Binary cross entropy with deep learning technique for Image classification

Article *in* International Journal of Advanced Trends in Computer Science and Engineering · October 2020

DOI: 10.30534/ijatcse/2020/175942020

Vamsidhar Yendapalli
GITAM University

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/ijatcse175942020.pdf
https://doi.org/10.30534/ijatcse/2020/175942020

Binary cross entropy with deep learning technique for Image classification

Dr.A.Usha Ruby[1], Prasannavenkatesan Theerthagiri[2], Dr.I.Jeena Jacob[3], Dr.Y.Vamsidhar[4]
[1,2,3,4]GITAM School of Technology, Bengaluru,
[1]uruby@gitam.edu, [2]prasanna91@gmail.com, [3]jeni.neha@gmail.com, [4]vyendapa@gitam.edu

ABSTRACT

This paper discussed about a unique neural network approach stimulated by a technique that has reformed the field of computer vision: pixel-wise image classification, which we combine with binary cross entropy loss and pre training of the CNN (Convolutional Neural Network) as an auto encoder. The pixel-wise classification technique directly estimates the image source label for each time-frequency (T-F) bin in our image, thus eliminating common pre- and-post processing tasks. The proposed convolutional neural network is trained by using the binary mask as the target output label. The binary mask identifies the dominant image source in each T-F bin of the magnitude spectrogram of a mixture signal, by considering each T-F bin as a pixel with a multi-label (for each image source). Binary Cross Entropy is used as the training objective, so as to minimize the average probability error between the target and predicted label for each pixel. The Inception V3 architecture is used to further boost ImageNet classification accuracy. The results show that the proposed algorithm has the highest accuracy.

Key words : Convolutional Neural Network, Inception V3, Softmax

1. INTRODUCTION

Nowadays, the artificial intelligence with deep learning has the tremendous growth in diverse fields with variety of rapid developmental functionalities. The impact of deep learning plays a vital role many research areas including image classification, natural language processing, speech recognition, computer vision, medical image processing, etc. The deep learning considers the large number of features, parameters and functions in order to solve critical problems and to arrive the decisions or to define the relationships on different classes dataset. Deep learning manages the dataset by mapping them with high dimension space. For image classification and processing, the deep convolutional neural network provides the extra-ordinary support with advanced techniques [1].

The high-resolution image classification are highly supported by the improved convolutional neural networks. Typically, the deep neural networks are employed for the larger input dataset model. In case of image input dataset, several parameters, features of images with large number of images are adopted during training by the deep neural networks. The deep neural networks works by extracting the semantic features and network fuses of the features among the image dataset in order to classify the given images.

The dropout and batch normalization methods are used to improvise the any model's generalization performance using convolutional neural networks. Indeed, the dataset with large number of training samples might improve the generalization performance. Further, the data enhancement approaches such as cropping, translation, flipping, and rotation on training samples also provides the better performance for convolutional neural networks [1].

The performance of an image classification model is estimated and measured using the cross-entropy loss. The probability output value of the cross-entropy loss function is between 0 and 1. The cross-entropy loss increases when the predicted probability of the image classification belongs to the actual class [6]. If the predicted probability diverges near to zero, then the current image classification model can be recognized as the bad classification model; it would produce a high loss value. The binary cross-entropy is a special class of cross-entropy, where the target of the prediction id 1 or 0. It applies the sigmoid activation for the prediction using deep neural networks. The cross-entropy can be used, even if the target value is not a probability vector [9].

In this work, we propose a novel binary cross loss function for image classification using deep CNN (Convolutional Neural Network). The remainder of the paper is organized as follows. Section 2 of this paper describes the recent literature works related to the image classification using cross entropy function, deep learning, and convolutional neural networks. The rest of the sections describes about proposed methods and experimental results.

2. RELATED WORKS

Two-stage deep convolutional neural networks image classification method was applied to overcome the issues of limited training dataset and learning parameters. This article [1] used the inner-move image enhancement method with augmented training samples for the image classification. It improves the generalization ability of image classification tasks using convolutional neural networks [1]. Mannor et al., (2005) had adopted the cross entropy method for the binary classification in support vector machines. This method searches the support vectors using linear programs in order to solve the optimization problem [2].

