Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2021.020407
ISSN 2790-2692 Vol. 2, Issue 4: 49-55

Image Emotion Recognition Supporting Fuzzy Neural Network

Rui Feng*

Wuhan Institute of Shipbuilding Technology, Wuhan, China

**corresponding author*

Keywords: Fuzzy Neural Network, Image Emotion Recognition, Selective Joint Fine-Tuning Strategy, Emotion Coefficient

Abstract: Among the information that humans obtain daily, humans recognize things and transform the world through visual senses. The advantage of visual information makes its memory more durable than other information carriers, and the information contained in the image itself is more intuitive and vivid, and hidden behind the image features is rich human emotional semantic information. Therefore, this paper studies and analyzes image emotion recognition (IER) based on fuzzy neural network(FNN). This paper firstly introduces the two basic concepts of FNN and image emotion feature, and then studies emotion recognition methods such as the overall framework of IER, selective joint fine-tuning strategy and multi-feature extraction, and finally the emotion recognition method. Coefficient and selective joint fine-tuning strategies are analyzed and conclusions are drawn.

1. Introduction

With the wide application of modern social media, many pictures are uploaded to the network every day, and people use pictures to share life and express emotions through pictures [1]. However, IER is a challenging task because many subjective factors are added to the recognition process, which makes its automatic recognition process more complicated [2]. IER has achieved remarkable research results in the fields of robot emotion analysis, network data understanding and public opinion analysis [3]. Especially when the FNN is applied to the IER task, it becomes possible to automatically recognize the high-level semantic information of the image [4, 5].

In recent years, many experts and scholars have conducted in-depth research on FNN and IER, and have achieved good research results. For example, experts and scholars such as Tehmina K proposed a technology based on domain adaptation to identify the emotions of images containing facial, non-face and non-human components. This technology detects the visual features of images based on emotion classification. The IER system has good performance, and the accuracy of IER is high [6]. Research scholars such as Manisha S proposed a lightweight multi-stream deep network, which performs image sentiment analysis by connecting the networks, and the experimental results demonstrate the effectiveness of additional contextual information in producing comparable performance to state-of-the-art sentiment models, But it has fewer parameters, which improves its practicality [7]. It is found that FNN can promote the development of IER technology by studying the data related to IER.

This paper studies and analyzes IER based on FNN. The structure of this paper can be roughly divided into three parts: The first part is a brief description of the relevant overview of FNN and IER. This part includes The description of FNN and image emotion features; the second part is the introduction and analysis of IER methods, which includes the design of the overall framework of image recognition, the introduction of selective joint fine-tuning strategy and the analysis of multi-feature extraction; the third part is the introduction and analysis of IER methods. Part of it is to analyze the research results, mainly the analysis of sentiment coefficient and selective joint fine-tuning strategy.

2. Related Overview

2.1. Fuzzy Neural Network

FNN are generally divided into two research types: the first is to form a FNN by fuzzifying its model structure and learning algorithm in a standard neural network; the second is to introduce a fuzzy logic system into the neural network. In the network structure, it is equivalent to adding some hierarchical structures to deal with fuzzy information. FNN can self-adjust and have strong adaptive ability for parameter learning and training [8, 9].

The structure of the FNN model is a feedforward multi-layer network structure, which is very similar to the structure level of the feedforward neural network model. A manifestation of sexual transmission [10]. Figure 1 shows the general principle flow chart of the FNN.

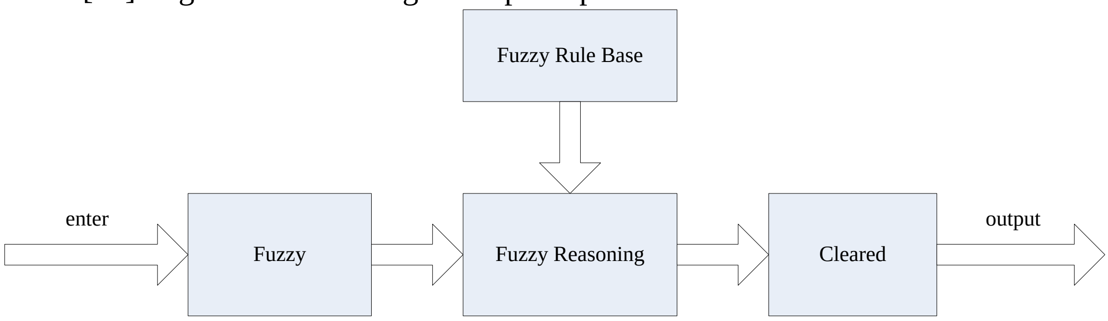

Figure 1. FNN principle flow chart

2.2. Image Emotion Features

For the task of visual sentiment analysis, it has been developed along three different dimensions to extract the emotional features required in model training for IER [11]. Among them, the image features used for visual sentiment analysis are mainly divided into the following three types:

