Real-life Dynamic Facial Expression Recognition: A Review

Sharmeen M. Saleem1*, Subhi R. M. Zeebaree2, Maiwan B. Abdulrazzaq3

1 Duhok Polytechnic University, 2 Duhok Polytechnic University, 3 University of Zakho,

Emails: sharmeenbarwary2020@gmail.com, subhi.rafeeq@dpu.edu.krd, maiwan.abdulrazzaq@uoz.edu.krd

Abstract. In emotion studies, critiques of the use of a static facial expression have been directed to its resulting from poor ecological validity. We conducted a study of studies in the present work, which specifically contrasted recognizing emotions using dynamic facial expressions. Brain imaging experiments and behavioural studies with associated physiological research are also included. The facial motion appears to be connected to our emotional process. The findings of laboratory brain injury experiments also reinforce the concept of a neurological dissociation between static and dynamic expression mechanisms. According to the findings of electromyography studies of dynamic expressions of affective signals, those expressions evoke more extreme facial mimic physiological responses. Studies significantly affirm the essence of dynamic facial gestures.

Keywords: Facial behavior analysis, Facial expression recognition, 3D facial surface, 3D facial surface sequences (4D faces).

1. Introduction

The most reliable and naturally important way for people to convey feelings, explain and explain variations in meaning, express intentions, and, in a more general context, control relationships with the world and others is the facial expression [1-3]. Mehrabian observed that 7% of knowledge moves between people through writing, 38% through voice, and 55% through facial expression [4, 5]. The Ekman & Friesen Facial Action Coding Scheme (FACS), released in 1978, describes seven main facial expressions expressed by people without language, including terror, coldness, surprise, disgust, good luck, truth, and neutrality. This system is known as the threshold for facial expression recognition (FER) [6].

Various applications involve the comprehension of human emotion by facial expression, which include human-computer interaction, robotics, and health care, among others [7, 8]. However, emotional recognition in our everyday lives is important for social contact, and emotions play an important role in deciding human actions [9]. However, the affective status of the pupil can be instantly interpreted by FER in the area of schooling. It lets teachers recognize their students’ academic interests to include appropriate instructional methods to increase teaching effectiveness [10, 11]. The method of observing movements of people such as tourists who enjoy museum shows and thinking and evaluating what they see is relevant for a long time, which provides important synanthropic information and time structure to identify its categories [12]. FER also has a wide variety of social life uses, such as smart protection, lie detection, and smart medical practice.

This specific area involves systems to classify the fundamental human emotions with current artificial intelligence algorithms, particularly neural networks (FACS)[13]. Image processing is a major step for
image recognition [14]. They contain unnecessary specifics that harm the performance of a system. Before preparing the feed-forward structure for the convolution neural network, there needs to be a preprocessing phase to normalize and integrate the face's information [15] in addition to several processes, such as face recognition, image resizing, rotation correction, and image cropping [16]. For example, eye movement correction ensures two eyes staying on a horizontal line [17]. Resizing, cropping, and border matching are done vertically at each eyebrow and jaw level and horizontally between the two ears [18]. This section includes facial orientation, gray image transfer, 2-D noise-removal adaptive filtering, image sharpening employing sharp masking, and data increase.

In preprocessing, input data quality (image) is improved, and redundancy is reduced or eliminated. Next, an image input RGB of m x n size is read and converted into a gray image with the standard equation [19, 20]. The circumference of the face was detected with Hear [21], Cascade pictures library. Those rectangular facial expressions were then cut off and reported to the same scale. The pictures' pixel values have also been transformed into 64x64 gray images to be put in neural networks. This is required to prevent the excess density of the neural networks [22]. In a real-life situation, data collection for images can be captured in various conditions, such as different directions, locations, sizes, and visibility [23]. Thus in such raw images, the conventional preprocessing technique such as standardization, cutting, and centralization enhances recognition of images during any experimental period [24].

In this article, we analyze recent advancements in detecting emotions with the aid of numerous methods, algorithms, and architectures to detect data from dynamic facial expressions. However, it provides a summary of various processes involved in real-life FER. The latest observations are discussed in our review of the topics and contributions.

The remaining part of this paper is organized as follows: Section 2 Background Theory. Section 3 Literature review Section 4 Discussion, then a general conclusion, is provided in section 5.

2. Background Theory

Analysis of facial expression covers both face motion estimation and expression recognition. Automatic Facial Expression Analysis (AFAE) consists of three steps: facial acquisition, facial and facial representation data removal, and facial expression identification [25, 26]. Face retrieval is a processing step for the input images or sequences that automatically find the Face Field [27]. You can detect faces for each frame or simply detect the faces in the first frame and then monitor the faces in the rest of the video series [28, 29]. The head-finder, head monitoring, and posture estimate can be used in a face expression analysis method for handling massive head movements [30]. The next move is to extract facial differences created by facial expressions and show them [31]. After the face is located, two major categories of facial extraction approach for emotion analysis: geometric feature-based methods and appearance-based approaches. The facial components' geometric features show form and position (including mouth, eyes, brows, nose, etc.) [32]. A function vector that describes the face geometry is derived from the facial components or face characteristics points [33]. Image filters like Gabor wavelets are applied to either the entire or particular regions in an image to removing a functional vector using appearance-based approaches [34]. Depending on the various methods of extraction of facial features, the head rotation's influence in the plane and different facial dimensions may be removed before the extraction or before the recognition process [35].

2.1. Facial Expression Recognition Methods

Many models to face are defined in two different ways: categorical and dimensional. A large portion of recent studies focused on the recognition of categorical feelings organizing emotions into different groups [36]. Although these methods will accurately recognize face expression, the categorical emotional representations do not encompass the complete spectrum of human emotions [37, 38].

2.2. Individual Differences in Subjects

The shape, texture, color, face, and scalp hair are different in terms of sex, ethnicity, and age. For example, children have smoother, less textured skin and often lack facial hair in their scalp or brows
[39]. The opening and contrast between iris and sclera vary significantly among Asians and Northern Europeans, which can have a more global effect on the strength of eye monitoring or facial analysis [40]. The dark facial characteristics of beards, eyeglasses, or jewelry. Such individual appearance differences can have significant effects on face analysis. There are few attempts to explore their influence. One exception was a survey by [37], which found that algorithms optimized for young adults for optical flow and high-gradient detection were less successful when used with children. All of the declared differences in the face analysis between children and adult patients have been influenced by reduced skin texture, increased fats, juvenile facial shape, and lack short furrows [41].

