

2204207

by Chenlong Jiang

Submission date: 04-May-2023 03:30PM (UTC+0100)

Submission ID: 204730062

File name: 42003_Chenlong_Jiang_2204207_498640_1468693907.docx (1.24M)

Word count: 2719

Character count: 14572

1. Introduction

In this project, we will conduct initial exploratory analysis of the picture data according to the data provided, and carry out operations such as downsampling, pre-processing and visualization of the sample data. Fully connected neural network and convolutional neural network were used to model and analyze the preprocessed data. In the modeling analysis, we will mainly discuss the performance comparison between fully connected neural network and convolutional neural network. The influence of different regularization methods on model training; The influence of the number of hidden layers on training and the influence of activation function on training. Through the analysis of model parameters, the best model is selected, the test set data is predicted, and the final performance of the model is analyzed.

2. Model instruction

2.1 Fully connected neural network

A fully connected neural network, also known as a feedforward neural network or a multilayer perceptron (MLP), is a type of artificial neural network where each neuron in one layer is connected to every neuron in the subsequent layer. In other words, the neurons are arranged in layers, and each neuron in a layer is connected to all the neurons in the next layer. This architecture allows the network to capture complex relationships between inputs and outputs.[1-4]

2.2 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a type of artificial neural network specifically designed for analyzing visual data. It is highly effective in tasks such as image classification, object detection, and image segmentation. CNNs leverage the concept of convolution, which involves applying filters to input data to extract meaningful features.[5-8]

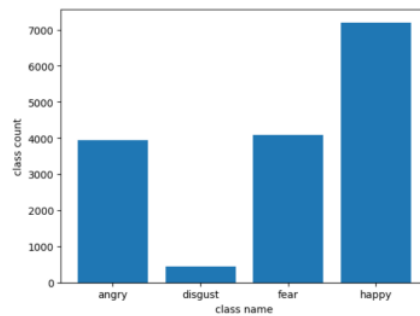
3. Exploratory analysis of data

Since there are many categories of initial data, we will choose four of them for research. They are: 'angry','disgust','fear','happy'. We use opencv to read some sample data, as shown below:

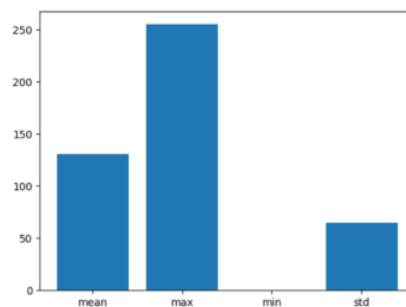


image.shape
(48, 48, 3)

The figure above is a sample data in angry data set. It can be seen from the figure that the size of sample data is 48x48x3. Although the color of the picture is gray, this picture belongs to the picture data of channel 3.



As shown in the figure above, we have counted the number of the four categories in the training set. It can be seen from the figure that the happy category has the largest number and the disgust category has the smallest number. We found that there was an obvious label imbalance in the training set. Therefore, when training the model, it is very difficult for the model to learn the characteristics of the disgust category, which leads to the Disgust recognition effect in this category becoming worse, while the model will learn more characteristics about the happy category, which makes the recognition effect of the model in this category become better.



As shown in the figure above, we calculated the mean value, maximum value, minimum value, variance and other information in the features of the training set. As can be seen from the figure, the extreme value of data features is large, and the variance is also large, indicating that the distribution of data features is uneven, which is easy to cause the machine learning model is not easy to converge during training, and the model learning effect becomes worse. Therefore, we need to normalize the data characteristics.



As shown in the figure above, we randomly output 25 sample pictures from the training set and mark the category of each sample. As can be seen from the figure, the feature difference between different categories is large, which is conducive to the training of the model. At the same time, the feature difference between samples in the same category is also large, which brings great challenges to the model training. The model needs to have a good feature extraction function to better learn the features of each category.

```
x_train=train/255
x_test=test/255
x_train, x_valid, y_train, y_valid= train_test_split(x_train, y, test_size=0.2, random_state = 1234, shuffle=True)
```

Through exploratory data analysis, we need to normalize the data. In this project, we use the Minmax scale principle to preprocess the data. Meanwhile, we only have the training set and test set, but no verification set. Therefore, we further divide the training set into 80% training sets and 20% training sets as verification sets.

