

Deep Learning based approach for Photographs and Painting Classification using CNN Model

Hitesh Kumar Sharma^{1*}, Tanupriya Choudhury^{2*}, Sachi Nandan Mohanty^{3†}, Shrabanee Swagatika⁴ and Satabdi Swain⁵

¹ School of Computer Science, University of Petroleum & Energy Studies (UPES), Energy Acres, Bidholi, Dehradun- 248007, Uttarakhand, India.

² School of Computer Science, University of Petroleum & Energy Studies (UPES), Energy Acres, Bidholi, Dehradun- 248007, Uttarakhand, India.

³ Adjunct Professor, Dept. of Computer Science, Singidunum University, Serbia and School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India.

⁴ Dept. of Computer Science and Engineering, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, Odisha-751030, India

⁵ Dept. of Information Technology, (Mtech IT), Utkal University, Bhubaneswar, Odisha -751004, India.

* are the corresponding authors

† authors contributed equally, and all are the first author

Abstract

One of the most powerful technologies for dealing with a wide range of real-world difficulties in the fast-paced world of the twenty-first century is Machine Learning (ML). Both regular and differently-abled persons benefit from machine learning. The Convolutional Neural Network (CNN) has been proposed for a variety of applications such as multimedia processing and so on. Here in this research paper, we have described the way and created a Binary Classification model using CNN for identifying the Paintings and Photographs. Each painting and Photograph have been warped using various procedures such as a convolutional layer, dense layer, and Flatten layers. The model is used for Binary Classifications. The wrap has been done at random on a large dataset for CNN training. We explore the architecture of CNN affects the accuracy of the identification. The proposed Model aims to increase the efficiency and accuracy of the model.

Keywords

CNN, Deep Learning, Image Processing, Classification, Machine Learning.

1. Introduction

With the advancement of machine learning algorithms [2], there has been development in the field of computer vision challenges, staying up to date on deep learning that is assisting in the quick evolution of AI that is becoming a necessity nowadays. Like how a youngster figures out how to perceive objects, we need to show a calculation and a great many pictures before it can sum up the information and make predictions for pictures.

CNN is a type of neural network that uses convolutional algorithms to CNN works by taking an image and assigning a weighting to it depending on the image's many objects. CNN [5][9] has mainly used image classifications [12], such as identifying/classifying photos vs. painting. It also has other functions, such as image segmentation and signal processing.

ACI'22: Workshop on Advances in Computation Intelligence, its Concepts & Applications at ISIC 2022, May 17-19, Savannah, United States
EMAIL: hkshitesh@gmail.com (A. 1); tanupriya1986@gmail.com (A. 2); sachinandan09@gmail.com (A. 3); shrabaneeswagatika@soa.ac.in (A. 4); satabdiswain1@gmail.com (A. 5)

ORCID: 0000-0001-6816-0324 (A. 1); 0000-0002-9826-2759 (A. 2); 0000-0002-4939-0797 (A. 3); 0000-0002-2821-4575 (A. 4); XXXX-XXXX-XXXX-XXXX (A. 5)



© 2020 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

A CNN can alternatively be constructed as a U-Net design, which consists of two nearly mirrored CNNs, in this image size output is similar to the image size input and is used in U-net architecture for image improvement and image segmentation. In astrophysics, CNNs are used to evaluate radio telescope data and forecast the most plausible visual representation of the data. One of the hottest IT topics is computer vision, which is to study the machines' ability to identify the images and videos [4]. Computer vision is utilized in the field of self-driving automobiles, robotics, and facial identification; for this the profoundly specific method known as CNN has gained monstrous headway in PC vision. CNN has been fabricated utilizing deep learning AI innovation [13].

