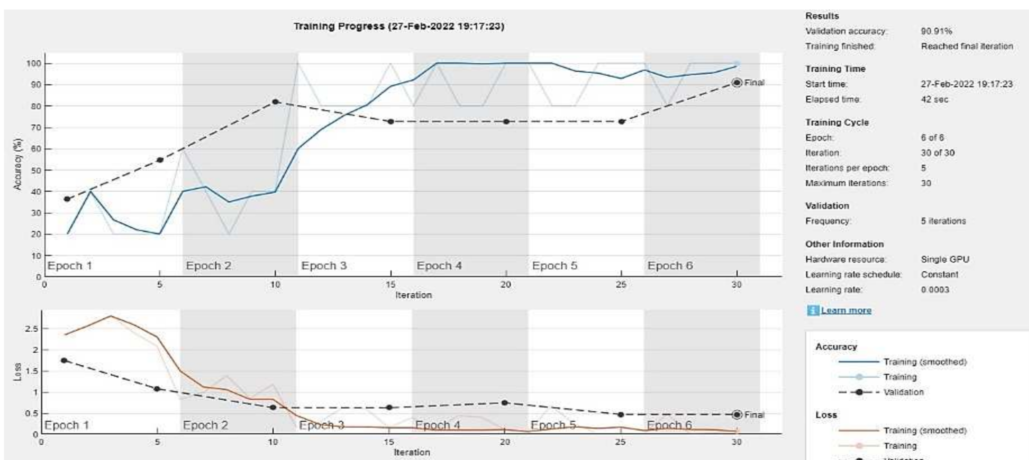

Figure 4: The effectiveness of the new technique and how to identify facial emotions.

group of six students participating in a science lecture with the happy, sad, neutral, surprised, angry, and afraid feeling. Figure 5 explains how emotional face detection recognizes the sum of the six features (a) and then evaluate the corresponding ratios (b). The results showed that the six presumed cases were detected with a happy feature at the maximum (about 50%), followed by three equally shared features of natural, angry, and sad sharing the remaining 50%. Furthermore, face recognition revealed that the two features of fear and surprise were not detected. These findings could be used in class or through e-learning to assist educators in changing their course of instruction, employing more advanced methodology, or returning dissatisfied students to the average or true learning path.

Figure 5: (a) Photos taken and (b) processing and percentage evaluation.

5 Conclusion and Future Work

Emotional feelings, in general, and specifically for students, are a form of advanced research work that contribute to modern life's needs and expectations. Students in the classroom who were involved in gaining knowledge from instructors or in any other way are prone to diverse behaviors for various reasons, including psychology, cognitive ability, and how instructors present the material. The sort of emotion reflects, to a certain degree, the student's circumstances or, more simply, the degree of the student's contentment. The approach utilized in this paper can be described by assuming a group of at least six students, depending on the number of emotional feelings being considered. Then, using a deep learning technique, the facial recognition was analyzed. This work used Alex Net and Google Net as training methods, followed by a convolutional neural network. Six emotions were assigned to the students' faces: natural, happy, neutral, surprised, angry, and afraid. This collection of emotions could be linked to students' typical feelings during the lecture. The convolutional neural network was developed and the results were improved using the MATLAB application. A new filter was built from Discrete Chebyshev Wavelet Transformations (DCHWT) that connected the images to the Chebyshev Wavelet convolutional neural network (CHWCNN). By achieving greater precision, ChWCNN enhanced the quality of the outcomes. The research examined a group of six students during lecture time and assessed the photos. Two sequential approaches were used to acquire the results. The first strategy was to conduct pre and post DCHWT tests on the Alex and Google networks. The results indicated that utilizing DCHWT increased the image's quality from 90% and 72% to 95% and 91% for Alex and Google Net, respectively. The results indicated that Google Net improved the image by 28.4 percent when using DCHWT, compared to 5.8 percent when using Alex Net. The results indicated that the cheerful characteristic received the highest rating of almost 50%, while the natural, furious, and sad characteristics received equal ratings of 17% each. Fear and surprised were not detected in any way. The study may aid in the learning process or similar endeavors by contrasting those who are content with those who have different emotions. The study could be extremely beneficial in evaluating the quality of courses or presenters by demonstrating facial emotion recognition.

References

[1] N. Mehendale, Facial emotion recognition using convolutional neural networks (FERC), SN Applied Sciences, **2**, no. 3, (2020), 1–8.

[2] L. Devillers, HumanRobot Interactions and Affective Computing: The Ethical Implications, Robotics, AI, and Humanity, Springer, Cham,(2021), 205–211.

[3] B. A. Erol, A. Majumdar, P. Benavidez, P. Rad, K. K. R. Choo, M. Jamshidi, Toward artificial emotional intelligence for cooperative social humanmachine interaction, IEEE Transactions on Computational Social Systems, **7**, no. 1, (2019), 234–246

[4] F. Vanderhaegen, Weak signal-oriented investigation of ethical dissonance applied to unsuccessful mobility experiences linked to humanmachine interactions, Science and engineering ethics, **27,** no. 1, (2021), 1–25.

[5] M. Egger, M. Ley, S. Hanke, Emotion recognition from physiological signal analysis: A review. Electronic Notes in Theoretical Computer Science, **343**, (2019), 35–55.

[6] M. Imani, G. A. Montazer, A survey of emotion recognition methods with emphasis on E-Learning environments. Journal of Network and Computer Applications, **147**, (2019), 102423.

