COMP 6721: Applied Artificial Intelligence

Animal Classification Using CNN Architectures

Team :- B

Outline

- Problem Definition
- Datasets
- Training CNN Architectures
- Comparisons and Results
- Conclusion
- References

- Efficient and reliable monitoring of wild animals in their Natural habitats is essential to inform conservation and management decisions regarding wildlife species, migration patterns, and habitat protection, and is possible, rehabilitation and grouping species of the same animals together. Processing a large volume of images and videos captured from camera traps manually is extremely expensive, time consuming, and monotonous. This presents a major obstacle to scientists and ecologists to monitor wildlife in an open environment.

Problem Definition

Datasets

- The first dataset, called Dataset 1, contains approximately 25,000 RGB images of dogs and cats with varying pixel resolutions, ranging from 320 x 200 to 498 x 480 pixels. Originally created by Microsoft Research for a Kaggle competition in 2013.
- The second dataset, Dataset 2, was provided by a user named Saumil Agrawal in 2018. This dataset contains around 15,100 images of 25 different animal classes, with a fixed pixel resolution of 1280 x 720 pixels.

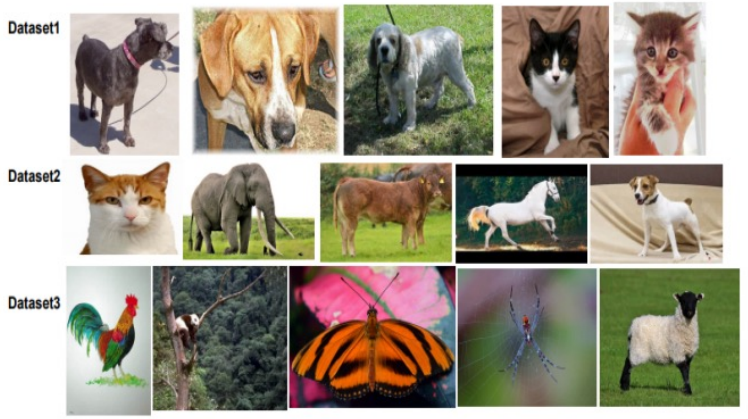

Figure 1.1

Datasets

- The third dataset, Dataset 3, was created by Kaggle user Piyush Kumar in 2019. It contains approximately 17,200 images of various animals, including butterfly, cats, cows, elephants, hens, horses, monkeys, pandas, sheep, spiders, squirrels, among others. Size of the images are ranging from 201 x 300 to 500 x 374 pixels.
- The dataset is divided into train validation and test sets in an 80:10:10 ratio, respectively. The images were randomly shuffled before the split. Test data is only used for evaluating the model.

Figure 1.2

Training CNN Architectures

- Architectures in our case are ResNet18, ShuffleNetV2, MobileNetV2.
- **Input Size** : The input size refers to the dimensions of the input data that is fed into CNN architectures for processing. Input size for all the architectures are 224 x 224 x 3.
- **Hyper Parameter Tuning :** In order to improve performance of above-mentioned CNN architectures process of selecting best hyperparameters for a machine learning model.
- Selected hyperparameters for animal image classification are
  - Batch Size = 32
  - Epochs = 30
  - Learning Rate = 0.001
- **Transfer Learning** : Used pre-trained model on IMAGENET1KV1.

  Using deep tuning and Fine tuning.
- **Optimization**:The choice of optimization algorithm depends on the specific problem and the nature of the data. We have used Adam optimizer for optimization.

Comparisons and Results

- Dataset wise Accuracy comparison
- Dataset wise Training time comparison
- Training loss
- Training Accuracy
- Confusion Matrix
- TSNE

Architecture	Dataset	Testing Acc.	Precision	Recall	F1-Scores	Training time
MobileNetV2	Dataset1	87.60	88	88	87	36 min
	Dataset2	80.22	80	81	80	2 hour
	Dataset3	73.69	75	74	74	5 hour
ShuffleNetV2	Dataset1	82	82	82	82	20 min
	Dataset2	79.82	82	79	79	1.2 hour
	Dataset3	72.98	72	73	72	1.3 hour
ResNet18	Dataset1	75.40	76	75	75	21 min
	Dataset2	61.71	67	56	61	1.1 hour
	Dataset3	53.50	52	54	53	1.3 hour