4. Model training

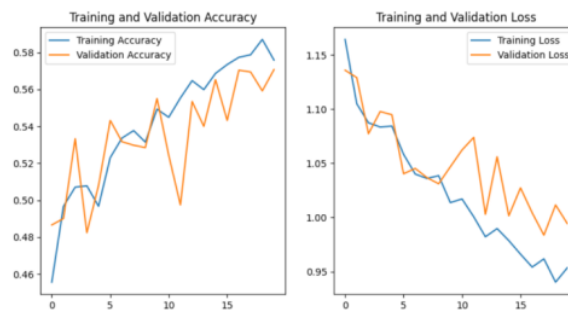
In this project, we will mainly discuss the performance comparison between fully connected neural networks and convolutional neural networks. The influence of different regularization methods on model training; The influence of the number of hidden layers on training and the influence of activation function on training. Through the analysis of model parameters, the best model is selected, the test set data is predicted, and the final performance of the model is analyzed.

For each study, the optimizer used in our fixed model was SGD with moment=0.9, the learning rate was 0.001 (during the training, the learning rate was fixed), the number of training steps was 20 and the batch size was 128.

4.1 Fully connected network

Layer (type)	Output Shape	Param #
flatten_8 (Flatten)	(None, 6912)	0
dense_18 (Dense)	(None, 500)	3456500
dense_19 (Dense)	(None, 500)	250500
dense_20 (Dense)	(None, 4)	2004
Total params: 3,709,004		
Trainable params: 3,709,004		

As shown in the figure above, the model belongs to the fully connected neural network. The model has four layers, including two hidden layers, one input layer and one output layer. The total number of parameters of the model is 3,709,004. Model training results are as follows:

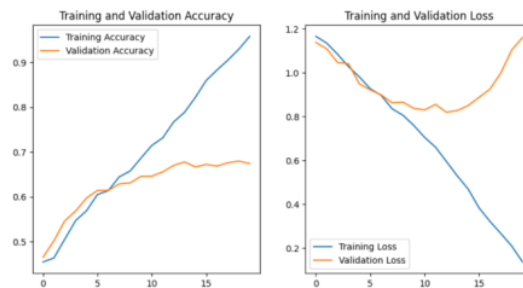


The figure above shows the training results of the fully connected neural network. It can be seen from the figure that the accuracy of the model basically increases stably in the verification set and training set, while the loss of the model basically decreases stably in the verification set and training set. This indicates that the overall training stability of the model is good, but the accuracy of the model in the training set and verification set is only about 0.58, indicating that the model is in a state of underfitting.

4.2 Convolutional neural network

Layer (type)	Output Shape	Param #
conv2d_15 (Conv2D)	(None, 46, 46, 20)	560
max_pooling2d_12 (MaxPooling2D)	(None, 23, 23, 20)	0
conv2d_16 (Conv2D)	(None, 23, 23, 30)	5430
max_pooling2d_13 (MaxPooling2D)	(None, 12, 12, 30)	0
flatten_9 (Flatten)	(None, 4320)	0
dense_21 (Dense)	(None, 500)	2160500
dense_22 (Dense)	(None, 4)	2004
Total params: 2,168,494		
Trainable params: 2,168,494		
Non-trainable params: 0		

As shown in the figure above, this model belongs to the convolutional neural network. The model has five layers, including two convolutional layers, a full connection layer, an input layer and an output layer, and a pooling layer (dimension 2x2) is added behind each convolutional layer. The kernel size of the first convolutional layer is 3x3 and the number of convolutional nuclei is 20; the kernel size of the second convolutional layer is 3x3 and the number of convolutional nuclei is 30. The total number of parameters of the model is 2,168,494. Compared with the fully connected layer, the training parameters of the convolutional network are much less. Model training results are as follows:



The figure above shows the training results of the convolutional neural network. It can be seen from the figure that the accuracy of the model basically increases stably in the verification set and training set, while the loss of the model decreases stably in the training set, but diverges in the verification set. The accuracy of the model in the training set reached more than 0.95, and the accuracy of the verification set also reached 0.67. However, the loss of the model in the verification set was found, indicating that the model was overfitting.

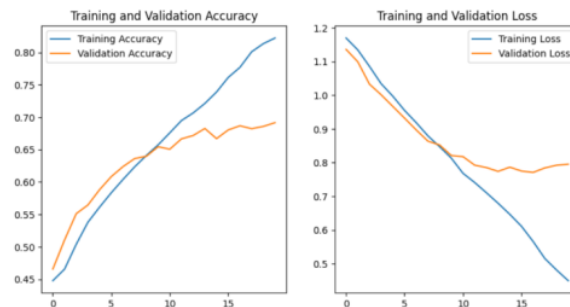
By comparing the fully connected neural network with the convolutional neural network under the same conditions, we find that the convolutional network is superior to the fully connected neural network in all aspects of performance, and the parameters required for training of the convolutional neural network are far smaller than those of the fully connected neural network, thus speeding up the training speed of the model. Meanwhile, the convolutional layer of the convolutional network can effectively extract features from the data. It improves the recognition effect of the model and makes the model have a good accuracy in the verification set. However, the fully connected neural network can only achieve the dimensionality reduction of the data and cannot extract the features of the data, so the recognition effect of the model is poor.