2.3. Attention Inference
Care is commonly referred to as a major role in the system of human perception. One important feature of a human visual system is that a whole scene is not processed at once [42]. Instead, people use a sequence of partial views and concentrate selectively on key components, to capture a better visual structure [37, 43].

3. Literature review
Mahmood, Mayyadah et al. in [39] provided a consistency assessment of several controlled classifiers to identify facial expression based on minimum chi-square characteristics. These are the most emblematic and prominent characteristics that have an actual determination meaning. Six classifiers, like MLP, SVM, decision-making board, a random forest, radial bias, and KNN, use the top six features to determine the most exact number using the minimum features. This is achieved by the analysis and assessment of the results of the classifiers. CK+ is used as a dataset for analysis. The most precise classification among the others is the random rain forest with a complete accuracy ratio of 94.23 percent.

Abdulrazaq, Maiwan et al. in [36] Focusing on using a minimal number of attributes aims to inspect the exactness of six classifiers based on the Reliever-F approach for function collection. MLP, Random Forestry, Decision Tree, SVM, KNN, and radial basis functions are all classifications under which the paper is being inspected. The experiment shows that, for the application of CK+ data collection, K-Nearest neighbor is the most reliable classifier with an overview of 94.93 percent.

[Wu et al.] in [9] suggested a two-stage Fuzzy Fusion based convolution neural network (TSFFCNN) network for the identification of complex emotions, the extraction of representative facial expression characteristics, and efficient modality fusion. Initially, the Local Binary Pattern (LBP) and the spectrogram extract complex low-level emotional characteristics to acquire Spatio-temporal information from face expression data. To identify strongly biased emotional recognition functions, two Deep CNNs have been built to derive high-level emotional semantic facial expression characteristics. Besides, the TSFFs are generated by the combination of canonical correlation analysis and a fuzzy broad learning system to consider the similarity and discrepancy between various modal characteristics. The TSFFCNN is then established. Not only can it extract discriminatory emotions that include Spatio-temporal information, but it can also effectively combine facial expression and language modalities at the function and decision-making level. BESIDES, the TSFFCNN will manage conditions where the contribution of each modality of data to emotional recognition is very imbalanced. Benchmark database tests are performed in order to determine the feasibility of the proposed approach, and experimental findings of the Surrey Audio-Visual Expressed Emotion (SAVEE), eINTERFACE05, and Acted Facial Expressions in the Wild (AFEW) datasets are carried out using the proposed TSFFCNN 99.79%, 90.82%, and 50.20%.

Dino, Hivi et al. in [44] Offered FER comparative methodology based on three feature selection methods: correlation, gains ratio, and data gain to determine the most outstanding characteristics of face images using MLP, Naive Bayes, decision tree, and KNN algorithms. These classifiers are used for the mission of expression recognition and for comparing their proportional performance. The main aim of the provided approach is to determine the most effective classifier based on the minimum acceptable number of features by analyzing and comparing their performance. The approach’s key objective is to assess the most efficient classifier by evaluating and measuring its efficiency based on a minimum reasonable number of characteristics. The solution presented was extended to CK+. The
experimental results indicate that KNN with 91 percent precision with just 30 features is best classified. Chen et al. in [12] built a Visual Geometry Group (VGG) model paired with a video over a long period to predict a client's emotional mixture in a service site. The key contributions are: 1) using the framework of Bidirectional Long Short-Term Memory (BiLSTM) to provide knowledge about input sequences, mixing long emotional variations with short facial output. 2) the dataset has been developed to address the issue of lack of facial expression data sets in device scenarios and service scenarios and is based on several studies. Comprehensive experiments have shown that long-term variations in expression and keyframe acquisition techniques previously overlooked by researchers have been successful. This helps the time structure of the phase of transition be investigated and the framework for concentrating on image sequences creation. Le, Trang et al. in [45] Presented the compact system used to obtain insightful features on specific regions of interest, in conjunction with a CNN concept of rank pooling, referred to as dynamic images, used to identify micromotion there. Only a certain number of localized facial areas based on observed dominant muscular changes are derived from the dynamic picture. The CNN models are fit for an emotional grouping challenge for the final feature portrayal. Liu et al. in [7] proposed the Siamese Action Units Network (SAANet) to construct a metric FER learning system that combines Space and time care modules: Action Unit Focus Module (AU Attention Module) and Vigilant Pooling Module. The CNN- Recurrent Neural Network (RNN) baseline models build a metric learning system for detailed and finely differentiated facial expressions. In the space market, they create a new AU mechanism for aggregating spatial qualitative knowledge on the critical regions from long-range dependences more accurately and efficiently. To fulfill this FER challenge, a particular pair-wise sampling technique for this metric learning system is proposed. A conscientious grouping module is used to capture temporal similarities in the video frames in the temporal domain. The four benchmark datasets (expanded Cohn-Kanade (CK+), Oulu-CASIA, Maja Pantic, Michel Valstar and Ioannis (MMI), and AffectNet) have experimented. The experimental findings suggest that the architecture introduced exceeds the state-of-the-art techniques. Chen et al. in [46] Construct a C3D-based network architecture for the extraction of space and time from complex facial expression sequences 3D-Incept-ResNet. A Spatial-time and Channel Care Module (STCAM) is proposed to leverage the integral spatial and channel-saved relationships between features extracted. The proposed STCAM is precisely designed to measure a channel and a spatial-temporal map to boost the functionality with more representative features in the corresponding dimensions. Three common dynamic facial expression recognition datasets, CK+, Oulu CASIA, and MMI, test this process. Experimental findings suggest that we do well or equal with state-of-the-art methods. Perveen et al. in [1] Proposed the dynamic kernel-based facial expressions, absorbed by local space-time representations in facial expressions captured in the universal Gaussian mix-model (uGMM). These complex kernels are used to preserve local correlations by using uGMM statistics to improve the same name's global meaning. The usefulness of dynamic kernel representation by three separate dynamic kernels is demonstrated: explicit mapping based on three common facial expression datasets, MMI, AFEW, and Binghamton–Pittsburgh 4D Spontaneous (BP4D). Evaluations show that the most biased among dynamic kernels are probability-based kernels. However, the intermediate matching kernels are more efficient than the other two representations in terms of computational complexity. Ni et al. in [47] proposed a DFFD-driven 3D face dynamic expression algorithm. By capturing authentic facial expressions to push a universal neutral 3D model of the virtual face, a convincing 3D animated facial expression is synthesized following the real actors' facial gestures. The 3D simulated face synthesized by this approach shows that there are certain limitations between visual, auditory, and content to be conveyed in the related process. The device also relies on the output of live actors. [Wen et al.] in [5] proposed a novel approach for understanding facial expression in various ways. Second, people are easy to understand those phrases, whereas some are harder to recognize. Driven by this intuition, the idea for a new loss function expands the distance between samples from easily confusing groups. Second, human learning is split into several stages, and each stage has a particular learning target. A new deep neural network for FER is suggested to realize the above ideas better,
incorporating the covariance body and residual network units into the deep convolution neural network to achieve dynamic objective learning better to describe the related loss function in the different levels. The suggested lack of domain knowledge alleviates the issue of directional uncertainty between recognition categories. The complex expectations of learning will better direct the neural network to prevent local optimization problems.

