Performance Comparison of Deep Learning Models for Natural Scene Images Classification

Abdulrahman Atif
dept. Computer Science
MSA
Cairo, Egypt
abdulrahman.atif@msa.edu.eg

Abdulrahman Salah
dept. Computer Science
MSA
Cairo, Egypt
abdulrahman.ahmed5@msa.edu.eg

Assem Omar
dept. Computer Science
MSA
Cairo, Egypt
asem.omar2@msa.edu.eg

Youssef Qotb
dept. Computer Science
MSA
Cairo, Egypt
youssef.mohamed81@msa.edu.eg

***Abstract*—In this study various image classification models will be evaluated on Intel Image Classification dataset in terms of their effectiveness especially DenseNet121, CNN, ResNet50V2, VGG16, Inception regarding accuracy, precision; recall; F1-score. was found from our experimentations that DenseNet121 gave the highest level of efficiency among them all."**

I. Introduction

From medical diagnosis to autonomous driving, Image classification plays a crucial role in computer vision. Significant improvements in accuracy and efficiency of image classification tasks have been achieved thanks to advancements in deep learning recently. By specifically using Intel Image Classification data set, this study evaluates different contemporary image classification models and contrasts their performances. The contributions that stem from this study involve a coherent introduction to issues such as object detection, segmentation, and localization in computer vision applications or retinal imaging techniques like RetCam systems among others.

II. Related Work

Various deep learning models have been explored in image classification domain but Rout et al.(2017) proposed a multi-label image classification system using convolutional neural networks (CNNs) for feature extraction and classification. A such kind of method performed better than the conventional machine learning algorithms. They showed how powerful deep learning is when it comes to categorizing images which are hard to understand and that have many different types of signs and things. There have been other similar investigations.

III. Data Description and Preprocessing Steps

The Intel Image Classification dataset comprises images categorized into six classes: buildings, forest, glacier, mountain, sea, and street [**?**]. The following preprocessing steps were performed:

- The dataset was extracted and divided into training and testing sets.
- Each image was resized to 150x150 pixels.
- Data augmentation techniques like, rescaling, zooming, and horizontal flipping were applied using the ImageDataGenerator class.
- The dataset was split into training, validation, and test sets with appropriate labels.

The dataset was visualized to ensure correctness, and class distributions were checked to identify any imbalances.

IV. Methodology/Approach

This section describes the models used for image classification:

A. DenseNet121

DenseNet121 connects every layer to every other layer through a feed-forward in order to enhance the flow of information and gradients in the network hence resulting in high efficiency of the model.

B. CNN

A basic Convolutional Neural Network (CNN) architecture was implemented, consisting of multiple convolutional and pooling layers followed by dense layers. This model is known for its simplicity and effectiveness in image classification tasks.

C. ResNet50V2

ResNet50V2 is a variant of the Residual Network (ResNet) architecture, which uses residual connections to allow gradients to flow more easily through deeper networks, thus mitigating the vanishing gradient problem.

D. VGG16

VGG16 is a deep network with 16 layers, known for its simplicity and depth, which makes it a strong baseline model for image classification tasks.

E. Inception

The Inception architecture employs a network-in-network approach, allowing for more efficient computation and better utilization of model parameters. It uses multiple convolutional filters of different sizes, concatenated along the channel dimension.

TABLE I
PERFORMANCE COMPARISON OF CLASSIFIERS

Model	Accuracy	Precision	Recall	F1-Score
DenseNet121	0.9083	0.91	0.91	0.91
CNN	0.8633	0.86	0.86	0.86
ResNet50V2	0.9053	0.91	0.91	0.90
VGG16	0.8680	0.87	0.87	0.87
Inception	0.8913	0.89	0.89	0.89

V. RESULTS

The performance of each model was evaluated using accuracy, precision, recall, and F1-score. Table **??** summarizes the results.

DenseNet121 with minimal differences crowned as the best-performing model with a balanced performance across all metrics.

VI. CONCLUSION

This study demonstrates that DenseNet121 is highly effective for the Intel Image Classification dataset, outperforming other models in terms of accuracy, precision, recall, and F1-score. The insights gained highlight the importance of model architecture in achieving high performance in image classification tasks.

REFERENCES

[1] M. Rout, S. Nandi, S. Biswas, and A. B. Roy, "An investigation on the impact of data augmentation and transfer learning in biological image classification," in *Proceedings of the 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Manchester, UK, 2017, pp. 1-7.

[2] P. Puneet, "Intel Image Classification," *Kaggle*, 2019. [Online]. Available: https://www.kaggle.com/datasets/puneet6060/intel-image-classification. [Accessed: 20-May-2024].

[3] Flygirl, "Intel Image Classification[90

[4] Vincee, "Intel Image Classification (CNN - Keras)," *Kaggle*, 2022. [Online]. Available: https://www.kaggle.com/code/vincee/intel-image-classification-cnn-keras. [Accessed: 20-May-2024].