Design and Implementation of Facial Expressions Recognition based on LBP & DRLBP with CNN

Minal Y. Barhate[1], Dr. Manoj Eknath Patil[2]

Research Scholar[1], Research Guide[2]

[1,2]Department of Computer Science & Engineering, Dr.A.P.J.Abdul Kalam University, Indore(M.P)

minaltkolhe@gmail.com[1], mepatil@gmail.com[2]

Article Info
Page Number: 9772 - 9783
Publication Issue:
Vol 71 No. 4 (2022)

Abstract

Facial Expression Recognition (FER), which is the main way to figure out what someone means without them saying it, is an important and promising area of research in computer vision and artificial intelligence. FER is also the main way that nonverbal intentions are figured out. This report gives a clear and concise summary of what's been going on in FER recently. To start, we will divide the current methods for FER into two main groups: those that are considered "traditional" and those that use "deep learning." We suggest a general structure for a classic FER method and look at the technologies that could be used for its different parts. Our goal is to show how they are similar and how they are different. Four different state-of-the-art neural network-based FER approaches are given as examples of deep learning-based techniques, and each one is then looked at in detail. We also give a description of seventeen FER datasets that are often used, as well as a summary of four characteristics of datasets that are linked to FER and could affect the choice of method and the way it is processed. We talk about evaluation methods and metrics, compare the results of different FER approaches using benchmark datasets, Design and implementation of Facial Expressions Recognition Based on Local Binary Pattern (LBP)&Dominant Rotated Local Binary Patterns (DRLBP).

Article History
Article Received: *15 September 2022*
Revised: *25 October 2022*
Accepted: *14 November 2022*
Publication: *21 December 2022*

Keywords:LBP , DRLBP ,Facial Expressions Recognition

Introduction

A key part of nonverbal communication is being able to recognise and understand a person's facial expressions. Help is given to make sure that what is said is clear and easy to understand. People's concentration can be seen in many ways, such as their behaviour, mental state, personality, criminal tendencies, dishonesty, and so on. No matter their gender, culture, colour, or where they live, the vast majority of people are born with the ability to comprehend what other people are feeling by looking at their faces. Automatically recognising and classifying facial expressions is hard to do, which is a shame. Most of the time, researchers will only use well-known emotions like anger, sadness, happiness, and fear. But it is very hard for robots to tell the difference between so many different feelings and moods. Also, robots need to be taught to understand the situation they are in and, more specifically, what a person wants them to do. When we say "machine," we can also be talking about things like computers and robots. Robots can talk to each other in new and creative ways because they are built to have some independence. Because people are different in terms of gender, age, race/ethnicity, and the quality of the photos or videos they use, it is hard to put emotions into categories. It is important to have a system that can understand human facial expressions at the same level as humans. In the last few decades, FER has grown into a new field of study that could be very useful. Face recognition systems are becoming more and more common in the security and healthcare fields. Most of these systems depend on computer vision, artificial intelligence, image processing, and machine learning.

Face detection is the first step in the FER process. It means looking through a video or still image for faces to identify. The photos don't just show the faces of the people in them; they also show their interesting backgrounds. In fact, people are born with the ability to see emotions and other facial traits in pictures, while computers, even with the best training, still have trouble doing this. The main goal of face detection is to tell apart people's faces from the places where they are found (non-faces). Face detection has many uses, such as gesture recognition, video surveillance systems, automated cameras, gender recognition, facial feature recognition, face recognition, tagging, and video conferencing. These are just a few of the many things that could happen. For these systems to work, the first step is to have a way to recognise people's faces as inputs. A full-color image is taken in every direction by a colour image sensor. Because of this, most face recognition algorithms used today work with black-and-white images, and only a small number can handle colour photos. In order to improve their overall level of performance, these systems use the two most common ways to recognise faces, which are called window-based and pixel-based. Pixel-based methods are slow when it comes to separating parts of the body, like the face from the hands.

When using a window-based method, on the other hand, you lose the ability to see faces from different angles. In the field of FER, face detection is one of the most common ways to use model matching techniques. With the window-based method, on the other hand, you can't see more than one face at once. Modern classifiers like artificial neural networks (ANN), Dominant Rotated Local Binary Patterns (DRLBP), support vector machines (SVM), kernel neural networks (KNN), and random forests are used in healthcare, biometrics, handwriting research, and face detection for security applications (RF).

