The Pharma Innovation Journal 2023; 12(2): 3643-3646

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2023; 12(2): 3643-3646

www.thepharmajournal.com
Received: 17-11-2022
Accepted: 27-12-2022

**Ambreen Hamadani**
National Institute of Technology Srinagar, Jammu and Kashmir, India

**Nazir A Ganai**
SKUAST-Kashmir, Jammu, Union Territory of Jammu and Kashmir, India

**Ayush Rajan Gupta**
National Institute of Technology Srinagar, Jammu and Kashmir, India

**Faizan Farooq Shah**
Institute of Hotel Management, Catering Technology & Applied Nutrition, Srinagar, Jammu & Kashmir, India

**Mohd Shakeb Umar**
Bristlecone India

**Kartik Jindal**
National Institute of Technology Srinagar, Jammu and Kashmir, India

**Saurabh Jajoo**
Equity Portfolio Manager, India

**Harsh Gopalika**
Birla Institute of Technology & Science, Pilani is a deemed university in Pilani, Jhunjhunu, Rajasthan, India

**J Bashir**
National Institute of Technology Srinagar, Jammu and Kashmir, India

**Corresponding Author:**
**Ambreen Hamadani**
National Institute of Technology Srinagar, Jammu and Kashmir, India

# Comparison of multiple deep convolutional neural networks for image classification

**Ambreen Hamadani, Nazir A Ganai, Ayush Rajan Gupta, Faizan Farooq Shah, Mohd Shakeb Umar, Kartik Jindal, Saurabh Jajoo, Harsh Gopalika and J Bashir**

**Abstract**
An attempt was made to train several deep convolutional neural networks to classify images from the Image Detect challenge into 200 distinct classes. Multiple CNN architectures were tried which included pretrained modes with feature extraction as well as development of models from scratch. Our models consisted of 3- 5 convolutional layers, 3 max-pooling layers, 2-3 fully connected layers, and a Soft Max classification layer. Among the multiple transfer learning techniques used, InceptionResNetV2 showed the most promising results with the highest accuracy and lowest loss. The test predictions were 56% accurate and our model was among the top 13 models among 39 submissions (top 36%) in the Image classification task.

**Keywords:** Image classification, models, pretrained models

## 1. Introduction

The area of deep learning is both challenging and intriguing and over the past few years it has found wide applicability. Image classification essentially is the methodology which assigns class labels from a fixed set of categories to each image given as input. This supervised learning technique defines a set of target classes (objects to identify in images) and trains a model to recognize them using labelled example photos.

This seemingly simple classification technique has a wide range of practical applications and image classification is in vogue in all spheres of learning. This technique has the potential to transform lives and have a great impact worldwide. Due to this, every research in the area of computer vision is a steppingstone towards a greater global mission of achieving the 4th industrial revolution. This machine learning techniques can potentially transform all major spheres of life, from self-driving cars to the creation of visual tools to the differential diagnosis of disease.

Image Classification challenge was taken up for the Capstone project considering all this as well as its applicability in our rapidly transforming world. An attempt was made to classify the models using different models using different approaches for achieving a reasonable accuracy of the validation dataset and subsequently the test dataset.

Since its ideation, one of the main issues in computer vision has been image classification, which is concerned with determining the presence of visual structures in an input image [1]. The fascination with image classification has led scientists and enthusiasts alike to contribute to this arena of deep learning. The MNIST digit-recognition is currently the gold standard for learning image classification, and it also approaches the best error rate [1].

