

Developing Artificial Intelligence Models for Image Classification of Cats and Dogs

Dr. Shiksha
CMRIT/ AI-DS ITPL, Bangalore, India
Email: shiksha13nov@gmail.com

Abstract— Deep learning has become popular these days for the image classification. In this paper, the most common dataset of images of dogs and cats have been analysed and implemented as there are very fewer known works for it. 2D CNN of deep learning has been implemented for image classification. A high performance of an accuracy of 94.8%, precision and recall of 95% with F1-score of 95% and MCC of 0.89 is obtained.

Keywords— image classification, deep learning, CNN.

I. INTRODUCTION

Deep learning models have shown a promising performance in the healthcare field. These models have the potential to select high level features from the input data with the help of several hidden layers. Among the deep learning models, CNNs are commonly used to process the clinical images (Ajit et al., 2020).

The cats and dogs dataset is obtained from Kaggle having 12500 images of cats and dogs, in total of 25000 images (Kaggle, 2013). Since the number for cats and dogs are same, it is a balanced classification problem, hence accuracy is used as a performance metric for the CNN model. The test data contains 12500 test records without any label for the images. The CNN model is trained on the training dataset and evaluated on the validation data and labels are predicted for the test data (Niu et al., 2019).

A. CNN model

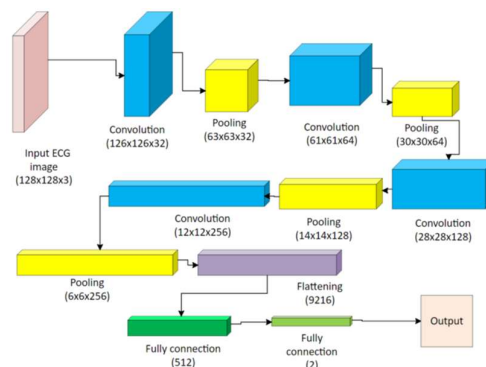


Figure 1. CNN model used in this research

CNN has the advantage of extracting features automatically through convolution and pooling compared to other discussed machine learning techniques, hence no need to extract features manually (Burger et al., 2020; Ajitesh Kumar, 2023; datacamp, 2024). The final CNN model used in this experiment for ECG images dataset shown in Figure 1 consists of 4 convolutional layers and 4 pooling layers. The convolution kernel of each convolution layer has same size of 3x3 and different numbers of 32, 64, 128 and 256 for different convolution layers. The size of the pooling layer is 2x2. The output from the convolutional layers is connected to a fully connected layer having 128 neurons. ReLU is used as the activation function for this fully connected layer. The output layer having 2 neurons and softmax as the activation function generates the final output as cats or dogs. The total number of parameters for this model is 5110017.

B. Experiment settings

CNNs are complex because of depth and parameters, it can be seen from Figure 5.3 that a huge number of trainable parameters is there for cats and dogs dataset, hence it requires GPU to accelerate training of CNN. On the windows 10 laptop of CPU with RAM 16GB, CNN implementation for this dataset was not working, hence GPU was required for it. The LJMU VPN “Big-IP Edge Client” was installed on the laptop and LJMU GPU “sad.cms.livjm.ac.uk” was enabled which was taking around 40-50 minutes for implementing CNN code. Different experiments have been conducted to obtain the best value for different layers of CNN in Python, programming language. OpenCV is used to work with images and CNN is implemented with TensorFlow. For training CNN, the training epoch is kept at 15 at start and varied later with a batch size of 32 with the performance metric as accuracy.

II. RESULTS

The loss and accuracy have been evaluated during training with different parameter settings for cats and dogs dataset. Figure 2 shows the changes in loss and accuracy for training and validation dataset for the best case. The validation accuracy in Figure 3. shows great fluctuation since only 5000 images have been used for validation. It can be seen from Figure 5.4b that validation accuracy was 83% for 5th epoch which has increased to 94% at 30th epoch. There is not much difference between training and validation accuracy. Figure 2 shows the loss has reduced more for 30 epochs compared to 5 epochs.

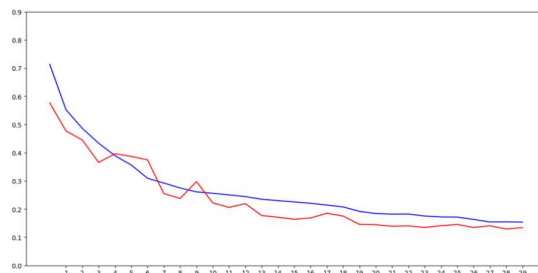


Figure 2. The change of loss for cats and dogs dataset.