4.3 Regular tuning

In this project, we will use three regularized methods of l1, l2 and dropout to tune the parameters of the convolutional neural network.

4.3.1 Dropout regularity

We add the dropout layer between the full connection layer and the output layer of the model with a probability of 0.5, which means that in each iteration, there is a 50% probability that some parameters will be deactivated. The training results of the model are as follows:

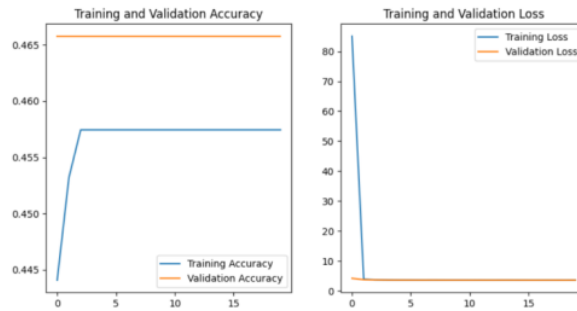


It can be seen from the figure that, compared with the convolutional neural network without regularization, after adding the dropout layer, the model's loss on the verification set decreases stably, with no obvious divergence trend. At the same time, the accuracy of the

model improves to 0.69 on the verification set, indicating that the model avoids overfitting after adding the regularization term. The recognition effect of the model on verification set is improved.

4.3.2 L1 regular

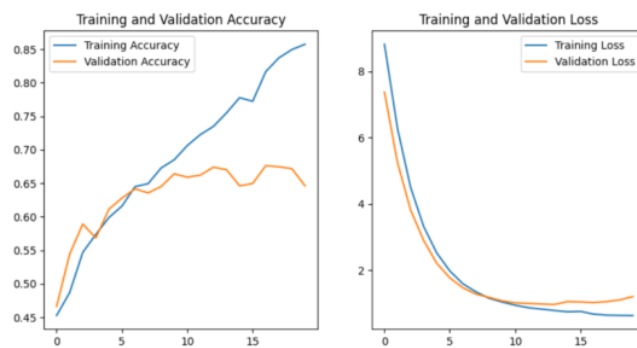
l_1 regularization was added to the fully connected layer, and the regularization coefficient was 0.01. The model training results were as follows:



As can be seen from the figure, after adding l_1 regularization to the model, the regularization coefficient is 0.01. In the process of iteration, the accuracy and loss of the model did not change significantly, indicating that when the regularization coefficient was 0.01 and l_1 regularization was used, the training of the model was severely inhibited, resulting in the model failed to learn any features.

4.3.3 L2 regular

l_2 regularization was added to the full-connection layer, and the regularization coefficient was 0.01. The model training results were as follows:



As can be seen from the figure, compared with the convolutional neural network without regularization, the loss of the model on the verification set decreased stably after 12 regularization was added, and there was no obvious divergence trend. However, the accuracy rate of the model on the verification set was still around 0.67, and the final accuracy rate was reduced to 0.64. Therefore, in this case, Although the loss change of the model is stable, the accuracy of the model decreases more.

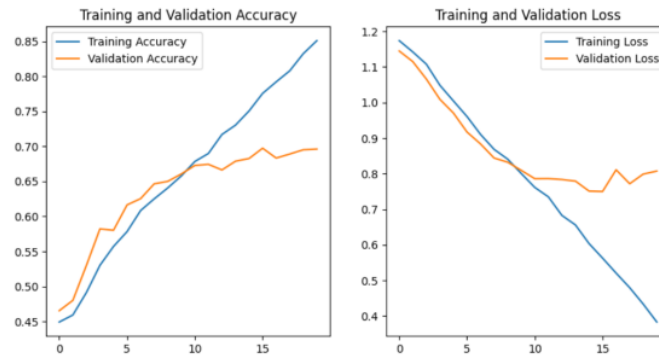
In summary, after trying different regularization methods, the method of adding dropout layer can more effectively avoid overfitting the model on the training set, and also improve the recognition accuracy of the model on the verification set.

4.4 Influence of model structure

In this project, we will study the influence of different convolutional layers on convolutional neural networks. In the initial model, the convolutional network constructed by us has two convolutional layers. We will study the influence of three convolutional layers and four convolutional layers on model training.