Over the years, multiple state-of -the-art and innovative algorithms were developed based on a popular dataset called the ImageNet. These algorithms offer varying amounts of accuracy upon the data and are also being used for transfer learning. The first deep learning model was developed in 2012 [2] which achieved an accuracy of 26.2% using a SIFT model. This model is popularly known as AlexNet. The architecture of AlexNet consists of five convolutional filters, max-pool layers and three fully connected layers. Krizhevsky *et al*. trained a large, deep convolutional neural network to classify 1.2 million high-resolution images into 1000 different classes and achieved an error rate of 15.3% [2]. Simonyan and Zisserman demonstrated the VGG16 model, which was made of 16 convolutional layers, many max-pool layers and 3 final fully connected layers [3]. They also introduced 3x3 filters for each convolution (as opposed to 11x11 filters for the AlexNet model) This significantly decreased the number of parameters during training. Lin *et al*. developed the concept of inception modules. Subsequently Inception V2 model and Inception V3 model we introduced [4]. They promised a higher precision. The model viz, ResNet, is composed of 152 convolutional layers with 3x3 filters using residual learning by block of two layers. Residual Learning created a connection between the output of one or multiple convolutional layers and their original input with an identity mapping. Ren *et al*., 2015 also introduced Region Proposal Network (RPN) that shares full-image convolutional features with the detection network while achieving state-of-the-art object detection accuracy on PASCAL VOC 2007 (73.2% mAP) and 2012 (70.4% mAP) using 300 proposals per image. He *et al*. (2015) combined inception modules and residual blocks into residual inception blocks. The inception modules have since been upgraded. Zoph *et al*. created a model with an architecture called the NASNet model [5].

The area of artificial intelligence is ever evolving as more and more developments occur, the field advances magnificently [19; 20; 21; 22]. These image recognition models are seeping into diverse areas of science and are transforming lives which in turn has a direct impact on the economy as well.

## 2. Materials and Methods

### 2.1. Data Set

Image Detect Challenge consisted of a dataset images belonging to 200 object classes. Each class had 450 training images, 50 validation images, and 50 test images. Class labels and bounding boxes were provided as annotations and the test classes had to be predicted and since it had no labels, the test accuracy could not be found.

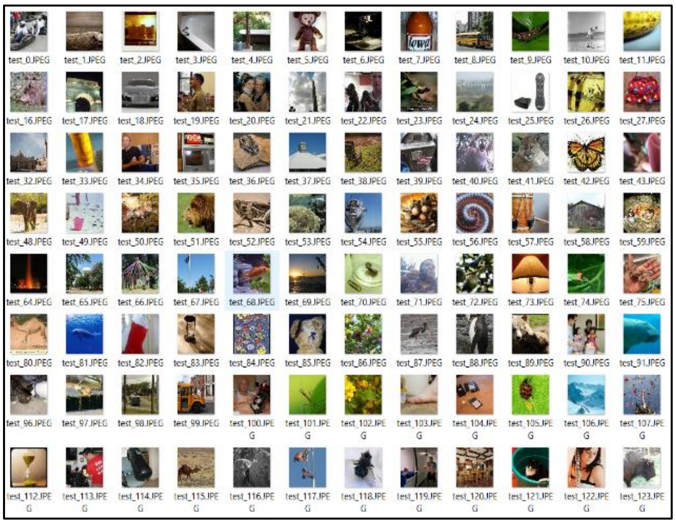

**Fig 1:** Random images within the dataset.

The dataset was downloaded into the local machine and data was extracted into the Jupyter notebook (Anaconda) and run on python version 3.7.7. This data was visualized using various libraries viz *os, pandas, NumPy, seaborn, matplotlib etc*. For the image classification task, multiple libraries were used which primarily included TensorFlow (CUDA enabled GPU version) from google [6]. Figure 1 depicts some of the random images within the dataset.

### 2.2 Building a model from scratch

There is a myriad of approaches to build an acceptable architecture for image classification. To build the most appropriate model for the classification task multiple models were tested for the highest accuracy. A total of 5 models using different numbers of multilayer Convolutional Neural Networks (Convolutional Layers, pooling layers as well as fully connected layers) using (Keras with TensorFlow backend) were built to run the train upon the train and validation dataset.