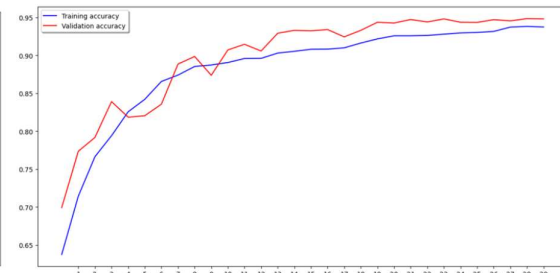


Figure 3. The change of accuracy for cats and dogs dataset.

The accuracy for the cats and dogs dataset was around 75% using 3 convolution layers and 3 pooling layers using 30 epochs. The CNN was trained for 5 epochs till 50 epochs. The best result is provided by 30 epochs. The best accuracy obtained by this model 94.8%.

Figure 4. shows the classification report and the confusion matrix for the best obtained CNN model for cats and dogs dataset. It can be seen that the precision and recall for identifying the cats and dogs are almost equal around 95% for this model. The confusion matrix in Figure 4b shows that FP cases are 113 and FN cases are 152, and total 4735 (2415+2320) images have been predicted correctly giving accuracy of 94%. The best hyper-tuned model achieved an MCC score of 0.89, indicating accurate predictions.

	Precision	Recall	F1-score	Support
Class 0	0.94	0.96	0.95	2528
Class 1	0.95	0.94	0.95	2472
Weighted avg	0.95	0.95	0.95	5000

Figure 4 (a) Classification Report

Confusion Matrix

True labels	cats	dogs
	cats	dogs
cats	2415	113
dogs	152	2320
	cats	dogs
	Predicted labels	

Figure 4 (b) Confusion matrix for cats and dogs dataset

The trained model is applied to the test data to predict the labels of the test samples. Fig. 5 displays the predicted labels for several images from the test data. It is evident that the best hyper-tuned model accurately identifies images compared to other models. Figs. 5a and 5c show the identification of test images by the initial CNN model, whereas Figs. 5b and 5d show the identification by the best hyper-tuned CNN model. It can be observed from Figs. 5a and 5c that the initial CNN model failed to correctly recognise the dog and cat images, which were accurately classified by the best model, as shown in Figs. 5b and 5d.

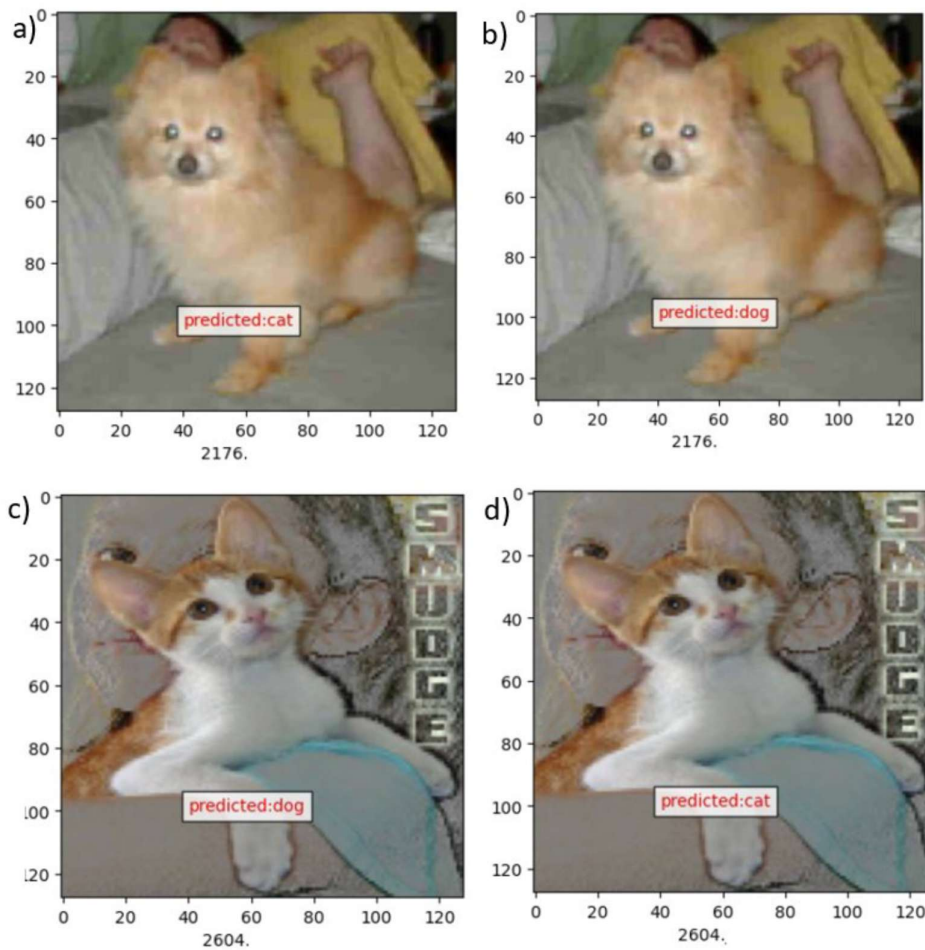


Figure 5 a, c) Predicted label from initial CNN model b, d) Predicted label from best hyper tuned CNN model.