Xue et al. in [48] learned expression-sensitive features that can not only deliver a comparable result with state-of-the-art identification but can also be understood by humans. Second, spatial time characteristics histogram of oriented gradients (HOG3D) are omitted from the patch sequences of local depth to reflect facial expression's dynamic. A two-phase selection process is then proposed to determine the facial components that better discriminate between gestures. In a hierarchical classifier for FER to validate the resulting facial components' efficacy, expressive features from the corresponding region are fed. In the Binghamton University 4D Facial Expression (BU-4DFE) benchmark database, the proposed approach is tested, and results show that studied expressive-sensitive features will achieve equivalent reconnaissance efficiency with current methods. Besides, the resulting HOG3D functionality can produce a semantic understanding of dynamic recognition after function discovery.

Deng et al. in [49] developed a model based on a deep convolution neural residual network consisting of stem layer, 3D reset configuration, GRU unit, and Iceland loss function. It has a high identification rate and a low time, and a good generalization capacity for preparation. Moreover, the model was evaluated on three sets of data (CK+, MMI, AFEW) and achieved high recognition rates. Moreover, the residual CNN structure extracts key characteristics. The link to the jumping layer extracts the time relation between the image sequences and applies the new loss function and an optimization algorithm to improve network recognition and generalization rates.

Ricciardi et al. in [50] A variety of geometrical features indicating distances in the bottom half of the video series between the markers can be taken to see if they shift over time when specifying a clear punishment. The author has training for a completely linked feed-forward neural network to verify whether this time-dependent descriptor adequately discriminates for precise and robust identification. The tests carried out show a state-of-the-art consistency of identification of more than 98 percent, with a remarked intensity to the sentence's pronouncement and strong freedom from the option of the sentence, demonstrating the validity of this compact and distinctive facet dynamic descriptor.

Alazrai et al. in [51] proposed a two-layered paradigm for human emotional states identification. To examine temporal human emotional dynamics using spatially decoupled contexts. The first layer assigns a mark to each frame of the input video that reflects the frame's time stage and the emotional state of the input video. To that end, they use the disconnected space background for training a selection of Support Vector Machine (SVM) classifiers to identify the ESTP of each video input frame. The second layer uses Dynamic Time Warping (DTW) to classify the sequence of frame labels created in the first layer into one of the following seven emotional conditions: frustration, indignation, disgust, anxiety, satisfaction, sadness, and surprise. The DTW's ability to resolve broad differences in periods in an emotionally moving video is critical in dealing with complex time complexities, such as the intensity with which one communicates emotion and the difference during each period. Such a recognition scheme will simplify the issue of the recognition of human emotions by working separately with spatial and temporal approaches.

Zhi et al. in [8] introduced the famous applied deep learning approach with sparse coding, i.e., Long-Short Term Memory (LSTM) for the complex facial learning of facial expression sequences is based upon Iterative Hard-Thresholding Algorithm (ISTA). This network structure called SLSTM will automatically learn facial features on the video sequence and create the corresponding sparse codes, and eventually, the product of recognition is a classification material. The sparse representation has been implemented into facial feature learning and dynamic FER. The experimental findings on CASME II and SMIC showed that this approach was superior.

Baddar et al. in [52] Try checking the functions Approximate expressive intensity is employed to encode the effect of the facial expression on the fly using the LSTM to encode appearance-deleted dynamics features, learning with a novel objectives feature consisting of three objective words.
Prediction improved by using incomplete sequences of expression. The proposed dynamics suppressed presence function enhances the prediction with incomplete video sequences required for prediction on the fly. In comparison, the finding shows that the dynamics suppressed by the suggested presence results in a greater generalization of previously unseen appearance variants (e.g., different races or variations in illumination conditions).

Verma et al. in [53] Proposed a complex micro-expression representation to retain video face movement details in a single frame. It also suggests the Lateral Accretive Hybrid Network (LEARNet) captures facial expression micro-level characteristics. LEARNet refines the excellent features of recognition by integrating AL in the network. The AL answer contains the hybrid feature maps created by previously linked layers of convolution. LEARNet architecture also incorporates the interconnection between convolution layers, which helps preserve small yet influential information about facial muscles. The visual answers of the proposed LEARNet demonstrate the method's functionality by retaining both high and micro-face functions. The performance of LEARNet is measured on four benchmarks: CASME-I, CASME-II, CAS(ME)*2, and SMIC. The experimental results after the research shows significant improvement in CASME-I, CASME-II, CAS(ME)*2, and SMIC datasets by 4.03 percent, 1.90 percent, 1.79, and 2.82 percent, respectively.