2. Literature Review

Neural Network Architecture (NNA) was researched as a technique for picture classification [1]. The system is comprised of two arrangements of natural eye imitates just as variety grouping auto-encoding. It included a lot of convoluted photographs, yet as the review advanced, the calculation continuously further developed the MNIST models. The open-source information base MNIST is used for the preparation set. It also investigated with the dataset of Street View House Numbers, which gave a more significant result on the grounds that even natural eyes can't differentiate. The ImageNet challenge is utilized to survey the adequacy of CNN models for picture classification. Since presenting AlexNet in 2012, they've made gradual enhancements to the plan, which have expanded execution. GoogLeNet was presented in 2015 by (Szegedy et al. [14], 2015), which was an enhancement for AlexNet, generally because of a decrease in the measure of boundaries included. Also, in 2014, (Simonyan and Zisserman, 2014 [6]) presented the VGGNet, which performed well because of the organization's profundity. The image classification technique dependent on the construction of CNN was talked about in the diary (CNN). The grounding was completed by taking out the extra face pictures from the face picture information so that a specific number of face and non-facial pictures were used for preparing the research information. The picture order framework uses a bi-scale CNN with 120 prepared information and auto-stage preparing on the Face Detection Data Set and Benchmark (Fddb) to achieve an 81.6 percent location rate with just six bogus up-sides. In contrast, the current status of the craftsmanship accomplishes around an 80% discovery rate with 50 bogus up-sides.

3. Proposed CNN model for Paintings and Photograph Identification

We have investigated and motivated bottom-up strategies for increasing the classification accuracy of CNN models for the picture classification problem. We have used the Keras and TensorFlow [11] deep learning libraries [3] to execute the model in Python. This model is open to the public to use and improve.

3.1. Convolutional Neural Network (CNN)

CNN's are a sort of profound neural organization that is habitually used to assess visual information. Clinical picture examination, picture and video acknowledgment, picture arrangement, recommender frameworks [8], normal language preparation, and monetary time series are mostly regions where CNN can be utilized. A Facial Recognition framework based on CNN is a Deep Learning calculation [7] that can take an information picture, allot learnable loads and inclinations to different parts of the picture, and recognize distinctive looks. Deep learning is the innovative way for this research innovation. The deep learning concept is delegated to artificial intelligence as it can imitate the human thought in an intelligent way. Ordinarily, the framework will be pre-stacked with hundreds, if not thousands, of information to make the 'instructional meeting' more productive and faster. It starts by furnishing some sort of 'preparing' with every one of the information input.

3.2. Photographs and Painting Classification Dataset Specifications

We have used Siddesh Sambasivam's Photographs and Painting Classification Dataset on Kaggle for our project [10]. This dataset can be found on Kaggle. There are 7041 images in all, including paintings and photography. It contains 3010 images of paintings and photographs for validation. (<https://www.kaggle.com/iiplutocrat45ii/painting-vs-photograph-classification-dataset>). Because of the limited computing capacity, we have checked the correctness after 10 Epochs. It has been taken random data samples and trained our model on the 7041 train dataset before testing it on 3010. The training accuracy was observed to be 94.1 percent after 10 epochs, while the validation accuracy was found to be 90.73 percent. Table 2 represents the labels distributions in 0's and 1's, and Basic CNN Model Configuration is represented in Table 2.

Table 1

Labels distribution in 0's and 1's

Data	Label Used
Painting	00
Photographs	01

3.3. The Architecture of the Proposed CNN Model

The layered distribution architecture is shown in Figure 3. It consists of three layers such as dense, Flatten, and Convo2d layer. Model used is Inception_resNet_V2.

