ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse100942020.pdf
https://doi.org/10.30534/ijatcse/2020/100942020

Emotion Recognition Using Artificial Intelligence And Neural Network

Elena A. Chelnokova [1], Olga I. Vaganova[2], Zhanna V. Smirnova[3], Maxim M. Kutepov[4], Ekaterina V.Vovk[5]
[1] Minin Nizhny Novgorod State Pedagogical University, Nizhny Novgorod, Russia
[2]Minin Nizhny Novgorod State Pedagogical University, Nizhny Novgorod, Russia
[3] Minin Nizhny Novgorod State Pedagogical University, Nizhny Novgorod, Russia
[4]Minin Nizhny Novgorod State Pedagogical University, Nizhny Novgorod, Russia
[5]V.I. Vernadsky Crimean Federal University, Simferopol, Russia

ABSTRACT

A scientific study is devoted to the problems of using artificial intelligence to recognize emotions. The basis for emotion recognition using artificial intelligence is considered on the example of using CNN. The main algorithms for classifying emojis are analyzed, and the CNN architecture is also built. Recognition of human expressions and emotions has attracted the attention of researchers as the ability to recognize human expressions helps in the world. There are many ways to test a person's recognition of expression, ranging from facial expressions, body posture, voice tone, and the like. Facial emotion recognition (FER) is a thriving research field that likes a lot of progress automatic translation systems, machine interaction between people are taking place in industries. In contrast to paper focus, review and view various features of face deletion, emotional databases, classifier algorithms, and the like.

Keywords:, Artificial intelligence, neural network, Python, FER, CNN.

1. INTRODUCTION

Facial recognition is a technology that deals with methods and techniques for detecting emotions from facial expressions. Various technological developments in the field of machine learning and artificial intelligence, makes it easier to recognize emotions. It is expected that expressions can be the next means of communicating with computers [4]. The need for automatic recognition of emotions from facial expressions increases extremely quickly. Research work is mainly focused on identifying emotions from pictures or acoustic information. Emotions can be detected by facial expressions, using language cues, and so on. Recognition of emotions is becoming increasingly important in research, which is of primary importance for solving many problems [1-3]. Primary requirement recognizing emotions on behalf of an expression is a complex task for artificial intelligence, where images are given as input data for systems [5].

Some of the classification algorithms like K-Nearest Neighbor, Random Forest are applied in classifying emotions. Deep RNN, like LSTM, bidirectional LSTM is modeled for audio-visual functions and is also used in emotion recognition. Facial emotion recognition is checked and analyzed in all areas of research. Emotion is detected by the face of the image using filter banks and CNN, which gives a high accuracy with which you can draw a high conclusion that CNN can also be used to detect emotions. Emotion recognition can also be performed using a spectrogram image with deep convolutional networks [6]. All of the methods listed above used some of the traditional selection methods in MFCC, wave parameters such as pitch. This article examines another database that is used to recognize a person's emotions, features selected from facial expressions, classifiers that are used to classify different classes of emotions. As the volume of the data set is taken and the bottleneck method, long-term short-term memory (LSTM) is used to recognize emotions on the face. Although language emotion recognition has been done and the desired results shown, real-time personality studies of emotions are still ongoing [7]. Real-time facial emotion recognition is performed by classifying RGB images using a transfer learning technique in which the knowledge gained is from solving a single problem, and this is implemented for the entire other problem.

2. MATERIALS AND METHODS

Various types of images and emotions were examined for facial expressions using various classifiers such as KNN, HMM, GMM, SVM [1-9]. This article explains important functions such as machine learning of reference vectors, locally invariant learning feature, and pronounced discrimination in analyzing features of emotion recognition. Various important features have been studied and trained to detect emotions using CNN, in which a dataset is obtained from various emotional databases such as SAVEE, Emo-DB, DES, MES. This article describes facial emotion recognition using neural networks. Although the probabilistic method of detecting emotions is generally accepted, it is used to recognize emotional changes. According to research in the field of recognized emotions, it was found that it is best to use CNN networks for artificial intelligence [14].