Given the complexity of the task at hand and multiple possibilities of improving the prediction, multiple models were tried. These were developed without any pretrained models. The Convolutional Neural Networks had the following architecture:

ReLu was used as a non-linearity which improves the network's learning dynamics, and this reduces the number of iterations and reduces computational budget. The Kernel size taken was 3x3; Pool size: 2x2; Activation function in the classification layer: SoftMax; Optimizers: Adam and Binary Cross entropy

Conv 2D layers varied from 3 to 5, Pooling layers: 3; Dense layers: 3 (400 neurons) and 2 (2156 and 200 neurons). A Dropout of 0.1 (one model only) was also tried. Care was taken to prevent overfitting by using a validation dataset for monitoring validation loss. We observed that as the number of the convolutional layers increased larger amounts of memory, so computational constraints limit the depth of the models. Max-pooling (pool) layers were useful in reducing the number of parameters in the network by the reduction of spatial size.

Upon experimenting with different numbers of pooling layers, we got many errors and therefore by heuristics, we settled for 3. The height and width of the convolutional filters are also important for the CNN architecture. These allow the CNN to build powerful input representations with fewer parameters we used a filter size of 3x3 so as to balance the complexity as well as to try to extract maximum information from the CNNs. Batch sizes of 32, 50, 60, 75 and 100 were were tried in our models. Beyond 100, the model became difficult to compute. Various learning rates viz: 0.001, 0.005, 0.009, 0.0001, 0.0005, 0.00005, 0.00009 were tried as well.

### 2.3 Transfer Learning

Realizing the difficulty and well as the computational expense involved in developing the networks from scratch in terms of computational power, we attempted to train our dataset on pre-trained models using transfer learning as well. This is an effective method for them to re-use the model weights from pre-trained models which were specifically developed and trained on huge datasets for computer vision [7].

Feature extraction was performed on the said model to customize it as per our requirements. Multiple pretrained models were tried for the purpose to see which one would suit our needs best. We experimented with ResNet101V2, ResNet152V2, EfficientNetB7, ResNet50, VGG19, InceptionResNetV2 and EfficientNetB7 Hyperparameter tuning was performed for each model and each model was run multiple times.

### 2.4 Final Model

We selected InceptionResNetV2 as the final model as it gave the highest accuracy among all the models tested. InceptionResNetV2 is a very deep network which makes use of a very deep network along with residual weights. With an ensemble of 3 residual and 1 Inception-v4, this network achieved 3.08 percent top-5 error during the on the ImageNet classification (CLS) challenge [8].

### 2.5 Initialization and Optimization

The ImageDataGenerator function was used to generate the image data. flow_from_directory was used to generate training images. For the validation images, flow_from_dataframe was used.

For the final model selected, InceptionResNetV2, include_top layer was given as false and the input image shape was resized to 200, 200. This was heuristically found to give the most accurate validation. Other sizes that were tried were 64, 100, 128. The optimizer was taken as Adam with a learning rate of .00005 which was heuristically found out to be most optimum (RMSprop was also tried). Categorical cross entropy was used as a loss function however binary cross entropy was also tried. Accuracy metric was used for the evaluation. The batch size was taken to be 32 (Batch sizes of 30, 50, 60 and 100 were also tried). The number of epochs were set to a 100 however early stopping mechanism was employed with a patience of 3 (2 was also tried).

All our models were able to generalize well and overfitting did not appear to be an issue. Looking at the validation loss and accuracy of the model, we realize that there is scope for improvement of both. However limited computational power is a major hurdle in training models of high complexity. Our models on an average five hours to train on a core i7-9750H GPU with NVIDIA GTX (4GB) GPU.

The model was saved, and the test images were run into the model to obtain predictions in a text file which were uploaded to Kaggle for obtaining test predictions and team rank.