Li et al. in [54] proposed a new 4D FER Dynamic Geometrical Imaging Network (DGIN) (FER). A 3D film representing a series of face scans first estimates their differential geometry quantities and produces geometric pictures, like Depth Images (DPI) (SID). These geometric images are then fed into DGIN for end-to-end prediction and preparation. DGIN consists of a temporal short-term pooling layer for dynamically producing geometric images, multiple repetitions of convolutions + ReLU + pooling layers for extraction of face-space features long-term time pooling layer, followed by completely linking layers and joint loss layer for dynamic feature map fusion. The long and short-term two-stage sliding window scheme is implemented for data growth and temporary pooling in the training process. Meanwhile, for the more discriminatory language, a mutual loss combining loss of cross-entropy with loss of three sections. The evaluation phase is used for the whole video series of certain geometric images, which have a stronger effect in the deeper network for calculating the similarity. In this stage, the sliding windows are only short-term. In the end, the projected expressive scores of different geometric images are integrated. Experimental findings published in the BU4DFE database demonstrate the feasibility of the proposed solution.

Vithanawasam et al. in [55] proposed a way to understand the complex face of the person and the emotion of the high body by the Microsoft Kinect sensor service robot. Four complex emotions will be classified simultaneously in the proposed framework. The unit has been validated in an indoor environment for preferred emotional contact. Emotional recognition is based on learning methods. He could effectively identify the rage, terror, neutral and bored emotions in both ways when he sat down and walked in front of the Kinect sensor. The depth sensor of the Kinect sensor is sensitive to sunlight and functions mainly indoor conditions. Since the face and the upper body are turned to a 15° side, the machine can recognize the feeling. When a pre-trained human turns his face to the Kinect sensor, the machine begins to detect emotions.

Dong et al. in [56] proposed the CNN structure with comparatively shallow densely connected short FER roads. We use dense networking through poolies to facilitate feature sharing in the shallow CNN layout instead of using transformation layers into down-sample feature maps between thick lines. Instead. Comprehensive tests demonstrate our approach to achieve competitive efficiency in CK+ and Oulu-CASIA benchmark datasets, inserting short distances for each convolutional layer from all earlier layers. By incorporating dense networking for sharing of features, we want to reduce the overcrowding of limited datasets. Short paths from the previous layers to the latter promote information flow and enable the network to acquire compact recognition features. However, we do not use dense intermediate blocks and transition layers, unlike conventional dense structures.

Wang in [57] proposed the smart image-based face recognition technology that could extract the signature features of the face and sample the face details such as pupils, eyelids, facial traits, and facial phrases. An adaptive prototype matching approach is used to divide the face frame, use the 3D dynamic scan method for extracting pupil and eyelid information, and use matching filter detection to obtain the best pixel image pixel function. Dynamic 3D material is derived from the human face.
These models may be the gray picture of each physiological function region's whole face and the red level image. Owing to the corresponding outcome of the template, facial recognition may be accomplished. The experimental research is eventually performed. The benefits of this approach are shown in the enhancement of dynamic facial recognition. The simulation results suggest that the suggested method is more reliable. It will improve facial recognition accuracy and resolution.

Li et al. in [58] proposed a modern approach for the identification of complex facial expressions using geometric and texture elements. The facial landmarks and texture adjustments on photographs in pairs are used to perform complex tasks of identification of the face. For one series of facial expressions, images are produced in pairs between the first and following frames. The incorporation of geometric and texture characteristics further improves the facial representation. Finally, the SVM is used to recognize facial expressions. Experiments in the (CK+) demonstrate that our proposed approach can compete with other approaches. This analysis's breakthrough is to document variations in expression sequences, which are used to represent expressional changes. The combination of these forms of features further improves facial representation. Finally, the SVM is used to recognize facial expressions.

Yu et al. in [59] proposed a monocular video FER device in real-time. First, the online statistical appearance model (OSAM) is enhanced to adaptive, weighted (AW-OSAM) for facial emotion monitoring to mitigate lighting variance complexity. In those, 13 basis point light locations are built to model each frame of film's illumination. Secondly, the static and dynamic information is derived and integrated by integrating the monitoring mechanism into the particle filtering process, which yields a higher facial expression recognition score than the static or dynamic form. The experimental findings affirm the system's benefit even though the head and occlusions are important.

Seo et al. in [60] suggested a method for synthesizing 3D dynamic face expression images based on regression. In this approach, the coefficients are calculated based on training data from a regression. They strive to find a function that maps two separate views or expressions for image coefficients. They also suggested using textural details in joint subspace learning to synthesize various facial forms. These two common subspaces learning methods have made it possible to achieve an accurate 3D dynamic picture synthesis.

Liu et al. in [61] developed three major extensions. Next, they generalize the structure for creating a mid-level term for different low-level 2D/3D descriptors. Secondly, they include a more detailed comparison and discussion of numerous universal manifold model (UMM) learning techniques, including local modes integration in the UMM training process and low-level UMM exercise approaches. Third, more in-depth tests are performed to test and compare each system's portion with other advanced algorithms. The proposed graphic representation demonstrated its dominance over standard approaches for identifying facial expression based on the video, based on four state-of-the-art facial expression benchmarks.

Zebari, Dilovan Asaad et al. in [29] Tried to create an improved way of preprocessing Computer-aided diagnosis (CAD) by using the conventional CAD technique to improve contrast and noise reduction approaches. A double increase in contrast after noise reduction improved mammogram image; for mammogram picture mdb072, the proposed method's mean-square error (MSE) was 1.44% decreased. The decrease in MSE raises the PSNR to 0.16%. Many mammograms were checked, and the results indicate that, compared with current approaches, they increase the contrast, decrease the mean square error and increase the signal-to-noise relation in the peak.

Dino, Hivi Ismat et al. in [62] Presented the automated face expression recognition device capable of identifying all eight main facial expressions (normal, joyful, furious, disdainful, unhappy, fearful, and unpleasant). For FER identification alone, several facial expressions were proposed. For device validation, the (CK+) dataset is used. The Viola-Jones algorithm is used for facial recognition. As a descriptor for extraction of feature images of expressional faces, the HOG is used. HOG was applied to minimize the size of the features and to acquire the most important characteristics. The presented approach was finally contrasted with the three classifiers: (SVM), K-Nearest Neighbor (KNN), and Multilayer Perceptron Neural Network (MLPNN). The experimental findings showed that the presented approach has an identification rate of 93.53% using the MLP classifier, 82.97% for the use of the MLP classifier, and 79.97% for using the KNN classifier.
4. Discussion

Technology to evaluate 3D facial expression is still in its infancy. Shortly, with many works planned as the present technological advancements make it simple and affordable to obtain high levels of 3D data quality. There are, however, some unresolved problems in this field.