Low-level features: This feature mainly uses traditional handcrafted features to solve the problem, which are related to the color value of image pixels to some extent, involving features such as color histogram, HOG, GIST, etc. [12]. In the context of visual sentiment analysis, some studies have extracted more slightly advanced features for sentiment prediction, such as color, texture, shape, composition, faces, and skin, inspired by research in psychology and art. Many studies have shown that these low-level features can be used to describe image emotion [13].

Intermediate-level features: Since hand-crafted features cannot solve the emotional gap problem well, intermediate-level features emerge as the times require and are developed from low-level visual features. This set of features brings more semantic information, related to objects or scenes, and thus has stronger interpretability and emotional relevance [14, 15].

Advanced Features: Due to the continuous development of deep learning techniques, researchers have devoted themselves to the study of deep architectures. High-level features describe semantic concepts in images and are mainly obtained by fine-tuning pre-trained neural networks to aid sentiment classification [16].

3. Methods of Image Emotion Recognition

3.1. Overall Frame Design

The IER system adopts a modular system, and the algorithm used in this paper is integrated into the system in modules, which is convenient for future additions, deletions and modifications. The general framework of the IER system is shown in Figure 2.

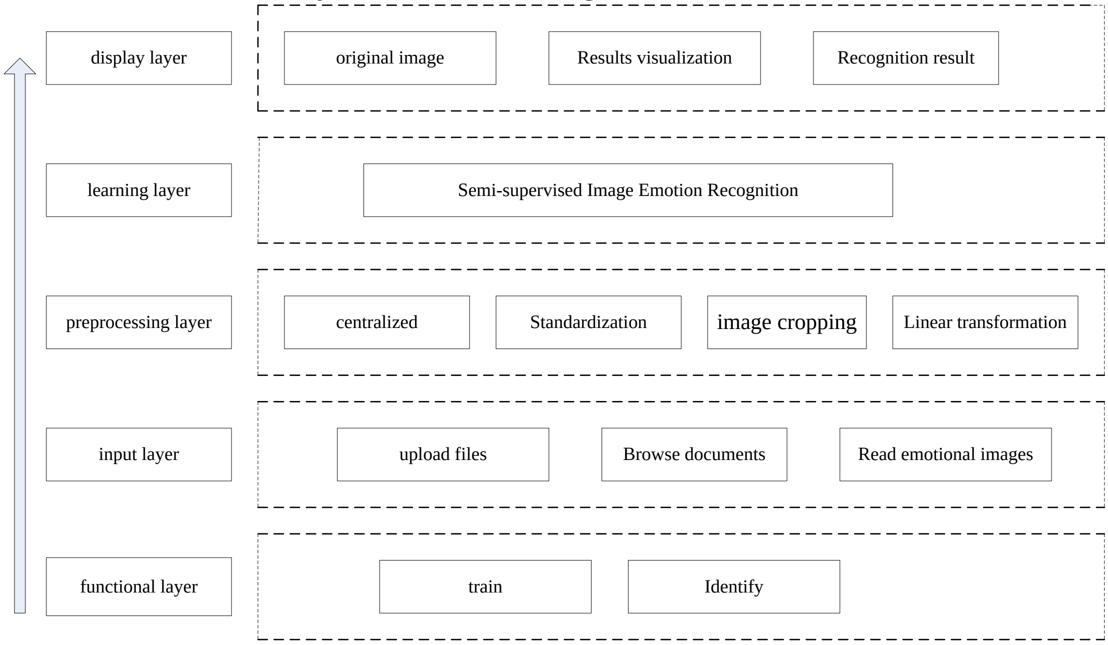

Figure 2. Overall framework diagram

In the overall framework of emotion recognition, the first layer is the functional layer, and the system will be set from two aspects of sample training and recognition; the second layer is the input layer, which directly selects samples from the existing image emotion data set as the system The system reads the relevant information of the samples input into the system at the input layer; the third layer is the preprocessing layer, which preprocesses the data to make the model have stronger generalization ability, which involves: data center The fourth layer is the learning layer, which uses the active learning method and the visual attention mechanism learning method to train the IER model for the preprocessed sample pictures, and extracts the salient emotional features. Emotion recognition; the last layer is the display layer, which is used to display the IER results and give the visual saliency localization results [17].