Authors Pasupa et al., (2020) had proposed a convolutional neural network deep learning technique to solve the class imbalance problem in the red blood cell morphology. The focal loss function was used to reduce the misclassification. Authors had shown this approach has better F1-scores with reduced bias for the classification of imbalanced dog red blood cell dataset [3]. The Single Logit Classification was developed for the fast real-time image search by Keren et al., (2018). The cross-entropy loss function was aligned with the principle of logit separation for output classification. This article had studied many loss functions and suggests that principle of logit separation is produced higher relative accuracy with lower losses [4].

A novel Real-World-Weight Cross-Entropy Loss Function was proposed by the authors Ho et al., (2019). This new loss function was used to measure the goodness-of-fit cost functions for the classifiers. The financial impact and information about a real world problem are considered as weights in this metric, whereas F1 scores and accuracy didn't considered these factors. The authors had suggested this metric can be used single-label multiclass classification, binary classification, and its variants [5].

A softmax cross entropy loss function by considering position of decision boundary is proposed in this article Cao et al., (2018). This loss function is derived in association with L2-norm parameter of neural network output in each class. Based on the experimental analyses the proposed loss function was shown with less bias as compared with conventional softmax cross entropy loss function [6]. Categorical Cross Entropy (CCE) loss was proposed by the authors to reduce the error performance rates. It used the noise-robust loss functions for the MAE and CCE. In this article [7], the categorical cross entropy loss function was applied with conventional DNN architecture.

In order to solve the image misclassification problems, rather than normal training and testing, the differential training method was proposed in the article [8]. The differential training method adopts the low-rank features and cross-entropy loss function to differentiate the opposite classes and to minimize the adversarial classification. The experimental results was shown as the differential training offers the greater margin among the training dataset and decision values of neural network. Thus the larger margin is suggested that it improves the performance of the classifier on binary classification task [8].

A new cross entropy loss function was proposed for effective learning of the binary descriptors and dubbed CE-Bits. In this method, the normal binary code is broke into several independent blocks and each of them are optimized separately. Its results are shown as improved as compared with L-2 and hinge loss functions for both classification problems [9].

Typically, uncertainty assessment techniques have been applied to evaluate the accuracy and error metrics of the algorithm such as confusion matrix, MSE, etc. Shadman et al., (2019) evaluated the uncertainty assessment techniques among deep neural network and random forest models. They employed the Shannon entropy to predict the classification probabilities on each pixels. By the RMSE error metrics analysis on deep neural network, the derived entropy was estimated as better uncertainty model with least errors [10].

The face liveness detection architecture was developed in this article [11] in order to detect and classify weather the detected image is a photograph of the face or real face in the face recognition application. This article adopted the deep convolutional neural network and texture analysis algorithms to reduce the face spoofing attacks on biometric authentication. The block-solver algorithm and nonlinear-diffusion splitting scheme was applied to enhance the images based on edges and surface texture. This algorithm classifies the images based on complex and deep features; also, its results are analyzed with inception network and residual network [11].

3. PROPOSED BINARY CROSS ENTROPY WITH DEEP LEARNING TECHNIQUE FOR IMAGE CLASSIFICATION

Deep learning is a type of machine learning in which a model learns to perform classification tasks directly from images, text, or sound. Deep learning is usually implemented using neural network architecture. The term "deep" refers to the number of layers in the network—the more layers, the deeper the network. Traditional neural networks contain only 2 or 3 layers, while deep networks can have hundreds. A deep neural network combines multiple nonlinear processing layers, using simple elements operating in parallel and inspired by biological nervous systems. It consists of an input layer, several hidden layers, and an output layer. The layers are interconnected via nodes, or neurons, with each hidden layer using the output of the previous layer as its input.