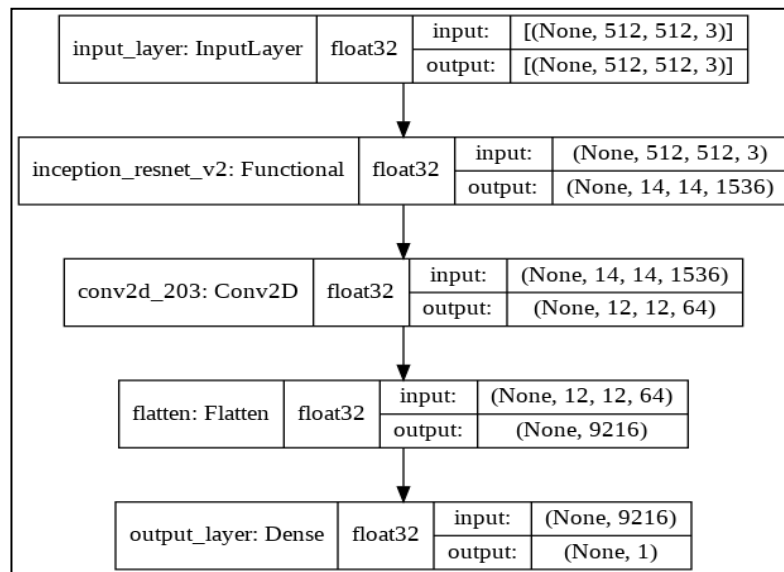


Figure 1: CNN layers for Photographs and Paintings Identification

In the Conv2d layer, the activation function that is used is ReLu (Rectified Linear initiation work), which is, in a way, a piecewise direct capacity that mainly yields positive info and returns 0 for any bad information. It is a direct numerical capacity that requires some investment to prepare and accomplishes higher exactness. This initiation work likewise supports the goal of disappearing inclination issues that are normal with other enactment capacities, for example, Sigmoid or TanH. SoftMax, otherwise called delicate argmax or standardized outstanding capacity, is one more enactment work utilized for yield. It's mostly used to change over a network's output to a likelihood distribution over projected yield classes. Figure 3 shows the layer configuration of the model.

Model: "model"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	[(None, 512, 512, 3)]	0
inception_resnet_v2 (Function)	(None, 14, 14, 1536)	54336736
conv2d_203 (Conv2D)	(None, 12, 12, 64)	884800
flatten (Flatten)	(None, 9216)	0
output_layer (Dense)	(None, 1)	9217
Total params: 55,230,753		
Trainable params: 894,017		
Non-trainable params: 54,336,736		

Figure 2: The layer configuration of the model

Table 2

Basic CNN Model Configuration

Layers	Details of Layers of Model
Model_optimization_Layer	Adam_Optimizer Layer(Layer1)
Model_loss_Layer	model_crossentropy_categorical(Layer2)
Model_metrics_Layer	['model_accuracy']
Model_Con2D_Layer	128 Model_filters, 3x3 Filter_size, ReLU Activation(Layer3)
Max-Pooling2d_Layer	2x2 size_kernel(Layer4)
Model_Dropout_Layer	20%(Layer5)
Model_Conv2D Layer	64 filters, 5x5 Filter_Size, ReLU Activation(Layer6)
Model_Max-Pooling2d_Layer	2x2 Size_kernel(Layer7)
Model_Dropout_Layer	20%(Layer8)
Model_Con2D Layer	356 model_filters, 5x5 Filter_size, Relu_Activation(Layer9)
Model_Max-Pooling2d Layer	2x2 kernel_size(Layer10)
Model_Dropout	20%(Layer11)
Model_Flatten layer	2404 Neurons(Layer12)
Model_Dense Layer	128 Neurons(Layer13)
Model_Batch Normalization	Relu_Activation(Layer14)
Model_Dropout	25%(Layer15)
Model_Dense Layer	512 Neurons(Layer16)
Model_Batch Normalization	Relu_Activation(Layer17)
Model_Dropout	20%(Layer18)
Model_Output layer	Softmax_Function 2 classes(Layer19)

4. Proposed CNN Model Enactment

The enactment of the proposed model in the research paper is defined. Kaggle provided the dataset for Photographs and Paintings. Here the image target size has been set to 512px and rescale it using the Image Data Generator. Due to a lack of resources, we have used 7041 example photos to train our model. Figure 3 shows some images that were chosen at random.