Related work

X. Zhang et al.[1]

The proposed solution, which uses something called a "Generative Adversarial Network" (GAN), has several advantages. First of all, the benefits of the unified model include both face image synthesis and recognising facial emotions. Second, unlike other face image synthesis networks, the model we've come up with doesn't need to use two photos together. Also, the problem of overfitting that we had in our FER study has been fixed because the generated facial images have greatly increased the size of the training set. Third, we don't have to use a set of predetermined expressions. Instead, we can make synthetic face images with any expression we want by swapping parts of their latent identifying qualities. This means we don't have to use a set of expressions that are already set.

Z. Hu et al.[2]

A lightweight multi-scale network with attention is shown with the goal of figuring out how people are feeling by looking at their faces. This is done to fix problems with the traditional convolutional neural network (CNN), such as the use of too many parameters, the use of a single scale feature, and the inefficiency caused by some features that aren't needed. This network can learn important facial traits by combining the lightweight convolutional neural network model Xception with the convolutional block attention module (CBAM). These traits are then combined to improve the receptive field and collect more detailed face feature data. A depth-separable convolution module with a convolution kernel of 3, 5, and 7 is used to find the facial expression picture's features.

Kabir, M. M., et al.[3]

In this study's proposed architecture, which is based on Convolutional Neural Network (CNN) and Long Short Term Memory, emotions like joy, anger, contempt, fear, sadness, surprise, and happiness are grouped together with neutral emotions like fear and surprise (LSTM). This all-in-one plan can be broken down into two separate steps. After a convolutional neural network (CNN) trained on the PNFE dataset pulls out the visual properties, an LSTM is used to limit the link between image sequences and emotions. As an extra quality control measure, confusion matrices are compared to the results of other well-known designs and FER datasets to find out what happened.

A. Kreinis and his colleagues [4] made a web-based video chat platform to help people with ASD, especially with recognising and practising emotions in real time. This app looks at video in real time to figure out how the speaker is feeling and put them into one of eight groups: neutral, surprised, happy, angry, disgusted, scared, or sad. After this sorting is done, a written label will be shown next to the picture of the speaker.

H. Zhang et al.,[5] it is suggested that an identity-expression dual branch network (IE-DBN) be used to figure out how someone is feeling by looking at their face. At first, two different pathways use the same image of a facial expression to learn different things about identity and expression. This picture is used all over. Then, our bilinear module is used to solve the problem of how to combine these two traits. The bilinear aggregate draws attention to both the differences and the similarities between the different groups. This shows how important it is to be yourself in any situation. To make the identity-related traits more uniform, the collaborative training method can be used. Our network can be made to produce identity-related traits that are set by expression, but at the same time, it can be used to stop negative identity traits from happening. Experiments with three of the most important face expression databases, two of the most popular posed face expression databases, and one of the most popular spontaneous face expression databases show that we are better at facial expression recognition.

Y. Xiaet al.[6]

In this paper, we propose a new end-to-end method for making fake facial expressions. We call it the Local and Global Perception Generative Adversarial Network (LGP-GAN). This method is made up of two stages that build on each other. It is meant to pull out and combine the details of the most important parts of the face. Also, this finding gave us the idea to suggest this method. Using both the LGP-global GAN network and a local network, it is possible to generate facial expressions.

D. Poux et al.[7]

In this study, we take advantage of this strength by showing how an auto-encoder with skip connections can be used to reconstruct the hidden facial features in the optical flow domain. This method is meant to be better than what has been tried before. This is the first time that anyone knows of that someone has tried to directly re-create the movement of the face to figure out how it looks. We used the regulated CK+ datasets to test the effectiveness of our method. For each of these tests, we used different levels of occlusions. Based on the results of our tests, we can see that the proposed method helps close the gap between the accuracy of recognition in settings with and without obstructions. Our method is different from the most cutting-edge ones that were used in the field before us. We also suggest a new experimental method that takes both the making of occlusions and the evaluation of reconstruction into account. This will set the stage for future comparisons that can be done again and are fair.

W. Liu et al. [8]

Researchers who used convolutional neural networks to try to identify facial emotions found that their work was wrong and slow because they didn't take into account how important the main links were to each other. In order to solve this problem and meet the recognition requirements, an algorithm model called "series cascade" has been made for educational robots to recognise facial expressions. Using the CK+ and Oulu-CASIA expression recognition databases, tests were run to compare this algorithm's cascade network to the industry standard STM-ExpLet and the FN2EN method.