3.2. Selective Joint Fine-Tuning Strategy

In the target domain learning task of IER classification, the number of training samples required for each emotion category is different [18]. In order to improve the accuracy of sentiment classification samples, the information entropy of the ith training sample in the target domain after the nth iteration is first calculated to measure the uncertainty of its sentiment classification category, and its expression is as follows:

$$K_i^n = -\sum_{a=1}^{a} P_{i,a}^n \log\left(P_{i,a}^n\right) \tag{1}$$

where a is the number of classes and P is the probability that the i-th training sample belongs to class a after the nth iteration. Training samples with high score uncertainty are considered hard to train, and in the next iteration, we will increase the number of source domain training samples that are nearest neighbors of these sample styles and continue to fine-tune the model trained in the current iteration. For a training sample x in the classification target domain, the number of source domain training samples that are closest to its style in the next iteration is defined as follows:

$$F_i^{n+1} = \begin{cases} F_i^n + b_0, \hat{z}_i^t \neq z_i^t \\ F_i^n + b_1, \hat{z}_i^t = z_i^t and K_i^n \geq c \\ F_i^n, \hat{z}_i^t = z_i^t and K_i^n < c \end{cases} \tag{2}$$

where b0, b1, and c are all constants, $\hat{z}$ is the predicted label of x, and F_i^n is the number of training samples in the source domain that are nearest neighbors of style x at the nth iteration.

3.3. Multi-Feature Extraction

In order to verify the effectiveness of the multi-feature description algorithm, this paper verifies the correct rate of IER. The whole image and multi-feature description algorithms are used to compare the accuracy of emotion recognition, and the relevant results are shown in Table 1.

Table 1. IER accuracy result table

Data Set	Recognition Accuracy Using the Entire Image(%)	Recognition Accuracy Using Multi-Feature Description Algorithm(%)
Dataset1	83.41	96.54
Dataset2	78.93	94.71
Dataset3	81.23	98.78

It can be seen from the results in Table 1 that the method of using the multi-feature description algorithm to extract the multi-level features of the image to describe the image emotion can effectively improve the accuracy of IER. The low-level basic features and high-level semantic information jointly describe the image sentiment, which greatly enriches the types and quantity of features. The interaction of the two features can better coordinate the image sentiment analysis. The multi-feature description algorithm proposed in this paper is flexible and can be adjusted according to the characteristics of the task. At the same time, such a multi-feature description structure is not limited to the use of FNNs. The use of artificially designed features can still be used for feature fusion, and then research that meets the research task.

4. Analysis of Research Results

4.1. Sentiment Coefficient Analysis

Using different sentiment coefficients for different images is used to analyze whether the image type will affect the emotion recognition. It can be analyzed from Figure 3 that the emotional coefficients of different categories are quite different. Among the emotional coefficients of human pictures, except the Amusement emotional coefficient is a positive value, other emotional coefficients are negative. Among the animal pictures, Excitement has the most obvious effect in positive emotion and is a positive value, Anger has the most obvious effect in negative emotion and is a negative value. In the emotional coefficients of other types of images, we can see that positive emotions have almost no effect, and they are all small values. Anger in negative emotions has a significant positive effect, and sadness has produced obvious effects as we expected. negative effect. It can be found that some of the emotional effects are not as we expected, positive emotions completely have positive effects, and negative emotions have completely negative effects. This illustrates that different image types can affect emotion recognition.