4.4.1 Three convolutional layers

We added a convolution layer into the second convolution layer of the initial model, with a convolution kernel of 3x3 and a number of 30. The model training results are as follows:

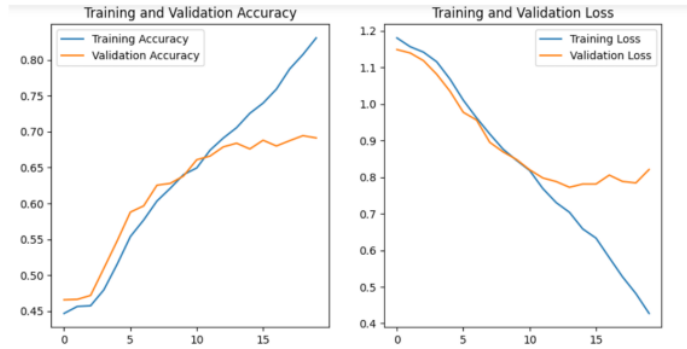


As can be seen from the figure, the performance of the model in the training set and the verification set is not much different from that of the initial model, indicating that adding a convolutional layer does not significantly improve the recognition effect of the model.

4.4.2 Four convolutional layers

We added two convolutional layers into the second convolutional layer of the initial model, and the convolutional kernel was 3x3 and the number was 30. The model training results were

as follows:

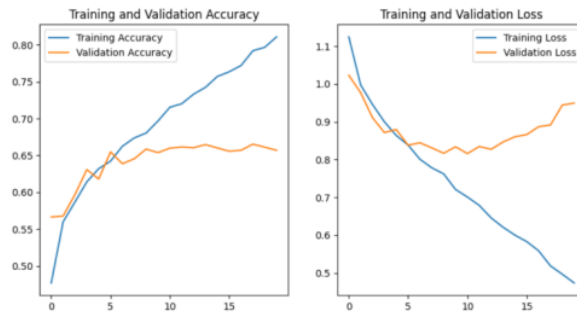


As can be seen from the figure, the performance of the model in the training set and the verification set is not much different from that of the initial model, indicating that adding two convolutional layers does not significantly improve the recognition effect of the model.

Through the above results, we found that increasing the number of convolutional layers did not improve the recognition effect of the model. As the depth of the network increases, the gradients propagated during backpropagation may diminish (vanishing gradients) or explode (exploding gradients). This can make it challenging to effectively update the weights of the earlier layers, hindering the convergence of the network.

4.5 Activation function study

Through the above studies, we determine that the optimal model is the convolutional neural network with the addition of dropout regularization term. In the initial model, the activation function we use is relu. In this section, relu is replaced by tanh activation function, and the model training results are as follows:

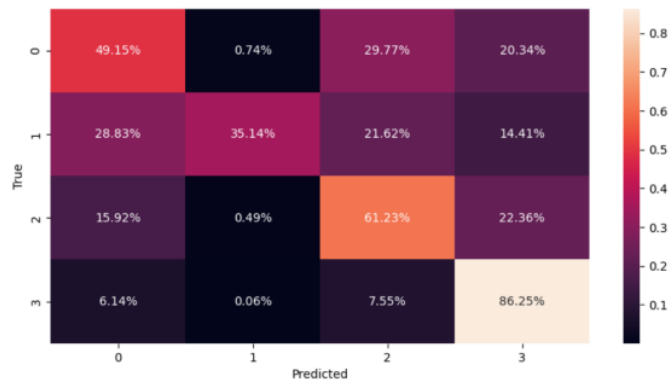


As can be seen from the figure, the performance of the model on the verification set and the training set is worse than that of the relu, and the loss of the model on the verification set still shows a divergence trend after the addition of the regular term, indicating that the tanh

activation function is not suitable for the modeling of this data set.

4.6 Test set evaluation

Classification report:					
	precision	recall	f1-score	support	
0	0.60	0.49	0.54	944	
1	0.75	0.35	0.48	111	
2	0.59	0.61	0.60	1024	
3	0.78	0.86	0.82	1774	
accuracy			0.69	3853	
macro avg	0.68	0.58	0.61	3853	
weighted avg	0.68	0.69	0.68	3853	



As shown in the figure above, we calculate the classification report and confusion matrix of the optimal model on the test set. As can be seen from the figure, the overall prediction accuracy of the model is 0.69, among which the prediction accuracy of category 3 is the highest, and that of category 1 is the lowest. This is mainly caused by the phenomenon of sample imbalance in the data. During model training, the model is likely to find it difficult to learn the characteristics of the disgust category, which leads to a deterioration in the recognition effect of the model in this category, while the model will learn more features about the happy category, which makes the recognition effect of the model in this category better.