Proposed methodology

shows a visual representation of the method, which shows how the process works. The three main parts of the technique are preprocessing, feature extraction, and classification. As we go through the methodology, we will talk about each of these in turn.

shows that the first step of the LBP & DRLBP,CNN framework that has been made is to preprocess the information that has already been gathered. This LBP & DRLBP -CNN uses a CNN architecture, specifically one with two hidden convolutional layers and four convolutional layers. After the data has been preprocessed, the next step is to get the features from the data. There are a number of steps involved in extracting features, and at the end of the process, only the features that will help the study the most are kept. At last, classification is used to find the correct label for the image input. Both the forward and backward propagation of errors are used to find the correct label for the image that is being read in.

Pre-processing

Using preprocessing, it is possible to turn an incoherent dataset into a consistent one, which makes the system run much faster. If you change the size of each picture to 150 pixels on each side, they will all look the same. At this point in the process, we also randomly zoom in and out on each image and rotate them between 0 and 180 degrees. The photos have been turned both horizontally and vertically from how they were before.

Architecture of LBP & DRLBP-CNN

For example, a convolutional neural network, or LBP & DRLBP- CNN, is a type of deep neural network that is used quite often. It takes pictures as input and uses weights and biases that can be trained to find differences between the pictures. These weights and biases are also used in layers that you can't see. Most of the time, LBP & DRLBP- CNN is used to analyse data about what can be seen.

Some of LBP & DRLBP- CNNs most important building blocks are the extraction and classification of features. There are several different subsets of feature extraction, such as convolution, padding, nonlinearity, and pooling. Figure 3 shows that the LBP & DRLBP-CNN

classification process is made up of a layer that is fully connected. This is clear from the figure. It has secret layers, and each of those layers has a certain number of neurons. This makes it a structure that is all connected. Every neuron in the layer below it is connected to every neuron in the layer above it through connections. This means that every neuron in every layer is connected to every neuron in the layer above it.

Feature Extraction

The goal of extracting features from an image is to get useful information that can be used to label the image. Using the picture is one way to do this. If we have a picture of a person's face, for example, we can use feature extraction to figure out its eyes, nose, and lips, among other things . We can use the same method if we have a picture of an animal's face. In the sections that follow, we'll talk in more depth about the next steps of feature extraction.

Convolution

In LBP & DRLBP-CNN

A filter or kernel is used to do convolution on an image that comes in. The size of the kernel used in this paper. In order to filter and convolve the image, we have to scan the whole thing . We start scanning from the top left corner of the image and move to the right side. Then we move one bit down and do the same thing from left to right until the

Padding

You can tell the difference between two types of padding : valid padding and identical padding. When using the right padding, the size of the object being padded should be smaller than its original size. When using the same padding, on the other hand, the size of the object being padded should either be bigger than its original size or stay the same. The LBP & DRLBP-CNNthat has been suggested uses the same padding method, which doesn't change the size.

Nonlinearity

Fully Connected Layer

After the convolutional layer of a LBP & DRLBP-CNN

has made the feature map, the next layer can be either a subsampling layer or a pooling layer. The pooling layer, like the convolutional layer, works to make the features that were convolved smaller.

After dimensionality reduction, it takes less time to do the calculations because there are now fewer dimensions to look at. Pooling is a method that can be very helpful when trying to find features in images that don't change when the image is rotated or moved. With the help of pooling, the length of the training session can be cut down, and overtraining can be avoided.

The proportional pooling method is an alternative to the proportional pooling method. The winner was the maximal pooling method, while the average pooling method had to settle for second place. When a photo is sent into the system, it is compared to the kernel or filter. The most meaningful result is then sent back. We take the average of the values in the part of the image being processed by the kernel and send that value back. On both the LBP and the DRLBP-CNN lines, the most pooling was used. This is because maximum pooling only looks at the data that is most important to the area in question, while average pooling looks at all the data and finds the average. This means that the calculation will not take into account any noise in the image.

The study suggested that a method based on traits be used. Since then, a lot of research has been done on 2D face recognition. In the new millennium, academics looked at faces for the first time in three dimensions (3D).