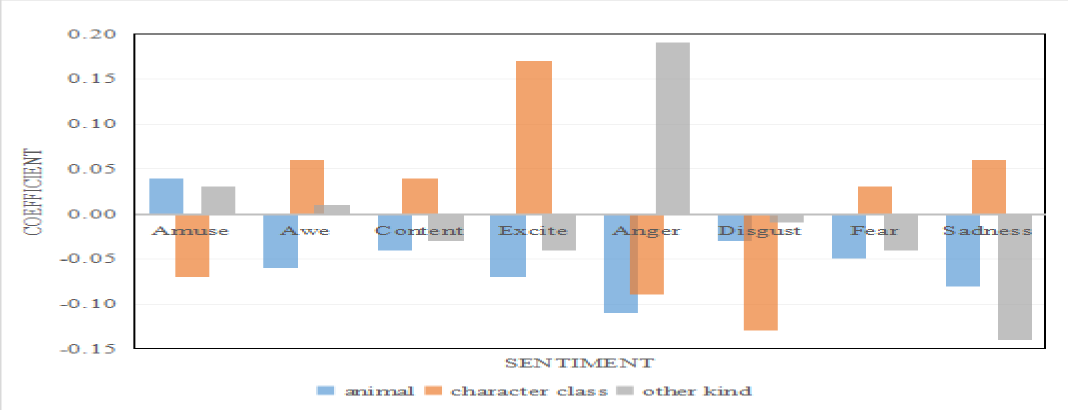

Figure 3. Situation diagram of emotional coefficients of different images

4.2. Analysis of Selective Joint Fine-Tuning Strategy

This paper compares and analyzes the models of IER using different learning methods for emotion recognition accuracy and the models of selective joint fine-tuning methods. The four models are the two-layer transfer learning strategy model TTL, the statistical machine learning model NLMC, Using the full training sample set of the source domain for joint fine-tuning of the model JF and the selective joint fine-tuning training method model SJF, the IER accuracy results of the four models are shown in Table 2.

Table 2. Image recognition accuracy results of different models (%)

Method	Dataset1	Dataset2	Dataset3
NLMC	80.51	89.12	90.69
TTL	84.25	90.78	92.65
JF	85.21	91.45	93.04
SJF	86.71	92.46	95.14

From Table 2 the following conclusions can be drawn:

Compared with the model NLMC, the selective joint fine-tuning method model SJF improves the emotion recognition accuracy of dataset1, dataset2 and dataset3 by 6.2%, 3.34%, and 4.45%, respectively.

Comparing the accuracy of the model TTL of the two-layer transfer learning strategy and the selective joint training method model SJF, the accuracy of TTL in dataset1, dataset2, and dataset3 are 84.25%, 90.78%, and 92.65%, respectively. The accuracy rates of SJF on dataset1, dataset2, and dataset3 are 86.71%, 92.46%, and 95.14%, respectively. Compared with TTL, the accuracy rates of SJF on dataset1, dataset2, and dataset3 are improved by 2.46, respectively. %, 1.68%, 2.49%, the reason for the improvement of SJF accuracy is to alleviate the negative transfer problem in the process of IER transfer learning.

Since the jointly trained convolutional neural network has already transferred the dataset pre-trained weights, the model JF fine-tuned with all source domain samples performs similarly to the two-layer transfer learning. The accuracy of emotion recognition of the model SJF using selective joint fine-tuning is 1.5%, 1.01%, and 2.1% higher than that of JF, respectively. This is because SJF selects similar style features from the source domain by selecting samples from the target domain. A subset of are used for transfer, which can effectively alleviate the negative transfer problem, thereby improving the recognition accuracy.

5. Conclusion

FNN is widely used in the field of IER, so this paper studies and analyzes IER based on FNN. This paper draws the following conclusions through research findings: the method of using multi-feature description algorithm to extract multi-level features of images to describe image emotion can effectively improve the accuracy of IER; IER has an impact; through the comparative analysis of the selective joint fine-tuning strategy, it is found that the selective joint fine-tuning training method model SJF has improved the accuracy of IER compared with other models, because SJF effectively alleviates IER. Negative transfer problem in learning process. There are many deficiencies in the text that need to be improved, but the research on IER based on FNN is a good research direction.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Radhi A. Transmission Line Protection based on Neural Network, Fuzzy Neural and Fuzzy Neural Petri Net. Australian Journal of Basic and Applied Sciences. (2021) 5(11):1466-1479.