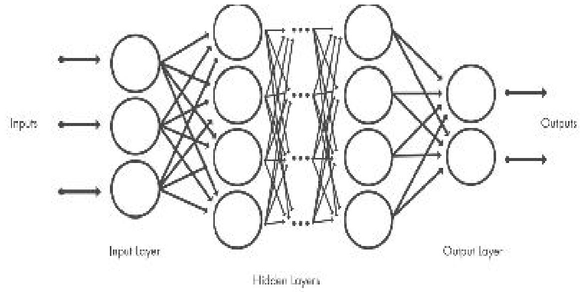

Figure 1: Inception v3 model

ImageNet is correctly meant as labeling and categorizing images into almost 22,000 separate object categories for the purpose of computer vision. However, when we hear the term "ImageNet" in the context of deep learning and Convolutional Neural Networks, we are likely referring to the ImageNet Large Scale Visual Recognition Challenge or ILSVRC for short. The goal of this image classification challenge is to train a model that can correctly classify an input image into 1,000 separate

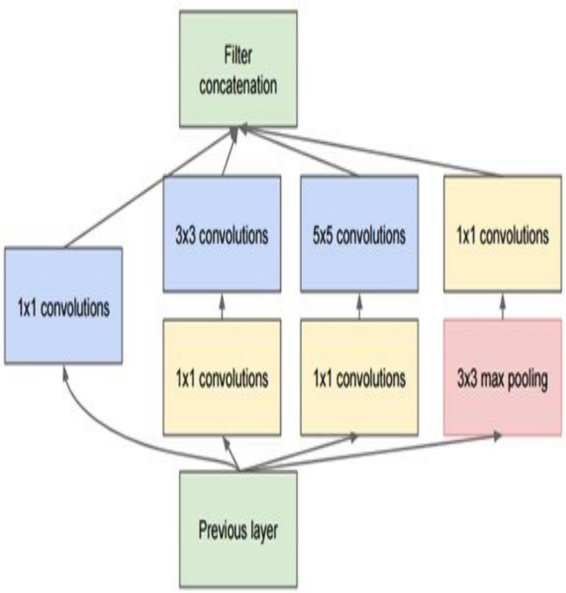

Figure 2: Inception V3 module

object categories. Models are trained on ~1.2 million training images with another 50,000 images for validation and 100,000 images for testing. These 1,000 image categories represent object classes that we encounter in our day-to-day lives, such as species of dogs, cats, various household objects, vehicle types, flowers and much more. When it comes to image classification, the ImageNet challenge is the de facto benchmark for computer vision classification algorithms — for this challenge has been dominated by Convolutional Neural Networks and deep learning techniques. The state-of-the-art pre-trained networks included in the Keras core library represent some of the highest performing Convolutional Neural Networks on the ImageNet. These networks also demonstrate a strong ability to generalize to images outside the ImageNet dataset via transfer learning, such as feature extraction and fine-tuning.

The goal of the inception module in Figure 1 and Figure 2 is to act as a "multi-level feature extractor" by computing 1×1, 3×3,and 5×5 convolutions within the same module of the network — the output of these filters are then stacked along the channel dimension of images and before being fed into the next layer in the network. The Inception V3 architecture is used to further boost ImageNet classification accuracy. The weights for Inception V3 are smaller than both VGG and ResNet, coming in at 96MB. The images are resized for the binary cross entropy models are shown in Figure 3 and Figure 4.

The Softmax classifier uses the cross-entropy loss. The Softmax classifier gets its name from the softmax function, which is used to squash the raw class scores into normalized positive values that sum