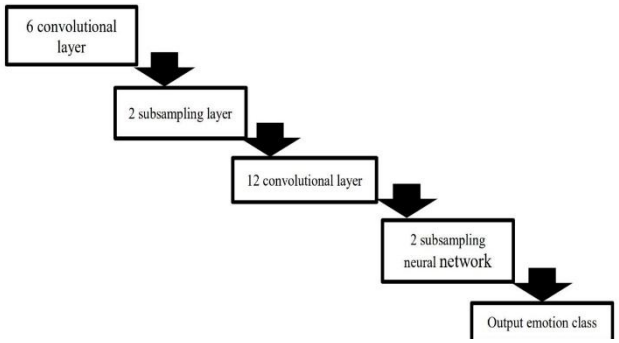

Figure 1: CNN Architecture

The architecture of the proposed CNN is shown in Fig. 1. It has two convolutional layers and two subsample layers. The first convolutional layer used six masks, or the so-called C1 layer. The next layer is the pid-sampling layer, which has two layers (s1). The second convolutional ball(or S2) has 12 masks. The last neural network has two layers. The latter is a fully connected layer [10-12].

Convolution neural network is a neural network consisting of convolution layers that performs a computational operation by performing a convolution algorithm. Convolution is a mathematical operation on two functions to obtain a third function. Note that images are not represented as pixels, but as numbers representing pixel values [28]. In terms of what the computer sees, it will just be a matrix of numbers. At this point, the number convolution operation is performed. We use both fully linked layers and convolutional layers. In a fully connected layer, each node is connected to every other neuron. They are layers that are used in standard feed neural networks. Unlike fully connected layers, convolutional layers are not connected by each neuron. Connections are made in localized regions [13]. The sliding window moves across the image. The size of this window is known as the kernel or filter. They help you recognize patterns in the data. For each filter there are two basic properties that you should consider is the padding and step. A straight line represents the stage of the collapse operation, i.e. the number of points through which the window moves. Overlay-adding zero points to increase the size of the image. Zero pixels here refer to pixels with the value 0. If we have a 5x5 image and a window with a 3x3 filter, step 1 and no gaskets, the output of the convolutional layer will not be a 3x3 image. This condensation of mapping functions is known as merging. In this case, "maximum join"is used. Here is the maximum value taken by Sud [14].

By measuring a data item using data measurement tools for each class, a summary was obtained. The accuracy level for an evil emotion is 87.73%; contempt 90.95%; disgust 93.46%; fear 91.75%; happy 96.38%; sad 91.15%; surprise 98.09%; and neutral 92.96%. The average accuracy level for all testing is 92.81%

The experiment scenario gives us an outstanding result about the system performance with an average Accuracy of 92.81%. The lowest accuracy level is the anger class, 87.73%, and the highest - surprise. Each class has misclassification results that indicate that the system requires further improvement. For the next study, we should consider changing the entire architecture to improve the result.

The Convolutional Neural Network architecture for facial expression recognition was proposed. Using the CK + database, there wasa I using different sizes of training data, and as a result, the standard deviation decreases with increasing amount of training data. From the experiment, it can be concluded that the standard deviation decreases with the growth of the adaptive control system. In addition, the system's performance reaches 92% [15-19]

The architecture of the MSA consists of many layers. There are two main types of layers: convolutional and subsampling. Convolutional and subsampling layers alternate with each other. But recently, there is a tendency to reduce the number of sub-samples. Therefore, there are often fewer layers than convolutional or not at all.

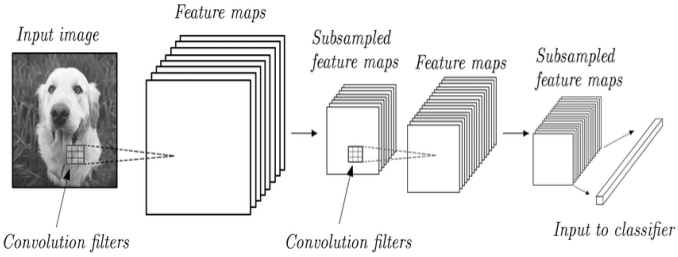

Figure 2: General view of the portal neural network

Figure 2 shows us the general view of the portal neural network.In each occlusal layer, there is a set of several filters (convolution cores) that have a small receptive field, but occupy the entire depth of the input volume, and the neurons of one filter have the same weights, which are applied to all local areas of the previous layer (thereby performing convolution). The image of the previous layer is scanned by a small window and multiplied by a set of weights, and the result is stored on the corresponding neuron of the current layer. Thus, a set of planes is a feature map, where each plane finds "its own" image features anywhere on the previous layer [20].