[2] Hajarolasvadi N, Demirel H. Deep Facial Emotion Recognition in Video Using Eigenframes. IET Image Processing. (2020) 14(14):3536-3546. https://doi.org/10.1049/iet-ipr.2019.1566

[3] Shinnosuke, Ikeda. Investigation of the Role of Emotion Words on Emotion Recognition from Facial Expression, Affective Voice, and Affective Music by Using Semantic Satiation. Japanese Journal of Research on Emotions. (2018) 26(1):12-18. https://doi.org/10.4092/jsre.26.1_12
[4] Mohammed S N, Karim A. Speech Emotion Recognition Using MELBP Variants of Spectrogram Image. International Journal of Intelligent Engineering and Systems. (2020) 13(5):257-266. https://doi.org/10.22266/ijies2020.1031.23
[5] Kim J H. Emotion Training: Image Color Transfer with Facial Expression and Emotion Recognition. Journal of the Korea Computer Graphics Society. (2018) 24(4):1-9. https://doi.org/10.15701/kcgs.2018.24.4.1
[6] Tehmina K, Muhammad A S, Muhammad M, et al. Emotion Recognition from Facial Expressions Using Hybrid Feature Descriptors. IET Image Processing. (2018) 12(6):1004-1012. https://doi.org/10.1049/iet-ipr.2017.0499
[7] Manisha S, Nafisa H S, Gopal N, et al. Bimodal Emotion Recognition using Machine Learning. International Journal of Engineering and Advanced Technology. (2021) 10(4):189-194. https://doi.org/10.35940/ijeat.D2451.0410421
[8] D N V S L S Indira, Babu C S, Rao C V, et al. An Enhanced CNN-2D for Audio-Visual Emotion Recognition (AVER) Using ADAM Optimizer. Turkish Journal of Computer and Mathematics Education (TURCOMAT) (2021) 12(5):1378-1388. https://doi.org/10.17762/turcomat.v12i5.2030
[9] Anandharasu A, Bharathi Raja S, Sakthivel A, Gopikrishnan R. Human Emotion Recognition System using Machine Learning for Videos. International Journal on Recent and Innovation Trends in Computing and Communication. (2019) 7(3):29-36. https://doi.org/10.17762/ijritcc.v7i3.5254
[10] Nguyen H D, Yeom S, Lee G S, et al. Facial Emotion Recognition Using an Ensemble of Multi-Level Convolutional Neural Networks. International Journal of Pattern Recognition and Artificial Intelligence. (2018) 33(11):139-149. https://doi.org/10.1142/S0218001419400159
[11] Ghali A, Kurdy M B. Emotion recognition using facial expression analysis. Journal of Theoretical and Applied Information Technology. (2018) 96(18):6117-6129.
[12] Gervasi O, Franzoni V, Riganelli M, et al. Automating Facial Emotion Recognition. Web Intelligence and Agent Systems. (2019) 17(1):17-27. https://doi.org/10.3233/WEB-190397
[13] Shabrina N H, Omar J, Bhakti A N, et al. Emotion Recognition using Convolutional Neural Network in Virtual Meeting Environment. Jurnal ULTIMA Computing. (2021) 13(1):30-38. https://doi.org/10.31937/sk.v13i1.2108
[14] Jayasimha Y, Reddy R. A Robust Face Emotion Recognition Approach through Optimized SIFT Features and Adaptive Deep Belief Neural Network. Intelligent Decision Technologies. (2019) 13(3):379-390. https://doi.org/10.3233/IDT-190022
[15] Liu Y, Dhakal S, Hao B. Coal and Rock Interface Identification Based On Wavelet Packet Decomposition and Fuzzy Neural Network. Journal of Intelligent and Fuzzy Systems. (2020) 38(4):3949-3959. https://doi.org/10.3233/JIFS-179620
[16] Yazid M, Talib M S, Satria M H, et al. Mortality Prediction for Acute Decompensated Heart Failure Patient Using Fuzzy Neural Network. Malaysian Journal of Fundamental and Applied Sciences. (2020) 16(4):469-474. https://doi.org/10.11113/mjfas.v16n4.1808
[17] Bhattacharjee P K. Mutual Authentication Technique with Four Biometric Entities Applying Fuzzy Neural Network in 5G Mobile Communications. IOSR Journal of Electronics and Communication Engineering. (2020) 15(3):38-46.
[18] Zyryanov S I. The Production Model of Fuzzy Neural Network in Information Security Systems. International Journal of Advanced Trends in Computer Science and Engineering. (2020) 9(4):5190-5192. https://doi.org/10.30534/ijatcse/2020/145942020