International Innovative Research Journal of Engineering and Technology
ISSN: 2456-1983 Vol: 5 Issue: 3 March 2020

Identification of Violence Images Using Convolution Neural Network in Social Network

Santhi Baskaran[1], Jahnavi Korrapati[2], K. Sooriya[3], R. Pavithra[4]

[1,2,3,4]Department of Information Technology, Pondicherry Engineering College, Puducherry, India
[1]santhibaskaran@pec.edu, [2]jahnavi.korrapati123@gmail.com, [3]sooriya2208@gmail.com,
[4]pavithraragu08@gmail.com

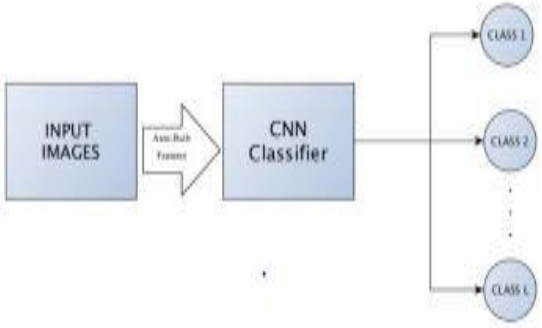
Abstract: In social networks some people may upload images that will contain violent content. In order to delete them, we have to go back to the database and search the offending image among all and clear them. Instead of doing this we use the software where we train it to learn features like bloodshed, by giving several datasets of violent images in order to identify the images containing alcohol, violence, weapons etc. This method aims to filter violent images by using a neural network to classify between the image categories. Leveraging the technique of image classification with neural nets, this project aims to filter inappropriate image content. This can be done through Convolution Neural Network (CNN) by processing an image.

Keywords: Image classification, Convolution Neural Network (CNN), Violence detection

1. INTRODUCTION

In today's modern world internet has become an day to day thing. There is a huge amount of data present in internet in form of images. Classification of images is very much important so that access to certain images can be restricted to certain age group. So it is necessary to identify harmful content in images. In this paper, we just focus on the recognition of violence images. This project aims to filter violent images social networks by using a neural network to classify between the images. Leveraging the technique of image classification with neural nets, this project aims to filter inappropriate image content. This can be done through Convolution Neural Network (CNN) by processing an image. We train the machine with diverse datasets containing images of violence. After training (supervised) given to the machine with datasets of images of violence. The machine will learn the common patterns of an violent image such as (bloodshed, weapons, expressions, mobfights), so that the machine can classify between the safer images to non-safer images and can deny the access of uploading these. Sometimes when there is a wrong output the machine has to be remodeled and the machine has to be trained more on that error output with more diverse datasets. We have used technique of Transfer learning (Densenet-50) to improve the accuracy. From the existing binary classifiers we have included multiclass classifiers.

In order to reduce the data lost in pooling process in CNN an alternative method of random shifting will be used. It not only reduces the training time but also improves the performance of the network. The purpose of image classification is to find the percentage of belongs of an given image to a label.



2. DATABASE

Since we were unable to find a proper publicly available database of violence images, we established our own database consisting of 1000 violence images and 2000 nonviolence images. Most of the images mainly come from the online searching engines like Google. Query words such as "violence", "horror", fight", "explosion", "blood", "and gunfire" and so on were used when collecting these images.

3. LITERATURE SURVEY

In table 1.1 various image classification algorithms and techniques have been compared with their respective advantages and disadvantages.

Figure 1. Samples of images in the database

Table 1. Comparative study on Various Existing Image Classification Techniques

SL. No	Name of The Journal, Year	Title of Paper	Technique used	Dataset used	Parameters	Results	Limitations
1.	IEEE TRANSACTIONS ON MEDICAL IMAGING, FEBRUARY 2018	A Deep Cascade of Convolution Neural Networks for Dynamic MR Image Reconstruction	CNN technique	Cardiac MR dataset	Learning rate α	Training CNN on comparatively smaller datasets and acquiring efficiency	Based on only one complexity.
2.	IEEE International Conference on Knowledge and Systems Engineering(KSE),2017	Advertisement Image Classification Using Convolution Neural Network	The processing component use CNN techniques.	online captured images.	Number of layers and number of filters.	Image classification is critical and significant research problems in computer.	The efficiency percentage is less than 90%
3.	IEEE International Conference on Systems, Man, and Cybernetics San Antonio, TX, USA - October 2009	Violent Web images classification based on MPEG7 color descriptors	Data mining techniques	Image vectors	kernel width	Image filtering with techniques of data mining and image analysis.	Efficiency is less than 90%