Next to the bottom layer, the top layer reduces the scale of data planes to make it easier to highlight the main features in subsequent layers. The most effective is the "maximum". The last part of the network is a feature classifier, usually represented by a perceptron or a support vector machine. The number of neurons in the classifier layer is determined by the number of classes that the input image should belong to.

Two of the most advanced convolutional neural network algorithms, SSD and YOLO, were also applied, which have high detection speed and accuracy.

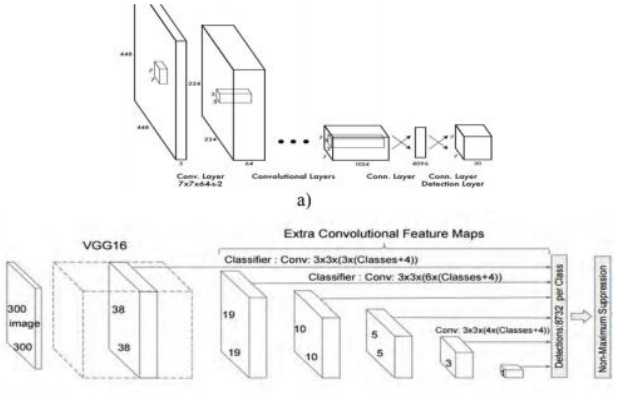

Figure 3: Architecture of YOLO (a) and SSD (b)

YOLO is very fast (fig 3). During testing on one of the most powerful GPUs (Nvidia Titan X GPU), the base network processes 45 frames per second. YOLO evaluates all images when it makes a forecast, unlike methods based on a sliding window or regional offers. The system divides images using a grid, and for each cell provides several limiting frames and the probability that they have an object of a certain class [21-25]. SSD is the next generation of object recognition software. it also looks at the entire image, but it looks at grids of different scales [26]. By adding a number of improvements, we were able to significantly increase the accuracy compared to the previous version, without losing the recognition speed. By applying these filters to different feature maps obtained from different layers at later stages, networks allow detection at different scales [27]. This solves a significant problem in the network and allows you to achieve high accuracy in detecting objects of various shapes and sizes in images with relatively low resolution.

3. CONCLUSION

As a result of the literature analysis, it was found that convolutional neural networks are the most effective with intellectual methods of recognizing objects from the video stream. Among them, the most advanced architecture is YOLO and SSD. Based on the SSD architecture, a system for recognizing objects in the image from the system's camera has been developed.