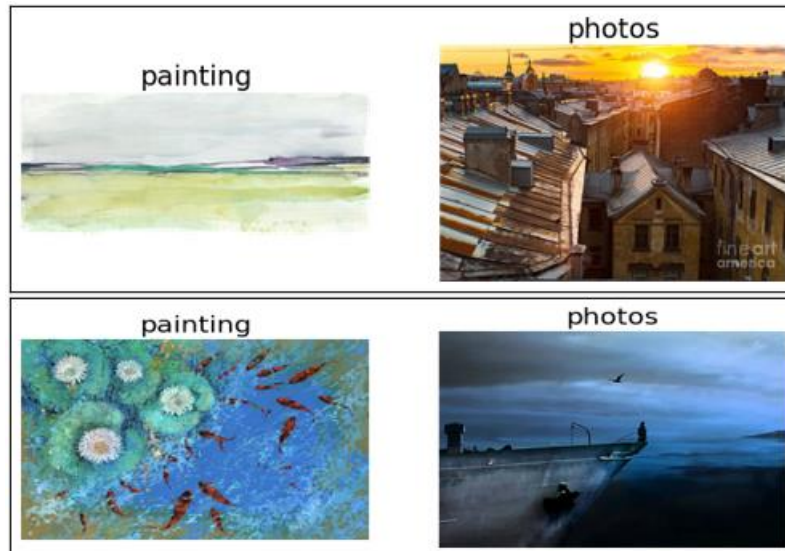


Figure 3: Random image from training datasets

5. Experimental Outcomes

The proposed model gave the outcomes that are being displayed in image form with great accuracy. The exhibition of our model is assessed utilizing the accompanying measurable boundaries.

Accuracy: Precision educates us about the rate regarding positive IDs that were genuinely correct.

Review: Recall educates us about the rate regarding real up-sides that were precisely distinguished.

Exactness is a measurement used to survey arrangement models. Exactness advises us about the rate regarding the right expectations made by our model.

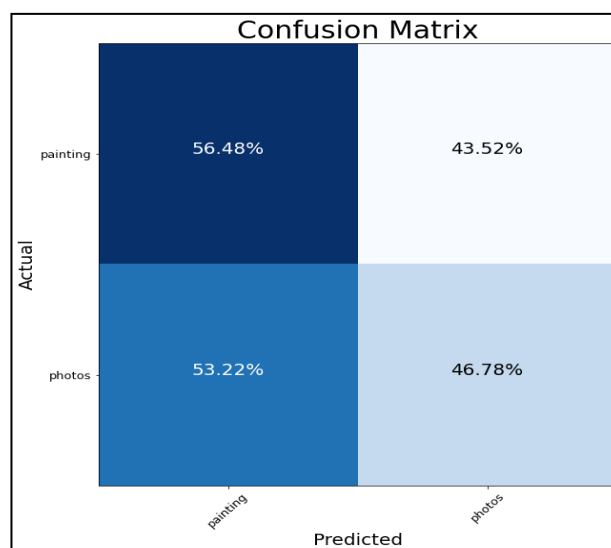


Figure 4: Accuracy of the proposed model as a Confusion matrix representation

Above Output (Figure 3) depicts the accuracy of the proposed model as a Confusion matrix representation. Actual findings are represented by vertical columns, whereas forecasted images are represented by horizontal rows.

	precision	recall	f1-score	support
painting	0.50	0.54	0.52	1505
photos	0.50	0.45	0.47	1505
accuracy			0.50	3010
macro avg	0.50	0.50	0.49	3010
weighted avg	0.50	0.50	0.49	3010

Figure 5: The Classification Report for our model, which is proposed

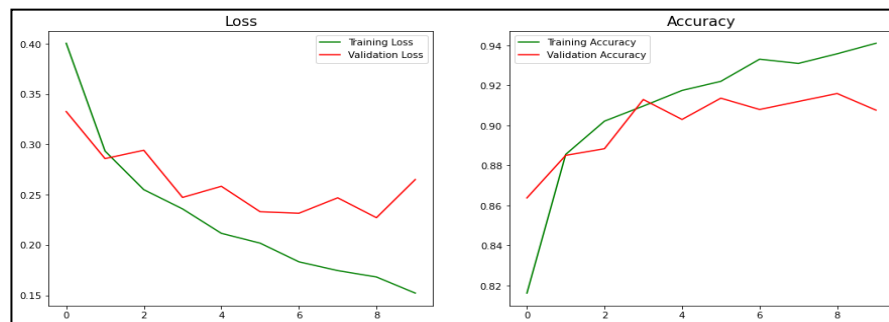


Figure 6: Accuracy and Loss Function Graph

The Accuracy (Figure 3 and Figure 3) is increasing, and loss is decreasing for both Training and validation datasets. After the 15 epochs, the accuracy will become constant, i.e., 97.1%.