## 3. Results and Discussion

Multiple results in terms of validation accuracy as well as the validation loss were obtained while testing the various models developed for the image classification task. All models were developed using convolutional neural networks (CNN) and were tested using the validation dataset for the purpose of getting optimal results from our model in terms of the validation loss and validation accuracy [9-10]. The final model was tested using test dataset as well.

We realized that shallow models give very little accuracy and deeper models are very hard to train both in terms of time as well as computational power. Among all the CNNs tried, InceptionResNetV2 gave the highest validation accuracy, lowest loss and the results were replicable multiple times.

The metrics for the models created from scratch ranged from 26.31 to 36% for validation accuracy. Validation loss varied from 3.45-2.9. For the models using transfer learning the results are summarized in Table 1.

For our final model selected, the validation accuracy was 0.5639 and loss 1.8791 (Figures 2 and 3). The test predictions were 56% and our model was among the top 13 models among 39 submissions (top 36%).

This network architecture has been shown to achieve has been shown to achieve good performance with low computational costs as was seen in our study too [12]. The use of deep learning techniques has been reported to improve classification accuracy of images by a number of authors for various fields like object-based land-cover image classification [13], Hyperspectral Image Classification [14], diabetic retinopathy detection [15], Breast Cancer Image Classification [16] etc. In fact convolutional Neural Networks are supposed to work very well for the later study [16]. Some researchers however have also reported decision tree classifier, and knowledge-based classification as important approaches for multisource data classification [17].

We infer that InceptionResNetV2 may be used for similar classification problems and that the accuracy may further be improved by using multiple networks for the same problem.

**Table 1:** Results of Transfer Learning (*Final Model)

| CNN | Validation Accuracy | Validation Loss |
|---|---|---|
| ResNet101V2 | 0.2133<br>0.5376<br>0.4953 | 3.9823<br>2.3041<br>2.3575 |
| ResNet152V2 | 0.5059 | 2.4791 |
| EfficientNetB7 | 0.05 | 5.123 |
| ResNet50 | 0.0378 | 5.0723 |
| VGG19 | 0.2441 | 3.4286 |
| InceptionResNetV2 | 0.4768<br>0.5200<br>0.5265<br>0.5639* | 2.7077<br>2.1028<br>2.0543<br>1.8791* |
| EfficientNetB7 | 0.050 | 5.03 |

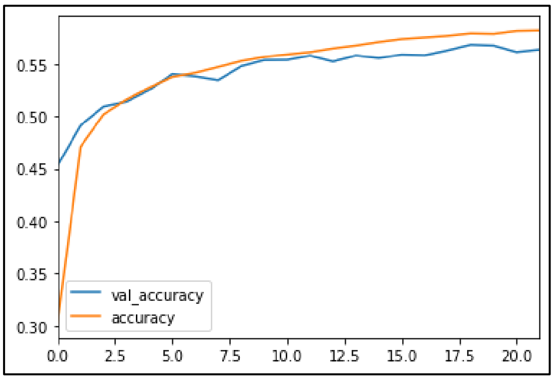


**Fig 2:** Plot of Training loss and validation loss for the final model: InceptionResNetV2

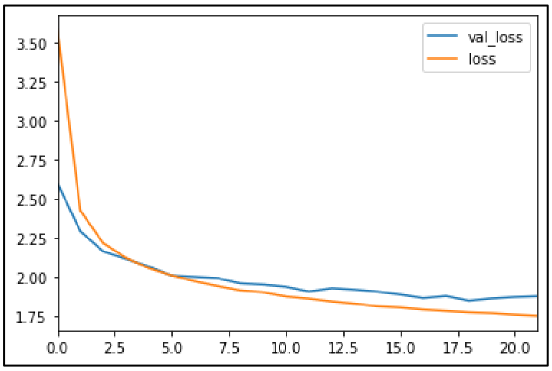


**Fig 3:** Plot of Training accuracy and validation accuracy for the final model: InceptionResNetV2

## 4. Conclusion

Of all the models trained for the present image classification task, we conclude that feature extraction with InceptionResNetV2 gave the highest validation accuracy, lowest loss and the results were replicable. This net can therefore be used for classification of similar images thereby for solving a number of problems in image classification.