4.	IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance 2012	Baseline Results for Violence Detection in Still Images	Bag of words method	new established dataset	Accuracy	Violence Detection with four representation and are tested.	Cannot recognize more diversified violent images
5.	IEEE Twenty-Sixth International Joint Conference on Artificial Intelligence	Random Shifting for CNN: a Solution to Reduce Information Loss in Down-Sampling Layers	Random shifting	CIFAR-10 , CIFAR-100 and ImageNet .	Accuracy	Data loss is reduced	Not many extensions are present.

4. MODEL DESIGN

Our convolution consists of following steps.

1. The feature arrays are lined up with the image array.
2. Each pixel in image is multiplied by corresponding pixel in feature.
3. The values thus obtained are added and sum is found.
4. The sum is divided by number of pixels in feature.

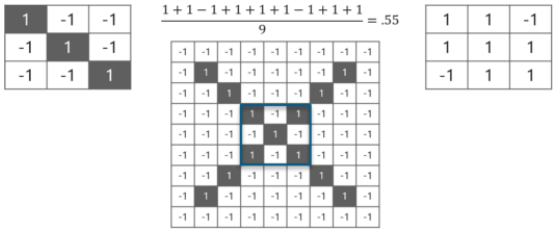


Figure 2. Convolution

Rectified Linear Unit (ReLU) is an activation function thatonly activates a node if the input is above a certain quantity. When the input value to it is below zero, the output is zero, but when the input rises above a certain threshold, it maintains a linear relationship with the dependent variable.

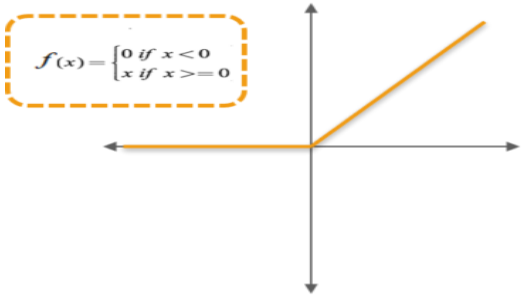


Figure 3. ReLU

Pooling reduces the spatial size of the representation to reduce the amount of parameters and computation in the network. Pooling layer operates on each feature map independently. We have used Max pooling and random shifting is an method that could be used to increase classification efficiency.

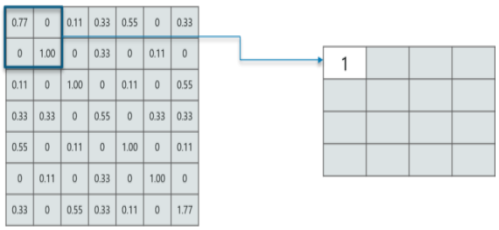


Figure 4. Pooling layer

Fully connected layer is the final layer where the prediction occurs. Final Image array is converted into one single list. The prediction is based on comparing the obtained values of output data with list of trained data.

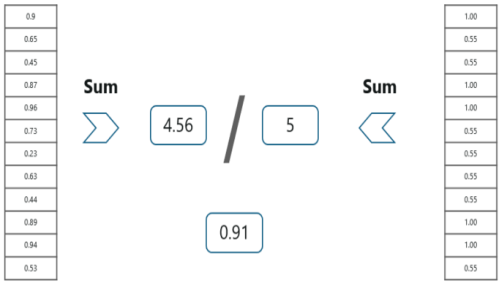


Figure 5. Fully connected layer

Back Propagation:

- The error from actual output to our model output is calculated.
- If the error value is minimum then, ignore.

International Innovative Research Journal of Engineering and Technology
ISSN: 2456-1983 Vol: 5 Issue: 3 March 2020

- Else update the parameters and the process is repeated.
- When the error becomes minimum the model will be ready for making prediction.