6. Conclusion

In a nutshell, this research work focuses on the classification of the picture by utilizing deep learning and the TensorFlow framework. It has two main objectives and one subobjective. The objectives are indistinguishably knotted to the conclusions, which is proved by us in this research study. Firstly, it can be quantified that all of the findings gained so far have been extremely impressive. Secondly, this research also focuses on the CNN, which is mostly useful in the image categorization technologies. To justify the sub-objective, the CNN technique was further examined in depth, beginning with the assembly, then the training model, and finally, the image was classified into classes. The epochs in CNN were allowed to control the accuracy in a way avoiding issues like overfitting. Furthermore, Python was used as the programming language all through this research work as it is viable with the TensorFlow structure, which in a way takes into account the plan of the framework which is to be done all together in Python.

7. References

- [1] Krishna, M & Neelima, M & Mane, Harshali & Matcha, Venu. (2018). Image identification using neural networks. 7. 614. 10.14419/ijet.v7i2.7.10892.
- [2] Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **70**, 489–501 (2006).
- [3] Nguyen, G. *et al.* ML and DL frameworks and libraries for substantial and ample data mining: A survey. *Artif. Intell. Rev.* **52**, 77–124 (2019).
- [4] Sharma, Hitesh & Ahmed, Md & Mor, Anurag & Paul, Gaurav & Gupta, Prashasti. (2021). Deep Learning Approach for Traffic Signs Detection. 10.13140/RG.2.2.19147.11048.

- [5] A. V. R, H. F. Mahdi, T. Choudhury, T. Sarkar and B. P. Bhuyan, "Freshness Classification of Hog Plum Fruit Using Deep Learning," 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2022, pp. 1-6, doi: 10.1109/HORA55278.2022.9799897.
- [6] K Simonyan, A Zisserman, Very deep convolutional networks for large-scale image recognition- arXiv preprint arXiv:1409.1556, 2014.(<https://arxiv.org/pdf/1409.1556.pdf>)
- [7] Vivank Sharma, B. Valaramathi, K. Santhi, Shobhit Srivastava, Sumit Jahagirdar. "Performance Analysis of the Classifiers for Optical Character Recognition", 2019 International Conference on Intelligent Computing and Control Systems (ICCS), 2019.
- [8] R. Biswas et al. "A Framework for Automated Database Tuning Using Dynamic SGA Parameters and Basic Operating System Utilities", Database Systems Journal vol. III, no. 4/2012.
- [9] R Sille, T Choudhury, P Chauhan, D Sharma - , A Systematic Approach for Deep Learning Based Brain Tumor Segmentation., Ingénierie des Systèmes d'Information, 2021.
- [10] Link : <https://www.kaggle.com/iiputocrat45ii/painting-vs-photograph-classification-dataset> (Last accessed on : 1st May 2022 19:58 IST)
- [11] M. A. Abadi, P. Barham, E. Brevdo, Z. Chen, C. Citro, for example, is one of the most well-known companies in the world. Using Tensorflow, we can do large-scale machine learning on heterogeneous distributed systems.. preprint arXiv:160304467 arXiv:160304467 arXiv:160304467 arXiv:1603044 (2016).
- [12] X. Long, H. Lu, Y. Peng, X. Wang, S. Feng, Image classification based on improved VLAD. Multimedia Tools Appl. **75**(10), 5533–5555.
- [13] Mishra, M., Sarkar, T., Choudhury, T. et al. Allergen30: Detecting Food Items with Possible Allergens Using Deep Learning-Based Computer Vision. Food Anal. Methods (2022). <https://doi.org/10.1007/s12161-022-02353-9>.
- [14] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1-9