## 5. Acknowledgement

The authors are grateful to IIT Kanpur, India for providing the dataset for the image classification.

**6. References**


1. Gavali P, Banu S. Deep Convolutional Neural Network for Image Classification on CUDA Platform", (Ed: Arun Kumar Sangaiah) Deep Learning and Parallel Computing Environment for Bioengineering Systems, Academic Press. doi.org/10.1016/B978-0-12-816718-2.00013-0, 1995.
2. Cire san D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification, Arxiv preprint arXiv:1202.2745, 2012.
3. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In NIPS, 2012
4. Simonyan K and Zisserman A. Very Deep Convolutional Networks for Large-Scale. Image Recognition. arXiv. https://arxiv.org/abs/1409.1556, 2014
5. Lin M, Chen Q, Yan S. Network in Network. Corr. abs/1312.44, 2014
6. Zoph B, Vasudevan V, Shlens J, Le QV. Learning Transferable Architectures for Scalable Image Recognition, arXiv. https://arxiv.org/abs/1707.07012, 2017
7. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, *et al*. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org., 2015
8. Brownlee J. How to Use Greedy Layer-Wise Pretraining in Deep Learning Neural Networks. https://machinelearningmastery.com/greedy-layer-wise-pretraining-tutorial/, 2020
9. Szegedy C, Ioffe S, Vanhoucke S, Alemi A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI Publications, Thirty-First AAAI Conference on Artificial Intelligence, 2017
10. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, 770-778, doi: 10.1109/CVPR.2016.90
11. Ren HK, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. In: NIPS, 2015
12. Szegedy C, Ioffe S, Vanhoucke V, Alex A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv. 2016, 1602.07261
13. Ma L, Manchun L, Xiaoxue M, Liang C, Peijun D, Yongxue L. A review of supervised object-based land-cover image classification, ISPRS Journal of Photogrammetry and Remote Sensing, 2017;30:277-293. https://doi.org/10.1016/j.isprsjprs.2017.06.001.
14. Wu T, Hu W, Huang Y, Wei L, Zhang F, Li H. Deep Convolutional Neural Networks for Hyperspectral Image Classification. Deep Convolutional Neural Networks for Hyperspectral Image Classification Journal of Sensors 2015;2(15):1687-725X
15. Shaohua W, Yan L, Yin. Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers & Electrical Engineering. 2018;72:274-282.
16. Nahid A, Kong Y. Involvement of Machine Learning for Breast Cancer Image Classification: A Survey, Computational and Mathematical Methods in Medicine. 2017, 29. https://doi.org/10.1155/2017/3781951,
17. Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance, International Journal of Remote Sensing. 2007;28:5. DOI: 10.1080/01431160600746456, 2007
18. Hamadani A, Ganai NA, Andrabi SM, Shanaz S, Alam S. Evaluation of Artificial Intelligence Algorithms for the Prediction of Genetic Merit. Preprint, 2021. https://doi.org/10.21203/rs.3.rs-1488946/v1
19. Onyijen OH, Hamadani A, Awojide S, Ebhohimen IE. Prediction of deaths from covid-19 in Nigeria using various machine learning algorithms. Sau science-tech journal. 2021;6(1):109-117.
20. Hamadani A, Ganai NA. Development of a multi-use decision support system for scientific management and breeding of sheep, Scientific Reports. 2022;12(1):19. DOI: 10.1038/s41598-022- 379 24091-y
21. Hamadani A, Ganai NA, Mudasir S, Shanaz S, Alam S, Hussan I. Comparison of artificial intelligence algorithms and their ranking for the prediction of genetic merit in sheep. Sci Rep. 2022;12:18726. https://doi.org/10.1038/s41598-022-23499-w