Face recognition methods based on three dimensions have grown in a different way than those based on two dimensions. This is because a three-dimensional face has more depth than a two-dimensional face. Because of this, it is best to talk about the different ways to recognise faces in terms of two-dimensional and three-dimensional space.

Face recognition research that uses a 2D method can be broken down into three main groups: analytical (feature-based, local), global (appearance), and hybrid. Within each of these approaches, there are also other types. When it comes to facial recognition, there are two main ways to think about it: the analytical and the global. Analytical tries to identify by comparing the features of the different parts of the face, while global tries to identify by using information from the whole face. Both ways of thinking are trying to get to the same place. Hybrid approaches might use both local and global methods together to get more accurate information about the situation.

Face recognition with this kernel can be compared to what can be done with other, more thorough methods.

Analytical methods for identifying faces use a number of factors, such as the distance between the calculated feature points and the angles between them, the shape of the facial features, or the variables that hold the regional features made from the image of the face. Photos of people's faces can mostly be put into two groups: those with patterns and those with geometric shapes. These strategies cut down on the amount of data needed to represent a face, which lowers the high cost of processing that comes with face recognition.

Researchers use global-based approaches to face recognition instead of feature-based approaches because it lets them recognise faces without having to extract facial features, which takes a lot of time. The global-based techniques have been used in face recognition because they make the field much more efficient. now thought to be the most reliable and accurate way to use principal component analysis to represent and recognise faces. Turk and Pentland used this method to turn the whole face image into a set of vectors. They then used a sample set to figure out the eigenfaces. In this case, LBP & DRLBP-CNN was able to take the image data and use it to its fullest, which gave them a perfect image of the face. One problem with LBP & DRLBP-CNN was how it handled the same person in different lighting situations.

Using the shape of the data to tell it apart and divide it into test and production settings

First, we need to separate the training phase from the testing phase so that we can better describe the goal. Why do we spend so much time and energy on this project?

Here's how the project should go from here:

First, you should use the data from the training set to make a model that can tell what a photo is based on its name.

It is recommended that the training set have a lot of pictures of the same people, so that the model can learn how to handle the same face in different situations. This will help the model learn how to use the same face in different situations.

Use the model to try to figure out what photos you don't know, and put it through some tests.

People who were in the training set, but not necessarily in the same photos, should also be in the test set. People who were not in the training set should also be in the test set. This will make it possible to check if the algorithm is right.

Results analysis

Before starting to build a model, it is important to define what it means to say that something is right. This is the last step before building the model. Please list the things that make a model look good. If tp is the number of correct predictions and tn is the number of correct guesses, the formula will show how accurate a prediction is. In the table that comes next, you can see what percentage of people got this question right. But when our model figures out how accurate its predictions are, it doesn't take into account the possibility that the costs of false positives and false negatives may be different. Also, the LFW dataset doesn't work well when there are big differences between classes. Not only is this a problem, but it's also a big one.

$$acc = \frac{tp + tn}{n}$$

Most of the time, the performance of future models will be judged by how well they can keep their accuracy and memory. You can figure out how accurate a model is by figuring out the ratio between the actual class x and the class x that was predicted.

$$prec_x = \frac{tp_x}{tp_x + fp_x}$$

The fpx variable keeps track of how many false positives are associated with class x. Accuracy is the number of correctly labelled pictures that have been shown to belong to class x. The results of the calculations showed that the model recall for class x was about what was expected. In this formula, the number of false negative predictions for class x is represented by the variable fnx. One way to figure out how well a classifier is doing at putting things in the right category is to look at its recall. Using this metric, one can figure out how good a classifier is. To find the F1 score, just add the number of correct answers to the total number of questions on the test. But for this notebook, it is more useful to look at the value of each statistic separately.

In the real world, which of these strategies do you think would be the most useful to use first? You need to think about how things will work in real life and what problems could come up. When dealing with situations where a false positive could have bad effects, accuracy is the most important thing. In a method for diagnosing disease, for example, a false positive result could have serious physical and mental effects on the patient. On the other hand, recall may be helpful in situations where the cost of making false negatives is high. If the model can't figure out who a wanted criminal is in surveillance footage because it missed a picture of the person in question, it has a lot of problems that need to be fixed. At the moment, two of the most interesting ways that image recognition could be used are to label news images and to add notes to live videos. False positives could have a big effect on the finances of both services, so they should try not to label things wrongly. To put it another way, they have something at stake in how things turn out. It is very important to protect the public's trust in the system at all costs and to avoid embarrassing the public at any time. It is also important that the public never has to face any kind of shame. Because of this, we take into account the recall for each projected class. However, the accuracy of our model is the most important thing to us.