Figure 1.3

Results

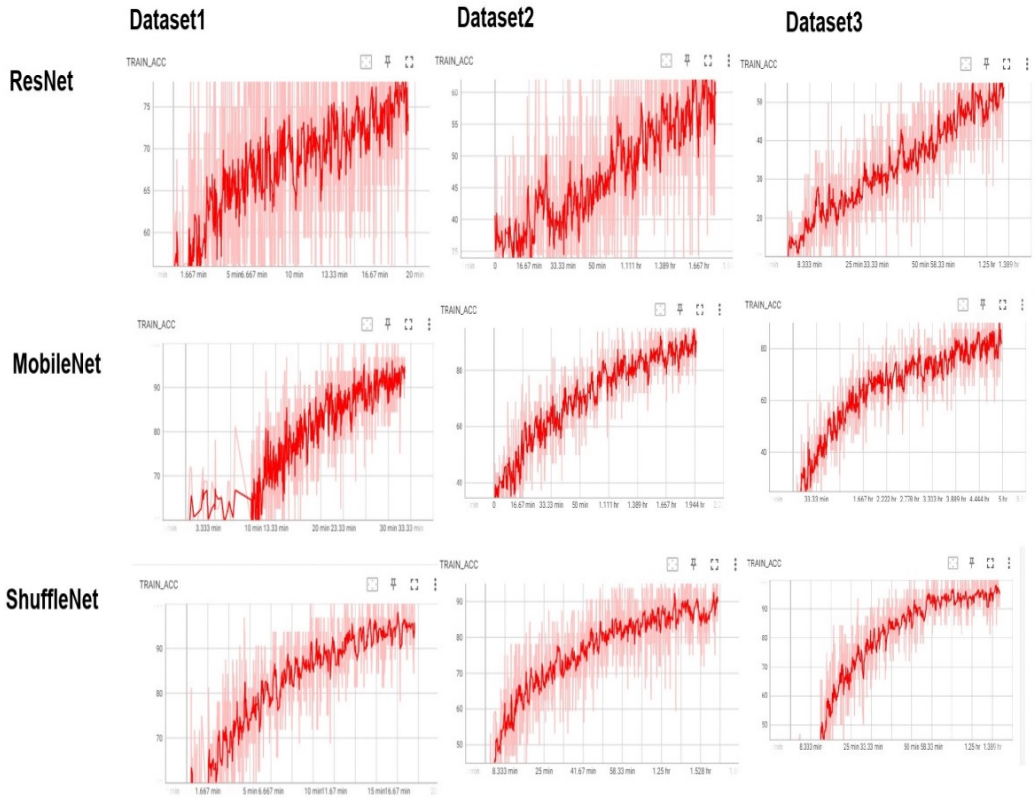

Figure 1.4

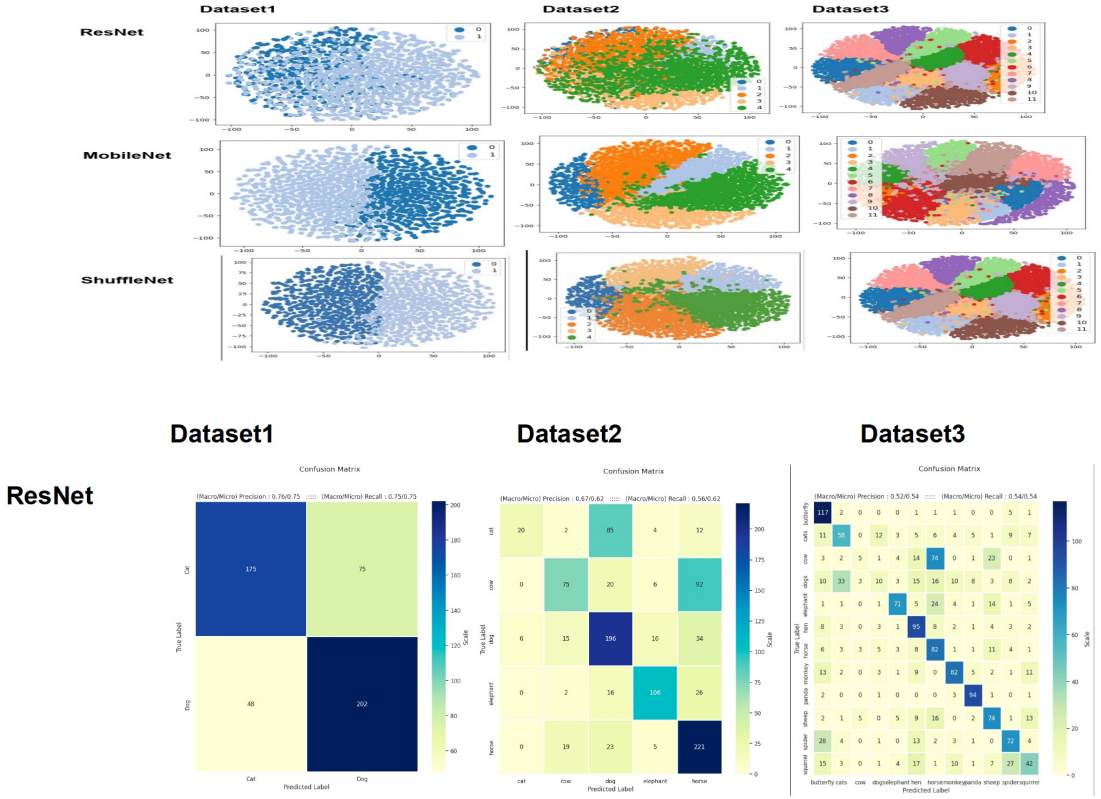

Figure 1.5

Conclusion

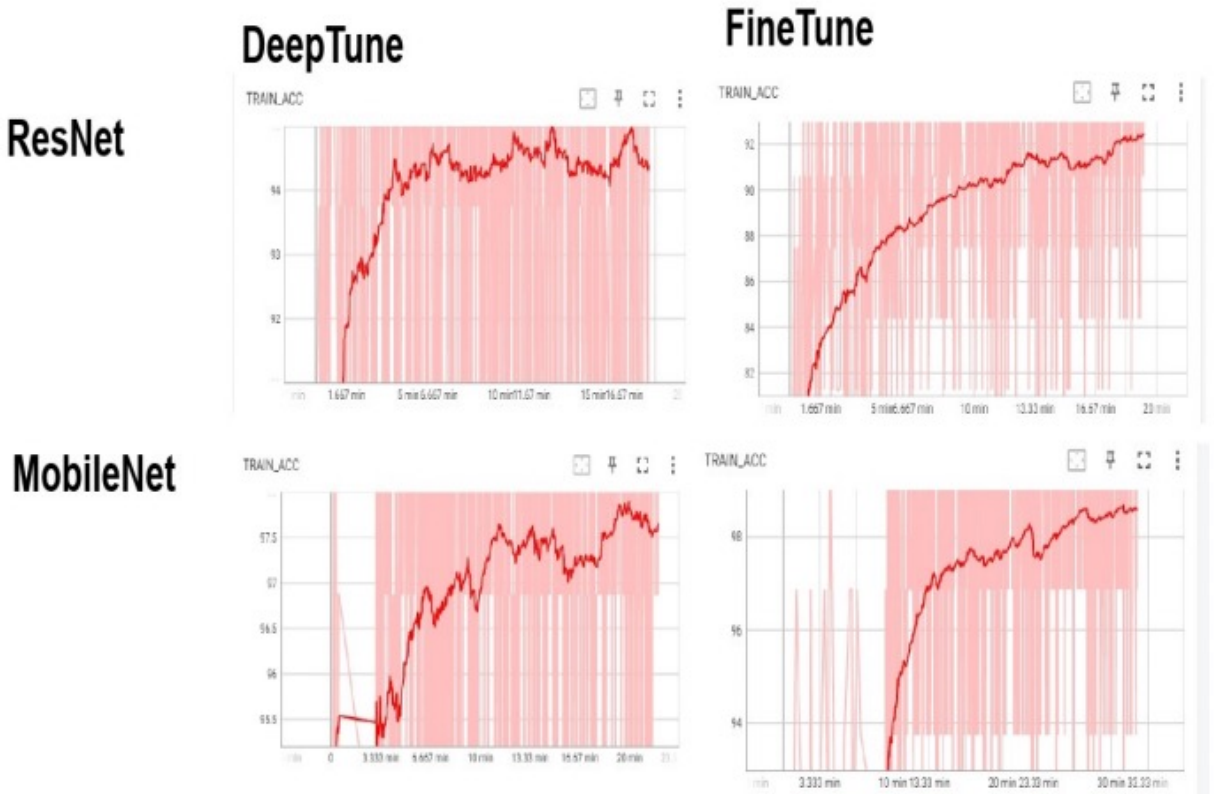

Figure 1.6

- Overall, after comparing different CNN architectures on datasets of different sizes, Different classes it can be stated that ResNet18 using transfer learning achieved highest-accuracy.

References

- Julia Larson. Assessing Convolutional Neural Network Animal Classification Models for Practical Applications in Wildlife Conservation. Master's Theses. 5184 (2021).
- Moallem G, Pathirage DD, Reznick J, Gallagher J, Sari-Sarraf H. An explainable deep vision system for animal classification and detection in trail-camera images with automatic post-deployment retraining. Knowledge-Based Systems. 2021 Mar 15;216:106815.
- Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen. MobileNetV2: Inverted Residuals and Linear Bottlenecks. arXiv:1801.04381v4 [cs.CV] 21 Mar 2019.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. Deep Residual Learning for Image Recognition.arXiv:1512.03385 [cs.CV] 10 Dec 2015.
- Ningning Ma, Xiangyu Zhang, HaiTao Zheng, Jian Sun. ShuffleNet Practical Guidelines for Efficient CNN Architecture Design. arXiv:1807.11164 [cs.CV] 21 Dec 2021.
- Will Cukierski.(2013). Dogs vs. Cats. Kaggle: https:/ /kaggle.com/competitions/dogs-vs-cats
- Saumil Agrawal. Animal Image Dataset resized. Kaggle: https:/ /www.kaggle.com/datasets/saumilagrawal10/animal - image-dataset-resized
- Piyush Kumar. Animal Image Classification Dataset. Kaggle: https:/ /www.kaggle.com/datasets/piyushkumar18/animal - image-classification-dataset