The previous section shows that researchers have employed different types of techniques and algorithms in Facial Expression Recognition fields. Scientists have issued a list of their recommendations upon studying theirs. Table 1 of the paper explains the explanations explained in the paper. The plan should include a comparison of the success and commonalities in the methodology of Facial Expression Recognition. The researcher used the Objectives, the Research Used Tools and Technique, used Algorithms, the Significant Satisfied Aims approach, and data set to analyze the results.

It is evident from the table that some researcher objectives are better to perform dynamic objectives learning, allow people to recognize and trust understanding of the meaning, improve the accuracy. In contrast, another work in the detect time phases of human facial expressions automatically, enhancement stage based on wavelet to produce an image with quality that can improve the segmentation and features extraction stage, rather than recognize all eight facial expressions. Depending on the scientific area of facial expressions, Algorithm, CNN, DNN, Histogram of Directed 3D Gradients, MLP, SVM, decision-making board, a random forest, radial bias, and KNN. Another method Tool/technique is used, such as LBP, (UGMM), (MIK), (IMK), and minimum features selected by chi-square. The Data set used SAVEE, CASMEII, CK+, and Oulu-CASIA. By using this methodology and techniques, both researchers have strong structures, frames, and functions. However, researchers’ trend has been oriented for modern Facial Expression Recognition fields.

Table 1. Formatting sections, subsections, and subsubsections.

| Ref. Year | Objectives                                                                 | Classifier                                                                 | Tool/technique                  | Significant results                                                                 | Data set used | Accuracy |
|-----------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|---------------------------------|-------------------------------------------------------------------------------------|---------------|----------|
| 2021      | a method of classification and feature selection for Facial Expression Recognition from sequence facial images. | MLP, SVM, decision-making board, a random forest, radial bias, and KNN. | minimum features selected by chi-square | provided a consistency assessment of several controlled classifiers to identify facial expression based on minimum chi-square characteristic s. | CK+           | 94.23%   |
| 2021      | Efficient system for Facial Expression collection and classification Sequence of face pictures | MLP, Random Forestry, Decision Tree, SVM, KNN, and radial basis functions | Reliever-F approach            | Focusing on the use of a minimal number of attributes, it aims at inspecting the exactness of six classifiers based on the Reliever-F approach for function collection | CK+           | 94.93%   |
| 2020      | To achieve effective modality                                           | Fuzzy Fusion-based two-stage neural network                                | LBP-TOP and spectrogram         | The condition in which each modality data SAVEE, eINTERFACE05, and AFEW            | SAVEE, AFEW   | 99.79%, 90.82% |


| Reference | Year | Description | Approach | Result |
|-----------|------|-------------|----------|--------|
| [44] | 2020 | The eight fundamental expressions of human faces to be recognized | using MLP, Naïve Bayes, decision tree and KNN algorithms | CK+ 50.28% |
| [12] | 2020 | To forecast a customer's collective emotions over some time in a service area. | - BiLSTMAttention-based network model - Combine VGG16 to propose a BiLSTM-Attention Network Model (VGG16-BiLSTM-Attention) known as VBA | CK+ |
| [45] | 2020 | To enhance the spontaneous detection of facial micro-expression by sophisticated extraction techniques by hand | (CNN) used for the detection of microexpressions in complex images requested - The magnification technique of the face movement is extended to input sequences - rank pooling to enter complex images is introduced. | CASMEII, SMIC and SAMM. |
| [7] | 2020 | To allow a single function to interpret the action units' features (AUs) (eyebrows, eyes, nose, and mouth). | - CNN and RNN -VGG16 structure - Standardization batch (BN) layers - Bidirectional Short Memory (BiLSTM) | CK+ OuluCASIA MMI, AffectNet |
| [46] | 2020 | To capture the complex evolution of the face | - A Channel and Spatial-Temporal Focus Module (STCAM) - 3D-Inception-ResNet network architecture | CK+ OuluCASIA and MMI |