In studies about facial recognition, photos from the Olivetti data set were used. Face recognition is now seen as a safer technology because a number of safety measures have been put in place.

Analysis of Unsupervised Face Classification with a Convolutional Neural Network for Forensics Using Labeled Images from the Wild

The dataset called "Labeled Faces in the Wild" has 13,233 pictures of the faces of famous people. These pictures came from different places on the Internet. This collection of pictures shows what the data set is (LFW). The set of images was put together to help researchers, developers, and testers of image recognition systems do their jobs a little bit better.

In this notebook, we will try to build an algorithm that can successfully recognise the faces in the dataset after being trained on other faces that are similar to the ones in the dataset. We'll also talk about what this strategy can't do when it comes to working with real-world data.

To sum up what's in the notebook, we can say the following:

The first step is to get any needed libraries and then install them.

Step 2: Reading in Data In the third step, you do some basic data cleaning to make sure that the structure of the data frame is easy to understand.

Data analysis, data representations, and separating the data into training, test, and validation sets are all part of Analyzing Data Freely (EDA).

You can choose the right evaluation measures in one of three ways.

Which metric would tell us the most about this situation?

Taking Users' Opinions into Account in the Model

Based on test data, we look at how well CNN approaches work and explain a wide range of different computer vision (CV) algorithms.

5. A look at the results and a discussion of them

Results and different ways to understand those results will be talked about.

6. Closing remarks and ideas for what to do next

A look at the potential of this method, as well as its current limitations and future growth potential

There are both supervised and unsupervised ways to do things in the field of machine learning. In the context of supervised learning, every data set has at least two main parts: the data that went in and what was expected to come out. "Data" is the values of the sample in the data set, while "input" is the class (for classification) or the value that was intended (for regression). A data collection should only include the data part to make it easier for people to learn on their own. Non-supervised learning is made up of two main parts: grouping data together and changing it. In this investigation, unsupervised learning will be the best way to change the data. With unsupervised transformation techniques, both machines and people can get a better understanding of the data. Most of the time, unsupervised transformations are used to make data smaller. When data is compressed, the size of the data shrinks. A method called P LBP & DRLBP- CNN that reduces the size of data representations. Before the selection process, the data is broken up into smaller pieces that are easier to handle. This leads to a smaller set of information that is easier to understand. After testing the George Bush model several times, we were able to get an accuracy of 81% and a recall of 98%. To restate, 95% of the photos we thought were of George Bush have been confirmed to be of him, and 81% of the photos have been correctly identified as being of George Bush. What does this mean when it's put into action. we care a lot about accuracy. A value of about 81% means that only 19% of the pictures we think show George Bush actually show someone else.