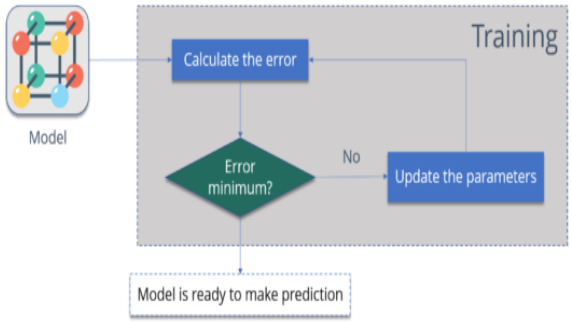


Figure 6. Back Propagation

5. PARAMETERS

1. Epoch: epoch is one complete presentation of the data set to be learned by the learning machine .For example in module 1 we are training 10 epoch and each epoch contain 60 images.

2. Number of images: Number of images as data that is used to train system. we have used nearly 3000 images.

3. Learning Rate: It is a configurable hyper parameter used in training of neural networks.

Weight = α (gradient).

α is learning rate and gradient is random function value function generated by python

4. Accuracy: It is the fraction of predictions our model got right

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where *TP* = True Positives, *TN* = True Negatives, *FP* = False Positives, and *FN* = False Negatives.

5. Loss Function: loss function is to compare how good or bad our prediction was in relation to correct result. The relation between loss and weight is

dl/dw

Where, l is loss and w is weight

binary cross entropy loss function

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

Where, y is the label, p(y) is the predicted probability, N is Number of images.

.

6. CONCLUSION

Thus in this paper we have proposed a model to filter violent images in social networks by using a neural network. The model classify between the images as violent and non-violent. We have established a new image database covering a large variety of violent sources, which will facilitate further research on violence image detection.

REFERENCES

[1] Ozan Ok tay, Enzo Ferrante, Konstantin's Kamnitsas, Mattias Heinrich, Wenjia Bai, Jose Caballero , Stuart A. Cook, Antonio de Marvao, Timothy Dawes, Declan P. O'Regan, Bernhard Kainz, Ben Glocker, and Daniel Rueckert," *Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation*", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 37, NO. 2, FEBRUARY 2018.

[2] TianmeiGuo, Jiwen Dong, HenjianLi'Yunxing Gao,"*Simple Convolutional Neural Network on Image Classification",* 2017 IEEE 2nd International conference on Big Data Analysis.

[3] Gangming Zhao, Jingdong Wang , Zhaoxiang Zhang," *Random Shifting for CNN: a Solution to Reduce Information Loss in Down-Sampling Layers",* Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17).

[4] M. D. Zeiler and R. Fergus, "*Visualizing and understanding convolutional networks*," in ECCV, 2014.

[5] K. Simonyan and A. Zisserman, "*Very deep convolution networks for large-scale image recognition*," in ICLR, 2015.

[6] Ciregan,D.,U.Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification.in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. 2012. IEEE.

[7] J. Nam, M. Alghoniemy, and A. H. Tewfik, "*Audio-visual content-based violent scene characterization,*" in Proc. of IEEE International Conference on Image Processing, 1998.

[8] T. Wang, D. Wu, A. Coates, and A. Ng, "*End-to-end text recognition with Convolutionalal neural*

networks," in International Conference on Pattern Recognition (ICPR), 2012.

[9] E. B. Nievas, O. D. Suarez, G. B. Garc´ıa, and R. Sukthankar, "*Violence detection in video using computer vision techniques*," in Proc. of the 14th International Conference on Computer Analysis of Images and Patterns - Volume Part II, 2011.

[10] B. Li, W. Hu, W. Xiong, O. Wu, and W. Li, "*Horror image recognition based on emotional attention*," in Proc. of 10th Asian Conference on Computer Vision (ACCV), 2010.

AUTHOR BIOGRAPHY

Dr.SanthiBaskaran

She received her B.E. degree in CSE from University of Madras, M.Tech. and Ph.D.degree in CSEfrom Pondicherry University. She was senior lecturer in women's polytechnic college, Pondicherry during 1997-2000. Shejoined Pondicherry Engineering College in 2012 and currently working as Professor in the Department of Information Technology. Her area of specialization is Real Time Systems and Cloud Computing.

R.Pavithra

She is pursuing her B.Tech degree in the Department of Information Technology, Pondicherry Engineering College, Pondicherry.

JahnaviKorrapati

She is pursuing her B.Tech degree in the Department of Information Technology, Pondicherry Engineering College, Pondicherry.

K.Sooriya

He is pursuing his B.Tech degree in the Department of Information Technology, Pondicherry Engineering College, Pondicherry.