Fusion (TSFFCNN) contribution to emotion recognition is strongly imbalanced is handled well by TSFFCNN.
|   |   |   |   |
|---|---|---|---|
| [1] 2020 | To explore the development of local expressions caught from various areas of the face. | Dynamic kernel-based facial expression representation | (UGMM), (MIK), (IMK) | For any expression recognition program based on accuracy or measurement time, a dynamic kernel is an important option |
| [47] 2019 | Construct functional three-dimensional model and synthesize true facial expression | - DFFD-driven algorithm (Dirichlet Free-Form Deformation) | In encounters between the person and the virtual beings, the system will replicate the most typical emotional contact scenes. | --- |
| [5] 2019 | To better perform dynamic objectives, learning -Enlarge the distances between samples from easily confused categories | - a new DNN for facial expression - a new loss function is proposed to enlarge the distances between samples | Integrates deep the pooling covariance and residual network units | The neural network built will prevent the loss of gradients and |
| [48] 2019 | To allow people to recognize and trust understanding of the meaning. | The Histogram of Directed 3D Gradients (HOG3D) Spatio-temporal properties were derived from local depth patch sequences. | - use the BU-4DFE benchmark database tool. | Learned expression-sensitive functions can produce a comparable result in recognition with current approaches. |
| [49] 2019 | To improve the accuracy. | - 3D Convolutional Neural Network (CNN) video FER method | - Stem Layer, 3D Layout-ResNets, GRU layer, Dropout layer, Island layer, Softmax layer Layer | can best disregard the consequences of face expression variability and adaptation and gain greater precision of identification and generalization? |
|   |   |   |   | (CK+, MMI, AFEW) | CK+ 94.39%, MMI 80.43%, AFEW 82.36% |
| Reference | Year | Description | Methodology | Evaluation | Dataset | Result |
|-----------|------|-------------|-------------|------------|---------|--------|
| [50]      | 2019 | This is a salient and ultimately better biometric marker since it is more difficult to forge a time variable descriptor than a static one. | - architecture of DFF NN (deep feed-forward neural network) - geometric time series descriptor arranged. | Identification modality tests resulted in 98.2% of average precision, 0.64% of equivalent error rates, and exceptional robustness in the phrase's pronunciation. | OuluVS visual-database | 98.2% |
| [51]      | 2019 | To detect time phases of human facial expressions automatically - define the equivalence of multiple video frames | - pyramid of histograms of oriented gradients (PHOG) - (ESTP), (SVM), (DTW) | Validate the method efficiency, carry out computer simulations extensively, and show an average grading precision of 93.5%. | CK+ | 93.53% |
| [8]       | 2019 | Reduce LSTM difficulty in sequence processing | (ISTA), SYSTEM Algorithm | In detailed results with low computational complexity, superiority showed the proposed system. | CASME II and SMIC | ------ |
| [52]      | 2019 | Reducing prediction time and making "on-the-fly" prediction (as frames are fed to the system) | recurrent neural networks (RNNs) -LSTM | The approach suggested will address a problem with the prediction delay and be applied on the fly (i.e., interactive environments). | ------ | ------ |
| [53]      | 2019 | -to retain details about facial movement in a video in a single frame. | Lateral Accretive Hybrid Network (LEARNet) in the area of facial expression to capture micro-level features -Constitute network accretion layers (AL) to produce hybrid characteristics maps | LEARNet achieved superior precision values compared to the latest state-of-the-art approaches. Our method has achieved outstanding efficiency. | CASME-I CASME-II CAS(ME)^2 SMIC | improvement of 4.03%, 1.90%, 1.79% and 2.82% as compared with ResNet on CASME-I, CASME-II, CAS(ME)^2 and SMIC datasets |
| [54]      | 2019 | To improve effectiveness. | -Dynamic Geometrical Image Network -Depth Images (DPI) | The findings obtained indicate the | BU4DFE database | 92.22% |
| Year | References | Description |
|------|------------|-------------|
| 2019 | [29]       | Focus on the enhancement stage based on wavelet to produce an image with quality that can improve the segmentation and features extraction stage. Computer-aided diagnosis (CAD) method. Try to create an improved way of preprocessing Computer-aided diagnosis (CAD) by using the conventional CAD technique to improve contrast and noise reduction approaches. 90.5% |
| 2019 | [62]       | Recognize all eight facial expressions, which are (normal, joyful, furious, disdainful, unhappy, fearful, and unpleasant). Viola-Jones, SVM, KNN algorithm. MLPNN, HOG, HOG. Presented the automated face expression recognition device capable of identifying all eight main facial expressions (normal, joyful, furious, disdainful, unhappy, fearful, and unpleasant). |
| 2018 | [55]       | To communicate with humans more effectively with service robotics. Pattern Recognition Neural Network (PRNN) classifiers. -Linear analysis of discriminants (LDA) -RGB-D (Kinect) Operation Robot Sensor. When a previously educated human turns his face towards the Kinect sensor, the machine begins detecting emotions. CK+ and Oulu-CASIA CK+ 93.53%, MLP 82.97%, KNN 79.97% |
| 2018 | [56]       | To decrease overfitting with minimal data from training. Shallow CNN structure. The planned network size can be changed easily without greatly impacting performance. CK+ 97.25%, Oulu-CASIA 76.25% |
| Year | Authors | Method/Approach | Details |
|------|---------|----------------|---------|
| 2018 | [57]    | Face signature characteristics are extracted. | To solve the issue of poor precision of dynamic FR identification. To boost dynamic FR efficiency. |
| 2018 | [58]    | Support Vector Machine(SVM) | To enhance the representation of facial expression. |
| 2017 | [60]    | Two-layer fuzzy support vector regression-Takagi–Sugeno (TLFSVR-TS) | It aims to allow robots to recognize and understand human emotions. |
| 2016 | [59]    | Adaptive weighted online statistical appearance model (AW-OSAM) | To reduce the lighting influence. To alleviate the illumination variation difficulty. |
| 2016 | [60]    | Regression-based joint subspace learning (RJSL) | To realize the accurate synthesis of 3D dynamic expression images. |
| 2016 | [61]    | The recent approach was introduced for the identification of complex facial expressions. | To Solve both problems (Dynamic representation of temporal |
5. Conclusion
This paper addressed important facial expression recognition. The literature review indicated that various active mechanisms contribute to facial expression recognition. There are many tools in this area, such as 3D method-matching dynamic scanning filter detector, geometrically normalized facial image (GNFI), minimum features selected by chi-square, LBP-TOP and spectrogram, fuzzy multiple support vector regression (SVR), Constitute network accretion layers (AL) to produce hybrid characteristics maps, and other types of computing tools (MCU). To attain the goal, various techniques and algorithms have been successfully used by the researchers. Therefore, efficient FER has been developed, such as providing a consistency assessment of several controlled classifiers for the identification of facial expression based on minimum chi-square characteristics. Offered FER comparative methodology based on three methods of feature selection: correlation, gains ration and data gain to determine the most outstanding characteristics of face images, proposed depiction demonstrated its dominance over standard approaches for facial expression identification dependent on video as measured on four state-of-the-art benchmarks for facial expression, allow people to recognize and trust understanding of the meaning. Adding to that, it can conclude that MLP, Random Forestry, Decision Tree, SVM, KNN, and radial basis functions are the most suitable technique for FER. Rather than Conditional Random Field (CRF) classifier, Dynamic Geometrical Image Network (DGIN), Pattern Recognition Neural Network (PRNN) classifiers.

6. References
[1] N. Perveen, D. Roy, and K. M. Chalavadi, "Facial Expression Recognition in Videos Using Dynamic Kernels," IEEE Transactions on Image Processing, vol. 29, pp. 8316-8325, 2020.
[2] R. Zhi, C. Zhou, T. Li, S. Liu, and Y. Jin, "Action Unit Analysis Enhanced Facial Expression Recognition by Deep Neural Network Evolution," Neurocomputing, 2020.
[3] S. Li and W. Deng, "Deep facial expression recognition: A survey," IEEE Transactions on Affective Computing, 2020.
[4] W. Mellouk and W. Handouzi, "Facial emotion recognition using deep learning: review and insights," Procedia Computer Science, vol. 175, pp. 689-694, 2020.
[5] G. Wen, T. Chang, H. Li, and L. Jiang, "Dynamic Objectives Learning for Facial Expression Recognition," IEEE Transactions on Multimdia, 2020.
[6] O. F. Mohammad, M. S. M. Rahim, S. R. M. Zeebaree, and F. Y. Ahmed, "A survey and analysis of the image encryption methods," International Journal of Applied Engineering Research, vol. 12, pp. 13265-13280, 2017.
[7] D. Liu, X. Ouyang, S. Xu, P. Zhou, K. He, and S. Wen, "SAANet: Siamese action-units attention network for improving dynamic facial expression recognition," Neurocomputing, vol. 413, pp. 145-157, 2020.
[8] R. Zhi and M. Wan, "Dynamic Facial Expression Feature Learning Based on Sparse RNN," in 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), 2019, pp. 1373-1377.
[9] M. Wu, W. Su, L. Chen, W. Pedrycz, and K. Hirota, "Two-stage Fuzzy Fusion based-Convolution Neural Network for Dynamic Emotion Recognition," *IEEE Transactions on Affective Computing*, 2020.