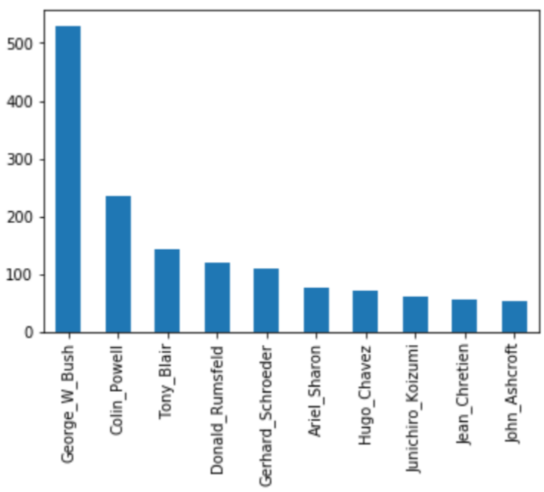

Figure 1: Results analysis

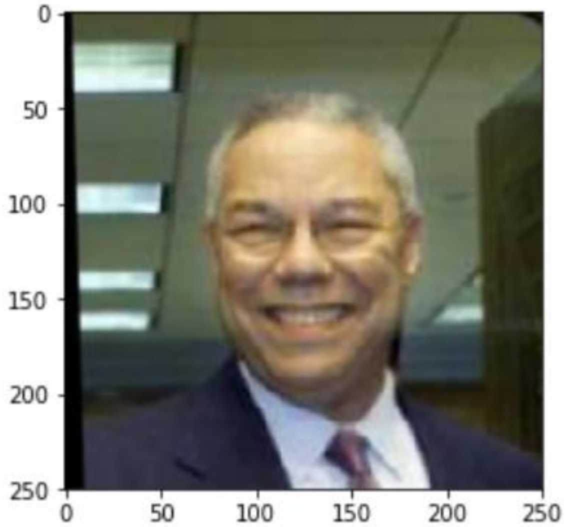

Figure 2: Detection of face through image

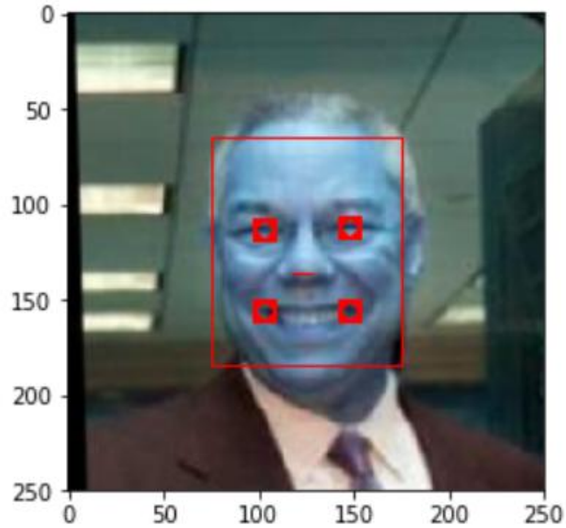

Figure 3: Detection of face through image

Does it not bother you if I do that? There is a very small chance that the answer will be "yes." In contrast to news labelling, where a person is on-site to check the results before they reach the public, live video labelling is likely to make too many mistakes to be useful and could lead to embarrassing mistakes. Before putting this model into production, more checks should be done to make sure it is safe. This is especially important if the rate of recalls is 99.5% or higher, or if there have been more than 51,000. this model's high recall rate may give us confidence in its use in a different situation where false negatives are expensive. How does the idea of precision work when there are six people in a group? There isn't much difference between classes in terms of accuracy and memory, but there is a lot of difference between them. Compared to the size of the sample as a whole, this is a small number. In fact, the model's performance on photos of Tony Blair was pretty bad (0.41*) in the last run before this notebook was posted. This is because a lot of the test pictures of Tony Blair were mistakenly called pictures of George W. Bush. Because of this, people from George W. Bush's generation also had a hard time remembering things. This is not just a funny old joke; it shows the limits of a shallow network trained on small data sets when it is asked to tell the difference between people who have similar traits and often appear in photos that are similar to each other. One way to solve this problem could be to make the model more complicated.

Conclusion

In this paper, we look at a number of different FER methods in depth and compare them. We basically put these techniques into two groups: (1) the traditional ML-based methods and (2) the DL-based methods. In a typical machine learning process, there are three steps: finding faces, pulling out features from those faces, and then using those features to figure out how someone is feeling. In regular machine learning for FER, classification techniques like random forest, adaboost, key-value network (KNN), and support vector machine are used (SVM). Compared to DL-based FER techniques, there is much less reliance on models that are based on how the face works. They also speed up the learning process by getting rid of the need for any intermediate steps in the making of the input photographs. On the other hand, the training and testing parts of these systems take a lot more time. Even if a hybrid design works better, it is still hard to solve the problem of micro-expressions because the face can move in other ways without the person's intention.

Reference

[1]. X. Zhang, F. Zhang and C. Xu, "Joint Expression Synthesis and Representation Learning for Facial Expression Recognition," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1681-1695, March 2022, doi: 10.1109/TCSVT.2021.3056098.

[2]. Z. Hu and C. Yan, "Lightweight Multi-Scale Network with Attention for Facial Expression Recognition," 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), 2021, pp. 695-698, doi: 10.1109/AEMCSE51986.2021.00143.

[3]. M. M. Kabir, T. A. Anik, M. S. Abid, M. F. Mridha and M. A. Hamid, "Facial Expression Recognition Using CNN-LSTM Approach," 2021 International Conference on Science & Contemporary Technologies (ICSCT), 2021, pp. 1-6, doi: 10.1109/ICSCT53883.2021.9642571.