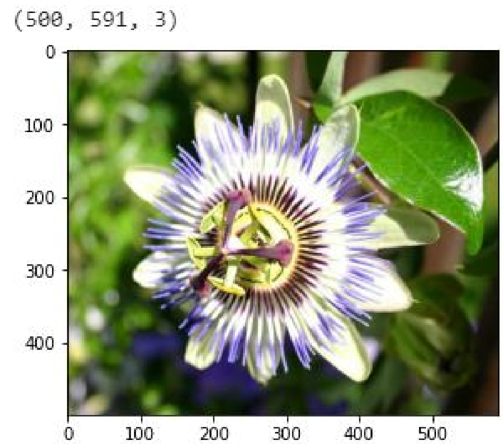

Figure 3: Image before cropping

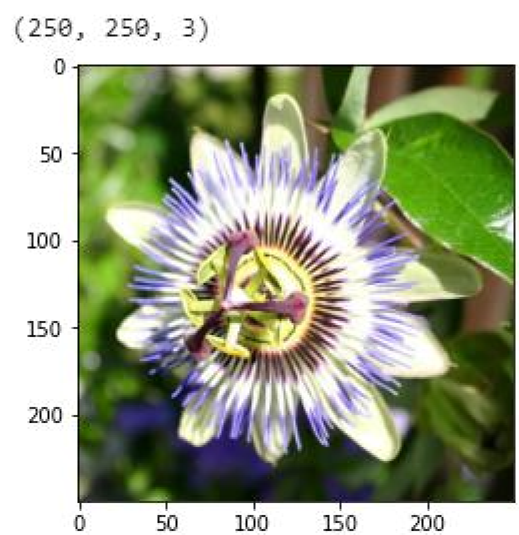

Figure 4: Cropped image

to one, so that the binary cross-entropy loss can be applied. The loss function was defined in such a way that making good predictions on the training data is equivalent to having a small loss. The Softmax regression is a form of logistic regression that normalizes an input value into a vector of values that follows a probability distribution whose total sums up to The output values are between the range [0,1] which is good because we are able to avoid binary classification and accommodate as many classes or dimensions in our neural network model. This is why softmax is sometimes referred to as a multinomial logistic regression. The function is usually used to compute losses that can be expected when training a data set. Known use-cases of softmax regression are in discriminative models such as binary cross entropy and noise contrastive estimation. However, it should be well-known that softmax is not preferably used as an activation function like Sigmoid or ReLU (Rectified Linear Units) but rather between layers which may be multiple or just a single one. This softmax function ς takes as input a C-dimensional vector z and outputs a C-dimensional vector y of real values between 0 and 1. This function is a normalized exponential and is defined as in the equation (1):

$$y_c = \varsigma(z)_c = \frac{e^{z_c}}{\sum_{d=1}^{C} e^{z_d}} \tag{1}$$

As the output layer of a neural network, the softmax function can be represented graphically as a layer with C neurons. We can write the probabilities that the class is t=c for c=1…C given input z as in the equation (2):

$$\begin{bmatrix} P(t=1|z) \\ \vdots \\ P(t=C|z) \end{bmatrix} = \begin{bmatrix} \varsigma(z)_1 \\ \vdots \\ \varsigma(z)_c \end{bmatrix} = \frac{1}{\sum_{d=1}^{C} e^{z_d}} \begin{bmatrix} e^{z_1} \\ \vdots \\ e^{z_C} \end{bmatrix} \tag{2}$$

Where P(t=c|z) is thus the probability that that the class is c given the input z. To use the softmax function in neural networks, we need to compute its derivative and it is given in the equations (3),(4) and (5).

$$i = j: \frac{\partial y_i}{\partial z_i} = \frac{\partial \frac{e^{z_i}}{\sum C}}{e^{z_i}} = \frac{e^{z_i}\sum C - e^{z_i}e^{z_i}}{\sum_c^2} = \frac{e^{z_i}}{\sum C}\frac{\sum C - e^{z}}{\sum C} \tag{3}$$

$$= \frac{e^{z_i}}{\sum C}\left(1 - \frac{e^{z_i}}{\sum C}\right) = y_i(1 - y_i) \tag{4}$$

$$i \neq j: \frac{\partial y_i}{\partial z_i} = \frac{\partial \frac{e^{z_i}}{\sum C}}{e^{z_j}} = \frac{0 - e^{z_i}e^{z_i}}{\sum_c^2} = -\frac{e^{z_i}}{\sum C}\frac{e^{z_j}}{\sum C} = -y_i y_j \tag{5}$$