[10] J. Chen, Y. Lv, R. Xu, and C. Xu, "Automatic social signal analysis: Facial expression recognition using difference convolution neural network," *Journal of Parallel and Distributed Computing*, vol. 131, pp. 97-102, 2019.

[11] H. Dino, M. B. Abdulrazzaq, S. R. Zeebaree, A. B. Sallow, R. R. Zebari, H. M. Shukur, et al., "Facial Expression Recognition based on Hybrid Feature Extraction Techniques with Different Classifiers," *TEST Engineering & Management*, vol. 83, pp. 22319-22329, 2020.

[12] L. Chen, Y. Ouyang, Y. Zeng, and Y. Li, "Dynamic facial expression recognition model based on BiLSTM-Attention," in *2020 15th International Conference on Computer Science & Education (ICCSE)*, 2020, pp. 828-832.

[13] P. Babajee, G. Suddul, S. Armoogum, and R. Foogooa, "Identifying Human Emotions from Facial Expressions with Deep Learning," in *2020 Zooming Innovation in Consumer Technologies Conference (ZINC)*, 2020, pp. 36-39.

[14] S. R. Zeebaree, O. Ahmed, and K. Obid, "CSAERNet: An Efficient Deep Learning Architecture for Image Classification," in *2020 3rd International Conference on Engineering Technology and its Applications (ICICTA)*, 2020, pp. 122-127.

[15] M. NASRI, M. Hmani, A. Mtibaa, D. Petrovska-Delacretaz, M. B. Slima, and A. B. Hamida, "Face Emotion Recognition From Static Image Based on Convolution Neural Networks," in *2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 2020, pp. 1-6.

[16] K. B. Obaid, S. Zeebaree, and O. M. Ahmed, "Deep Learning Models Based on Image Classification: A Review," *International Journal of Science and Business*, vol. 4, pp. 75-81, 2020.

[17] A. A. Salih and M. B. Abdulrazaq, "Combining best features selection using three classifiers in intrusion detection system," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 94-99.

[18] Y. Tang, X. Zhang, X. Hu, S. Wang, and H. Wang, "Facial Expression Recognition Using Frequency Neural Network," *IEEE Transactions on Image Processing*, vol. 30, pp. 444-457, 2020.

[19] P. Kaviya and T. Arumugaprakash, "Group Facial Emotion Analysis System Using Convolutional Neural Network," in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)*, 2020, pp. 643-647.

[20] M. B. Abdulrazzaq and K. I. Khalaf, "Handwritten Numerals' Recognition in Kurdish Language Using Double Feature Selection," in *2019 2nd International Conference on Engineering Technology and its Applications (ICICTA)*, 2019, pp. 167-172.

[21] K. Mohan, A. Seal, O. Krejcar, and A. Yazidi, "Facial Expression Recognition using Local Gravitational Force Descriptor based Deep Convolution Neural Networks," *IEEE Transactions on Instrumentation and Measurement*, 2020.

[22] M. A. Ozdemir, B. Elagoz, A. Alaybeyoglu, R. Sadighzadeh, and A. Akan, "Real Time Emotion Recognition from Facial Expressions Using CNN Architecture," in *2019 Medical Technologies Congress (TIPTEKNO)*, 2019, pp. 1-4.

[23] M. B. Abdulrazzaq and J. N. Saeed, "A comparison of three classification algorithms for handwritten digit recognition," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 58-63.

[24] G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, and H. Wang, "Enhanced deep learning algorithm development to detect pain intensity from facial expression images," *Expert Systems with Applications*, vol. 149, p. 113305, 2020.

[25] S. Han, J. Hu, W. Li, S. Zhao, M. Chen, P. Xu, et al., "From structure to concepts: The two stages of facial expression recognition," *Neuropsychologia*, vol. 150, p. 107700, 2021.

[26] Y. Tian, T. Kanade, and J. F. Cohn, "Facial expression recognition," in *Handbook of face recognition*, ed: Springer, 2011, pp. 487-519.
[27] R. Ibrahim, S. Zeebaree, and K. Jacksi, "Survey on Semantic Similarity Based on Document Clustering," *Adv. sci. technol. eng. syst. j.*, vol. 4, pp. 115-122, 2019.

[28] T. Hussain, K. Muhammad, W. Ding, J. Lloret, S. W. Baik, and V. H. C. de Albuquerque, "A comprehensive survey of multi-view video summarization," *Pattern Recognition*, vol. 109, p. 107567, 2021.

[29] D. A. Zebbari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 100-105.

[30] P. Malviya and P. Tripathi, "DETECTION OF FACIAL EXPRESSIONS USING MACHINE LEARNING IN MATAILB."

[31] D. A. Hasan, B. K. Hussan, S. R. Zeebaree, D. M. Ahmed, O. S. Kareem, and M. A. Sadeeq, "The Impact of Test Case Generation Methods on the Software Performance: A Review," *International Journal of Science and Business*, vol. 5, pp. 33-44, 2021.

[32] Z. M. Khalid and S. R. Zebbari, "Big Data Analysis for Data Visualization: A Review," *International Journal of Science and Business*, vol. 5, pp. 64-75, 2021.

[33] A. R. Hazourli, A. Djeghri, H. Salam, and A. Othmani, "Multi-facial patches aggregation network for facial expression recognition and facial regions contributions to emotion display," *Multimedia Tools and Applications*, pp. 1-24, 2021.

[34] A. B. Sallow, S. R. Zeebaree, R. R. Zebbari, M. R. Mahmood, M. B. Abdulrazzaq, and M. A. Sadeeq, "Vacine Tracker/SMS Reminder System: Design and Implementation."