5. Conclusion

In this project, according to the data provided, we conducted an initial exploratory analysis of the picture data, conducted downsampling, pre-processing visualization and other operations on the sample data, and found that there was obvious label imbalance in the training set. Fully connected neural network and convolutional neural network were used to model and analyze the preprocessed data. In the modeling analysis, we mainly discuss the performance comparison between fully connected neural network and convolutional neural network. The influence of different regularization methods on model training; The influence of the number of hidden layers on training and the influence of activation function on training. Through the analysis of model parameters, the best model is selected, the test set data is predicted, and the final performance of the model is analyzed. The results show that the convolutional network is superior to the fully connected neural network in terms of performance, and the parameters required for training of the convolutional neural network are much smaller than those of the fully connected neural network, which speeds up the training speed of the model. At the same time, the convolutional layer of the convolutional network can effectively extract features from the data, improve the recognition effect of the model, and make the model have a good accuracy in the verification set. However, the fully connected neural network can only achieve dimensionality reduction of data, and cannot extract features from data, so the model recognition effect is not good. The method of adding dropout layer can effectively avoid overfitting the model on the training set and improve the recognition accuracy of the model on the verification set. Increasing the number of convolutional layers did not improve the recognition effect of the model. As the depth of the network increases, the gradients propagated during backpropagation may diminish (vanishing gradients) or explode (exploding gradients). This can make it challenging to effectively update the weights of the earlier layers, hindering the convergence of the network; tanh activation function is not suitable for modeling this data set; The overall prediction accuracy of the model on the prediction set is 0.69, among which the prediction accuracy of category 3 is the highest, and that of category 1 is the lowest. This is mainly caused by the phenomenon of sample imbalance in the data. During model training, the model is likely to find it difficult to learn the characteristics of the disgust category, which leads to a deterioration in the recognition effect of the model in this category, while the model will learn more features about the happy category, which makes the recognition effect of the model in this category better.

Reference

- [1]Neural Networks and Deep Learning by Michael Nielsen:
<http://neuralnetworksanddeeplearning.com/>
- [2]Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville:
<http://www.deeplearningbook.org/>
- [3]Pattern Recognition and Machine Learning by Christopher M. Bishop:
<https://www.springer.com/gp/book/9780387310732>
- [4]Deep Learning with Python by François Chollet:
<https://www.manning.com/books/deep-learning-with-python>
- [5]Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville:
<http://www.deeplearningbook.org/>
- [6]Convolutional Neural Networks for Visual Recognition (CS231n) by Stanford University:
<http://cs231n.github.io/>
- [7]Deep Learning with Python by François Chollet:
<https://www.manning.com/books/deep-learning-with-python>
- [8]Deep Learning for Computer Vision with Python by Adrian Rosebrock:
<https://www.pyimagesearch.com/deep-learning-computer-vision-python-book/>

FINAL GRADE

58/100

GENERAL COMMENTS

Instructor

Please include an abstract.

Your work could benefit from a final check to catch typos and improve grammar, etc.

Structured report submitted with use of references. Try to use refs more often to support arguments or assert facts. Don't be afraid to use refs if needed. Currently you overuse electronic references. Try to improve the quality of your refs - locate seminal papers, etc and also identify work that attempts similar experiments and/or uses the same dataset. This will provide you will points of reference and can help validate you designs.

Make sure you label figures and include titles, label axes, etc. This is important as it helps the reader understand your work better.

Currently it is difficult to understand your discussions regarding, for example, the relevance of descriptive statistics when reviewing the dataset.

Explain further your architectural choices for you're MLP.

Bar graphs will help you compare results. Use them.

You do not compare your final CNN with your MPL - what is the purpose of experimenting with the MLP prior to the CNN? Explain this and discuss your results - draw appropriate conclusions.

You experiment with several parameters in a structured way and present results reasonably clearly. Good. Try to bring the whole together in a more succinct manner.

Conclusions generally appropriate.

Did you consider dimensionality reduction etc for pre-processing in the MPL?

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

GRADING FORM: SMU GENERIC_GJ

CHENLONG JIANG

0

RESEARCH-INFORMED LITERATURE

KNOWLEDGE AND UNDERSTANDING

ANALYSIS

PRACTICAL APPLICATION AND DEP.

SKILLS FOR PROFESSIONAL PRAC.