4. EXPERIMENTAL RESULT

The Figure 5 shows the result of epochs using binary cross entropy with adamx and softmax classifier for 102flowers.tgz database and Figure 6 shows the result of test accuracy with 0.95625 with Inception V3 model.

```
Epoch 10/16
25/25 [==============================] - 345s 14s/step - loss: 0.2265 - accuracy: 0.9538 - val_loss: 0.2715 - val_accuracy: 0.9245
Epoch 11/16
25/25 [==============================] - 342s 14s/step - loss: 0.1422 - accuracy: 0.9762 - val_loss: 0.2474 - val_accuracy: 0.9375
Epoch 12/16
25/25 [==============================] - 341s 14s/step - loss: 0.1272 - accuracy: 0.9737 - val_loss: 0.2997 - val_accuracy: 0.9271
Epoch 13/16
25/25 [==============================] - 334s 13s/step - loss: 0.0757 - accuracy: 0.9900 - val_loss: 0.2446 - val_accuracy: 0.9453
Epoch 14/16
25/25 [==============================] - 335s 13s/step - loss: 0.0817 - accuracy: 0.9887 - val_loss: 0.2438 - val_accuracy: 0.9245
Epoch 15/16
25/25 [==============================] - 345s 14s/step - loss: 0.0555 - accuracy: 0.9950 - val_loss: 0.1854 - val_accuracy: 0.9531
Epoch 16/16
25/25 [==============================] - 346s 14s/step - loss: 0.0406 - accuracy: 0.9962 - val_loss: 0.2777 - val_accuracy: 0.9115
<tensorflow.python.keras.callbacks.History at 0x7fb71e23ad30>
```

Figure 5: Result of epochs using binary cross entropy

```
0.956250011920929
```

Figure 6: Test accuracy

5. CONCLUSION

Binary cross entropy is a loss function that is used in binary classification tasks. The binary cross entropy is very convenient to train a model to solve many classification problems at the same time, if each classification can be reduced to a binary choice. In this paper, our work shows the importance of loss function for muti-class classifier learning in convolutional neural networks. We have proposed a binary cross entropy loss function with softmax classifier that utilizes the maximum probability of predictive value to reduce the cross entropy loss of each iteration. In order to demonstrate the performance of this method on experiments with 102flowers.tgz, training accuracy is 95.62% is achieved from various epochs.

REFERENCES

1. Tang C, Zhu Q, Wu W, Huang W, Hong C, Niu X. **PLANET: Improved Convolutional Neural Networks with Image Enhancement for Image Classification. Mathematical Problems in Engineering.** 2020 Mar 11;2020. https://doi.org/10.1155/2020/1245924
2. Mannor S, Peleg D, Rubinstein R. **The cross entropy method for classification. In Proceedings of the 22nd international conference on Machine learning**, 2005 Aug 7 (pp. 561-568).
3. Pasupa, K., Vatathanavaro, S. & Tungjitnob, S. **Convolutional neural networks based focal loss for class imbalance problem: a case study of canine red blood cells morphology classification**, J Ambient Intell Human Comput (2020). https://doi.org/10.1007/s12652-020-01773-x
4. Keren G, Sabato S, Schuller B. **Fast single-class classification and the principle of logit separation**, In

2018 IEEE International Conference on Data Mining (ICDM) 2018 Nov 17 (pp. 227-236). IEEE.
5. Ho Y, Wookey S. **The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling**, IEEE Access. 2019 Dec 27;8:4806-13.
6. Cao J, Su Z, Yu L, Chang D, Li X, Ma Z. **Softmax cross entropy loss with unbiased decision boundary for image classification**, In 2018 Chinese Automation Congress (CAC) 2018 Nov 30 (pp. 2028-2032). IEEE.
7. Zhang Z, Sabuncu M. **Generalized cross entropy loss for training deep neural networks with noisy labels**, In Advances in neural information processing systems 2018 (pp. 8778-8788).
8. Nar K, Ocal O, Sastry SS, Ramchandran **K. Cross-entropy loss and low-rank features have responsibility for adversarial examples,** arXiv preprint arXiv:1901.08360. 2019 Jan 24.
9. Liu L, Qi H**. Learning effective binary descriptors via cross entropy**, In 2017 IEEE Winter Conference on Applications of Computer Vision (WACV) 2017 Mar 24 (pp. 1251-1258). IEEE. https://doi.org/10.1109/WACV.2017.144
10. Shadman Roodposhti M, Aryal J, Lucieer A, Bryan BA. **Uncertainty assessment of hyperspectral image classification: Deep learning vs. random forest. Entropy**, Entropy. 2019 Jan;21(1):78. https://doi.org/10.3390/e21010078
11. Koshy R, Mahmood A. **Optimizing Deep CNN Architectures for Face Liveness Detection**. Entropy. 2019 Apr;21(4):423. https://doi.org/10.3390/e21040423