[35] K. Juneja and C. Rana, "An Extensive Study on Traditional-to-Recent Transformation on Face Recognition System," *Wireless Personal Communications*, pp. 1-54, 2021.

[36] M. B. Abdulrazzaq, M. R. Mahmood, S. R. Zeebaree, M. H. Abdulwahab, R. R. Zebbari, and A. B. Sallow, "An Analytical Appraisal for Supervised Classifiers’ Performance on Facial Expression Recognition Based on Relief-F Feature Selection," in *Journal of Physics: Conference Series*, 2021, p. 012055.

[37] J. Lee, S. Kim, S. Kim, and K. Sohn, "Multi-modal recurrent attention networks for facial expression recognition," *IEEE Transactions on Image Processing*, vol. 29, pp. 6977-6991, 2020.

[38] B. S. Osanaiye, A. R. Ahmad, S. A. Mostafa, M. A. Mohammed, H. Mahdin, R. Subhi, et al., "Network Data Analyser and Support Vector Machine for Network Intrusion Detection of Attack Type."

[39] M. R. Mahmood, M. B. Abdulrazzaq, S. Zeebaree, A. K. Ibrahim, R. R. Zebbari, and H. I. Dino, "Classification techniques’ performance evaluation for facial expression recognition," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, pp. 176-1184, 2021.

[40] D. A. Haider Fawzia, Salama A. Mostafac, Mohd Farhan Md Fudzeeec, Mazin Abed Mahmoood, Subhi R.M. Zeebaree, Dheyaa Ahmed Ibrahimf, "A REVIEW OF AUTOMATED DECISION SUPPORT TECHNIQUES FOR IMPROVING TILLAGE OPERATIONS," *REVISTA AUS*, vol. 26, pp. 219-240, 2019.

[41] Y. Tian, T. Kanade, and J. Cohn, "Handbook of Face Recognition, eds. SZ Li and AK Jain," ed: Springer, London, 2011.

[42] I. A. Khalife, S. R. Zeebaree, M. Ataş, and F. M. Khalifa, "Image steganalysis in frequency domain using co-occurrence matrix and Bpnn," *Science Journal of University of Zakho*, vol. 7, pp. 27-32, 2019.

[43] I. E. Nicolei and M. Ivanovici, "Preparatory Experiments Regarding Human Brain Perception and Reasoning of Image Complexity for Synthetic Color Fractal and Natural Texture Images via EEG," *Applied Sciences*, vol. 11, p. 164, 2021.

[44] H. I. Dino and M. B. Abdulrazzaq, "A Comparison of Four Classification Algorithms for Facial Expression Recognition," *Polytechnical Journal*, vol. 10, pp. 74-80, 2020.
[45] T. T. Q. Le, T.-K. Tran, and M. Rege, "Dynamic image for micro-expression recognition on region-based framework," in 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), 2020, pp. 75-81.

[46] W. Chen, D. Zhang, M. Li, and D.-J. Lee, "STCAM: Spatial-Temporal and Channel Attention Module for Dynamic Facial Expression Recognition," IEEE Transactions on Affective Computing, 2020.

[47] H. Ni and J. Liu, "3D Face Dynamic Expression Synthesis System Based on DFFD," in 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2019, pp. 1125-1128.

[48] M. Xue, A. Mian, X. Duan, and W. Liu, "Learning Interpretable Expression-sensitive Features for 3D Dynamic Facial Expression Recognition," in 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), 2019, pp. 1-7.

[49] L. Deng, Q. Wang, and D. Yuan, "Dynamic Facial Expression Recognition Based on Deep Learning," in 2019 14th International Conference on Computer Science & Education (ICCSE), 2019, pp. 32-37.

[50] S. Ricciardi, D. Iengo, M. Nappi, and D. Vanore, "DYNAMIC FACIAL FEATURES FOR INHERENTLY SAFER FACE RECOGNITION," 2019.

[51] R. Alazrai, K. M. A. Yousef, and M. I. Daoud, "Emotion Recognition Based on Decoupling the Spatial Context from the Temporal Dynamics of Facial Expressions," in 2019 International Symposium on Networks, Computers and Communications (ISNCC), 2019, pp. 1-6.

[52] W. J. Baddar, S. Lee, and Y. M. Ro, "On-the-Fly Facial Expression Prediction using LSTM Encoded Appearance-Suppressed Dynamics," IEEE Transactions on Affective Computing, 2019.

[53] M. Verma, S. K. Vipparthi, G. Singh, and S. Murala, "LEARNet: Dynamic imaging network for micro expression recognition," IEEE Transactions on Image Processing, vol. 29, pp. 1618-1627, 2019.

[54] W. Li, D. Huang, H. Li, and Y. Wang, "Automatic 4d facial expression recognition using dynamic geometrical image network," in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018, pp. 24-30.

[55] T. Vithanawasam and B. Madhusanka, "Dynamic Face and Upper-Body Emotion Recognition for Service Robots," in 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 2018, pp. 428-432.

[56] J. Dong, H. Zheng, and L. Lian, "Dynamic facial expression recognition based on convolutional neural networks with dense connections," in 2018 24th International Conference on Pattern Recognition (ICPR), 2018, pp. 3433-3438.

[57] M. J. Wang, "Face Feature Dynamic Recognition Method Based on Intelligent Image," in 2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), 2018, pp. 57-60.

[58] M. Li and Z. Wang, "Dynamic facial expression recognition based on geometric and texture features," in Ninth International Conference on Graphic and Image Processing (ICGIP 2017), 2018, p. 106150T.

[59] J. Yu and Z. Wang, "A monocular video-based facial expression recognition system by combining static and dynamic knowledge," in 2016 IEEE/ACM 9th International Conference on Utility and Cloud Computing (UCC), 2016, pp. 353-357.

[60] M. Seo and Y.-W. Chen, "Joint subspace learning for reconstruction of 3D facial dynamic expression from single image," in 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2016, pp. 820-824.

[61] M. Liu, S. Shan, R. Wang, and X. Chen, "Learning expressionlets via universal manifold model for dynamic facial expression recognition," IEEE Transactions on Image Processing, vol. 25, pp. 5920-5932, 2016.
[62] H. I. Dino and M. B. Abdulrazzaq, "Facial expression classification based on SVM, KNN and MLP classifiers," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 70-75.