Table 1 compares the results obtained from our research on the cat and dog dataset with those from other published works on the same dataset. It is evident from the table that our work includes various performance metrics for this simpler dataset. The study (Chen, 2020) utilised ResNet for identifying cat and dog images; however, its accuracy is lower than that of our model. Our implemented model has yielded better results compared to other published works.

TABLE I. COMPARISON OF THE OBTAINED RESULT WITH OTHER REFERENCE WORKS FOR CAT AND DOG DATASET

Reference Work	Accuracy	Precision	Recall	F1 score	MCC
(Tushar Jajodia and Pankaj Garg, 2014)	90.1%	Not mentioned	Not mentioned	Not mentioned	Not mentioned
(Chen, 2020)	92%	Not mentioned	Not mentioned	Not mentioned	Not mentioned
Our work	94.8%	95%	95%	95%	0.89

III. CONCLUSION

In this research paper, a 2D CNN has been developed and implemented for the image classification of cats and dogs dataset. Not only accuracy, but different performance metrics have also been obtained. This research has obtained a high performance of an accuracy of 94.8%, precision and recall of 95% with F1-score of 95% and MCC of 0.89 is obtained.

REFERENCES

- [1] Ajit, A., Acharya, K. and Samanta, A., (2020) A Review of Convolutional Neural Networks. International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2020, pp.1–5.
- [2] Ajitesh Kumar, (2023) Different Types of CNN Architectures Explained: Examples. [online] Data Analytics. Available at: <https://vitalflux.com/different-types-of-cnn-architectures-explained-examples/>.
- [3] Burger, A., Qian, C., Schiele, G. and Helms, D., (2020) An Embedded CNN Implementation for On-Device ECG Analysis. 2020 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2020.
- [4] datacamp, (n.d.) A Complete Guide to Data Augmentation. [online] datacamp. Available at: <https://www.datacamp.com/tutorial/complete-guide-data-augmentation>.
- [5] Kaggle, (2013) Dogs vs. Cats. [online] Kaggle. Available at: <https://www.kaggle.com/c/dogs-vs-cats/data>.
- [6] geeksforgeeks, (2023) Introduction to Convolution Neural Network. [online] geeksforgeeks. Available at: <https://www.geeksforgeeks.org/introduction-convolution-neural-network/>.
- [7] geeksforgeeks, (n.d.) CLAHE Histogram Equalization – OpenCV. [online] GeeksforGeeks. Available at: <https://www.geeksforgeeks.org/clahe-histogram-equalization-opencv/>.
- [8] Hayati, M., Muchtar, K., Roslidar, Maulina, N., Syamsuddin, I., Elwirehardja, G.N. and
- [9] Hubel, D.H. and N., T.W., (1959) RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX. J. Physiol, pp.574–591.
- [10] Kiranyaz, S., Ince, T., Hamila, R. and Gabbouj, M., (2015) Convolutional Neural Networks for patient-specific ECG classification. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2015-NovemAugust, pp.2608–2611.
- [11] Lee, Y., (2021) Image classification with artificial intelligence: Cats vs dogs. Proceedings - 2021 2nd International Conference on Computing and Data Science, CDS 2021, 80, pp.437–441.
- [12] Manav Mandal, (2021) Introduction to Convolutional Neural Networks (CNN). [online] Analytics Vidhya. Available at: <https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/>.
- [13] Marimuthu, P., (2022) Image Contrast Enhancement Using CLAHE. [online] Analytics Vidhya. Available at: [https://www.analyticsvidhya.com/blog/2022/08/image-contrast-enhancement-using-clahe/#:~:text=Contrast Limited AHE \(CLAHE\) is,high accuracy and contrast limiting](https://www.analyticsvidhya.com/blog/2022/08/image-contrast-enhancement-using-clahe/#:~:text=Contrast Limited AHE (CLAHE) is,high accuracy and contrast limiting).
- [14] Nazia Aslam, (2022) Training a CNN from Scratch using Data Augmentation. [online] Analytics Vidhya. Available at: <https://www.analyticsvidhya.com/blog/2022/09/training-a-cnn-from-scratch-using-data-augmentation/>.
- [15] Open Data Science, (2019) Image Augmentation for Convolutional Neural Networks. [online] Medium. Available at: <https://odsc.medium.com/image-augmentation-for-convolutional-neural-networks-18319e1291c>.
- [16] Phani Ratan, (2021) What is the Convolutional Neural Network Architecture? [online] Analytics Vidhya. Available at: <https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/>.
- [17] Sagar Biswas, (2021) Pooling in Neural Networks. [online] Medium. Available at: <https://biswas-sagar97.medium.com/pooling-in-neural-networks-c783d9c3a35c>.