[4]. A. Kreinis, T. Damri, T. Leon, M. Litvak and I. Rabaev, "Telemoji: A video chat with automated recognition of facial expressions," 2021 International Conference on Visual Communications and Image Processing (VCIP), 2021, pp. 1-1, doi: 10.1109/VCIP53242.2021.9675330.

[5]. H. Zhang, W. Su, J. Yu and Z. Wang, "Identity–Expression Dual Branch Network for Facial Expression Recognition," in IEEE Transactions on Cognitive and Developmental Systems, vol. 13, no. 4, pp. 898-911, Dec. 2021, doi: 10.1109/TCDS.2020.3034807.

[6]. Y. Xia, W. Zheng, Y. Wang, H. Yu, J. Dong and F. -Y. Wang, "Local and Global Perception Generative Adversarial Network for Facial Expression Synthesis," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1443-1452, March 2022, doi: 10.1109/TCSVT.2021.3074032.

[7]. D. Poux, B. Allaert, N. Ihaddadene, I. M. Bilasco, C. Djeraba and M. Bennamoun, "Dynamic Facial Expression Recognition Under Partial Occlusion With Optical Flow Reconstruction," in IEEE Transactions on Image Processing, vol. 31, pp. 446-457, 2022, doi: 10.1109/TIP.2021.3129120.

[8]. W. Liu and J. Fang, "Facial Expression Recognition Method Based on Cascade Convolution Neural Network," 2021 International Wireless Communications and Mobile Computing (IWCMC), 2021, pp. 1012-1015, doi: 10.1109/IWCMC51323.2021.9498621.
[9]. C. Wang, Z. Wang and D. Cui, "Facial Expression Recognition with Attention Mechanism," 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2021, pp. 1-6, doi: 10.1109/CISP-BMEI53629.2021.9624355.
[10]. Z. Iksanova, "A new method of developing emotional intelligence in preschool children through interactive digital facial expression recognition technology," 2022 Fourth International Conference Neurotechnologies and Neurointerfaces (CNN), 2022, pp. 51-54, doi: 10.1109/CNN56452.2022.9912512.
[11]. F. M. Sadikoğlu and M. Idle Mohamed, "Facial Expression Recognition Using CNN," 2022 International Conference on Artificial Intelligence in Everything (AIE), 2022, pp. 95-99, doi: 10.1109/AIE57029.2022.00025.
[12]. N. I. Abbasi, S. Song and H. Gunes, "Statistical, Spectral and Graph Representations for Video-Based Facial Expression Recognition in Children," ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022, pp. 1725-1729, doi: 10.1109/ICASSP43922.2022.9747102.
[13]. W. Huang, S. Zhang, P. Zhang, Y. Zha, Y. Fang and Y. Zhang, "Identity-Aware Facial Expression Recognition Via Deep Metric Learning Based on Synthesized Images," in IEEE Transactions on Multimedia, vol. 24, pp. 3327-3339, 2022, doi: 10.1109/TMM.2021.3096068.
[14]. Y. Xia, H. Yu, X. Wang, M. Jian and F. -Y. Wang, "Relation-Aware Facial Expression Recognition," in IEEE Transactions on Cognitive and Developmental Systems, vol. 14, no. 3, pp. 1143-1154, Sept. 2022, doi: 10.1109/TCDS.2021.3100131.
[15]. H. Cao and C. Qi, "Facial Expression Study Based on 3D Facial Emotion Recognition," 2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS), 2021, pp. 375-381, doi: 10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00067.
[16]. J. Cai et al., "Identity-Free Facial Expression Recognition Using Conditional Generative Adversarial Network," 2021 IEEE International Conference on Image Processing (ICIP), 2021, pp. 1344-1348, doi: 10.1109/ICIP42928.2021.9506593.
[17]. Ram Kumar, Manoj Eknath Patil ,” Improved the Image Enhancement Using Filtering and Wavelet Transformation Methodologies”, Turkish Journal of Computer and Mathematics Education ,Vol.13 No.3(2022), 987-993.
[18]. Vibhor Mahajan, Ashutosh Dwivedi, SairajKulkarni,Md Abdullah Ali, Ram Kumar Solanki,” Face Mask Detection Using Machine Learning”, International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:05/May-2022