[7] K. Shao, R. Pekrun, L. J. Nicholson, Emotions in classroom language learning: What can we learn from achievement emotion research? System, **86**, (2019), 102121.

[8] R. Sarkar, S. Choudhury, S. Dutta, A. Roy, S. K. Saha, Recognition of emotion in music based on deep convolutional neural network, Multimedia Tools and Applications, **79**, no. 1, (2020), 765–783.

[9] B. J. Abbaschian, D. Sierra-Sosa, A. Elmaghraby, Deep learning techniques for speech emotion recognition, from databases to models. Sensors, **21**, no. 4, (2021), 1249.

[10] A. Aslam, K. Hayat, A. I. Umar, B. Zohuri, P. Zarkesh-Ha, D. Modissette, B. Hussian, Wavelet-based convolutional neural networks for gender classification. Journal of Electronic Imaging, **28**, no. 1, (2019), 013012.

[11] A. M Abduldaim , A. A. Abdulrahman, F. S. Tahir, The effectiveness of discrete hermite wavelet filters technique in digital image watermarking, Indonesian Journal of Electrical Engineering and Computer Science, **25**, no. 3, (2022), 1-1x.

[12]] A. A. Abdulrahman, F. S. T. Al-azawi, Face Detection for Color Image Based on MATLAB. In Journal of Physics: Conference Series, IOP Publishing, **1879**, no. 2, (2021), 022129.

[13] A. A. Abdulrahman, F. S. Tahir, Face recognition using enhancement discrete wavelet transform based on MATLAB, Indonesian Journal of Electrical Engineering and Computer Science, **23**, no. 2, (2021), 1128.

[14] L. Galib, F. S. Tahir, A. A. A Abdulrahman, Detection Face parts in image using Neural Network Based on MATLAB, Engineering and Technology Journal, **39**, no. 1, (2021), 159–164.

[15] R. I. Bendjillali, M. Beladgham, K. Merit, A. Taleb-Ahmed, Improved facial expression recognition based on DWT feature for deep CNN, Electronics, **8**, no. 3, (2019), 324.

[16] M. B. Priatama, L. Novamizanti, S. Aulia, E. B. Candrasari, Hand gesture recognition using discrete wavelet transform and convolutional neural network. Bulletin of Electrical Engineering and Informatics, **9**, no. 3, (2020), 996–1004.

[17] K. Mohan, A. Seal, O. Krejcar, A. Yazidi, Facial expression recognition using local gravitational force descriptor-based deep convolution neural networks, IEEE Transactions on Instrumentation and Measurement, **70**, (2020), 1–12.

[18] M. Z. Alom, T. M. Taha, C. Yakopcic, , S. Westberg, P. Sidike, M. S. Nasrin, V. K. Asari, A state-of-the-art survey on deep learning theory and architectures. Electronics, **8**, no. 3, (2019), 292.

[19] M. Sahu, R. Dash, A survey on deep learning: convolution neural network, Intelligent and Cloud Computing, Springer, Singapore, (2021), 317–325

[20] V. Franzoni, G. Biondi, D. Perri, O. Gervasi, Enhancing mouth-based emotion recognition using transfer learning. Sensors, **20**, no. 18, (2020), 5222.

[21] Q. T. Ngoc, S. Lee, B. C. Song, Facial landmark-based emotion recognition via directed graph neural network. Electronics, **9**, no. 5, (2020), 764.

[22] A. Khan, A. Sohail, U. Zahoora, A. S. Qureshi, A survey of the recent architectures of deep convolutional neural networks. Artificial intelligence review, **53**, no. 8, (2020), 5455–5516.

[23] E. Ledinauskas, J. Ruseckas, A. Juršź2nas, G. Buračas, Training deep spiking neural networks. arXiv preprint arXiv, (2020), 2006.04436.

[24] J. Chai, H. Zeng, A. Li, E. W. Ngai, Deep learning in computer vision: A critical review of emerging techniques and application scenarios. Machine Learning with Applications, **6**, (2021), 100134.

[25] Q. I. Xiao-xu, J. Wei, Application of wavelet energy feature in facial expression recognition, 2007 International Workshop on Anti-Counterfeiting, Security and Identification (ASID), IEEE, (2007), 169–174.

[26] L. Zhao, G. Zhuang, X. Xu, Facial expression recognition based on PCA and NMF, 2008 7th World Congress on Intelligent Control and Automation, IEEE, (2008), 6826–6829).

[27] X. Feng, M. Pietikinen, A. Hadid, Facial expression recognition based on local binary patterns. Pattern Recognition and Image Analysis, **17**, no. 4, (2007), 592–598.

[28] F. Y. Shih, C. F. Chuang, P. S. Wang, Performance comparisons of facial expression recognition in JAFFE database. International Journal of Pattern Recognition and Artificial Intelligence, **22**, no. 3, (2008), 445–459.

[29] H. Alshamsi, V. Kepuska, H. Meng, Real time automated facial expression recognition app development on smart phones, 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference, IEEE, (2017), 384–392.

[30] C. F. Liew, T. Yairi, Facial expression recognition and analysis: a comparison study of feature descriptors, IPSJ transactions on computer vision and applications, **7**, (2015), 104–120.

[31] M. A. H. Akhand, S. Roy, N. Siddique, M. A. S. Kamal, T. Shimamura, Facial Emotion Recognition Using Transfer Learning in the Deep CNN, Electronics, **10**, no. 9, (2021), 1036.