REFERENCES

1. Andrienko, O.A. (2019). **On the need to apply gaming training technologies**. *Balkan Scientific Review*, 2 (4), 5-8.
2. Klinkov, G.T. (2018). **The specificity of manifestation of pedagogical communication as a special construct**. *Scientific Vector of the Balkans*, 1, 51-52.
3. Vaskovskaya, G.A. (2018). **Features of implementation of pedagogical technologies of profile training**. *Balkan Scientific Review*, 1, 76-79.
4. Denisova, O.P., Smirnova, Z.V., Vaganova, O.I., Gladkova, M.N., Tsarapkina, J.M. (2019). **Development of partnerships in additional education**. *International Journal of Innovative Technology and Exploring Engineering*, 8 (8), 639-643.
5. Boyko, K.Yu. Skripkina, A.V., Smartly K.Yu. (2019). **The use of syllabus funds in the technology of time management of self-educational activities of bachelors,***Multilevel system of continuous professional education in the socio-cultural sphere: problems of continuity and integration*, 73-78.
6. Smirnova, Z.V., Vaganova, O.I., Loshkareva, D.A., Konyaeva, E.A., Gladkova, M.N. (2019) **Practice-oriented approach implementation in vocational education,***IOP Conference Series: Materials Science and Engineering*, 1, 483.
7. Cirdan, A.P.(2019). **Innovative technologies of professional training of future economists in the system of continuous education**. *Humanitarian Balkan Research*, 2(4), 27-30.
8. Filchenkova, I.F. (2019). **Educational management of innovative activity of teachers as an object of pedagogical research**. *Vestnik Mininskogo universiteta*(*Vestnik of Minin University)*, 2019. 7 (4), 3. (in Russ.).
9. Myalkina, E.V. (2019). **Diagnostics of the education quality in the higher educational institution**. *Vestnik Mininskogo universiteta (Vestnik of Minin University)*, 7, (3), 4. (in Russ.)
10. Oros, I.I. (2018) **The role of international connections in the development of the adult education system**. *Humanitarian Balkan Research*, 1, 57-59.
11. Osadchenko, I.I. (2019). **Key concepts of situational training technology in preparing future teachers**. *Scientific Vector of the Balkans*, 1 (3),46-49.
12. Pichugina, G.A., Bondarchuk, A.I. (2019a). **Structure of the training case in the organization of the educational process**. *Humanitarian Balkan Research*, 2(4), 5-7.
13. Pichugina G. A., Zhilyakova D. A. (2019b) **Structuring the organization of the process of creativity**. *Scientific Vector of the Balkans*, 3, 3 (5), 55-58.
14. Pliushch, V.M. (2018). **Independent work of students as a factor of improving education quality**. *Balkan Scientific Review*, 1, 69-71.
15. Aleksieienko-Lemovska, L.V. (2019). **The activity approach as a basis for preschool teachers' methodological activities**, *Humanitarian Balkan Research*, 3, 4(6), 10-14. https://doi.org/10.34671/SCH.HBR.2019.0304.0001
16. Chirva, A.N., Chirva, O.G. (2018). **Contents and method of professionally oriented training of informatic disciplines of future teachers of technologie**s. *Scientific Vector of the Balkans*, 1, 27-31.
17. Kamenez, N.V., Smirnova, Zh. V., Vaganova, O. I., Bystrova, N.V., Tsarapkina, J.M., (2019). **Development of Instructing Techniques in Professional Training**, *International Journal of Mechanical Engineering and Technology*, 10(02), 899–907.
18. Sharonin, P.N., Kozlova, E.V. (2017). **The role of instant messengers in the modern media space**. *Media economics of the 21st century*, 2, 12-15
19. Ilyashenko, L. K., Smirnova, Zh. V., Vaganova, O. I. Chelnokova, E. A., Kaznacheeva, S.N. (2019). **Methods of Conducting Practical Training on the Subject** "Power Sources for Welding", *International Journal of Mechanical Engineering and Technology*, 10 (02), 908–917.
20. Andriushchenko, T. K., (2018). **Personal aspects of pedagogue's innovative culture**, *Scientific Vector of the Balkans*, 1, 13-16.
21. Bakharev, N. P., (2019). **Creativity - a prerequisite for the formation of professional competences in specialists of technical direction of training**, *Scientific Vector of the Balkans*, 3, 4 (6), 17-21.
22. Pisarenko, D. A. (2019). **Evaluation of extracurricular activities of university students with a competency-based approach**, *Scientific Vector of the Balkans*, 3, 3 (5), 37-40.

23. Grigoriev S. G., Shabunina V. A., Tsarapkina Ju. M., Dunaeva N. V. **Electronic library system as a means of self-development of students of digital generation Z** (on the example of studying the course "Basics of the counselor activity") - Scientific and technical libraries. 2019. No. 7. Pp. 78-99. 29.
24. Andrienko, O.A. (2019). **Network educational technologies and their use when working with students**. *Humanitarian Balkan Research 2019,* 1(3), 5-7.
25. Koshechko, N.V. (2018). **Innovations from educational discipline "Pedagogical conflictology" in professional preparation of students**. *Scientific Vector of the Balkans*, 1, 59-63.
26. Halatsyn, K.A., Feshchuk, A.M. (2019).**Diagnosing motivational-and-valuable component of the communicative culture of students in higher technical educational institutions**, *Balkan Scientific Review*, 3, 2 (4), 17-20.
27. Chokkadi, S., Sannidhan, M. S., Sudeepa, K. B., & Bhandary, A. (2019). **A study on various state of the art of the art face recognition system using deep learning techniques**. International Journal of Advanced Trends in Computer Science and Engineering, 8(4), 1590–1600. https://doi.org/10.30534/ijatcse/2019/84842019
28. Chramate, I., & Belmir, F. (2019). **Consumption energy saving of an energy audit within industry in Morocco.** International Journal of Advanced Trends in Computer Science and Engineering, 8(1.5 Special Issue), 136–140. https://doi.org/10.30534/ijatcse/2019/2781.52019
29. Christ, M. J., Tri, R. N. P., Chandra, W., & Mauritsius, T. (2019). **Lending club default prediction using Naïve Bayes and decision tree**. International Journal of Advanced Trends in Computer Science and Engineering, 8(5), 2528–2534. https://doi.org/10.30534/ijatcse/2019/99852019