**PROJECT REPORT ON**

**“FINE-GRAINED CLASSIFICATION”**

Submitted By:

SHUBHAM THAKUR (23M0306)

BIJAY CHANDRA DAS (23M0313)

DHEERAJ KUMAR (23M0325)

**Under The Guidance Of:**

Prof. BIPLAB BANARJEE

Centre of Studies in Resources Engineering

**March 2024**



**CENTRE OF STUDIES IN RESOURCES ENGINEERING**

**Indian Institute of Technology, Bombay**

**Mumbai, Maharashtra, India**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **No.** | **Topics** | **Page** |
| 1. | Introduction | 3 |
| 2. | Methodology | 4 |
| 3. | Results | 6 |
| 4. | Conclusion | 7 |
| 5. | References | 8 |

**INTRODUCTON**

Fine-grained classification is a challenging task in computer vision where the goal is to classify objects into fine-grained categories, often belonging to the same superclass. The CUB-200-2011 dataset is a widely used benchmark dataset for fine-grained classification, consisting of 200 bird species with high inter-class similarity. In this project, we aim to build a deep learning model to classify bird species using the CUB-200-2011 dataset.

**Dataset Description:** The CUB-200-2011 dataset contains 11,788 images of birds belonging to 200 different species. Each image is annotated with bounding boxes and part locations. The dataset is divided into training and test sets, with 5,994 images for training and 5,794 images for testing. The dataset poses several challenges, including intra-class variation and similarity between different species.

**METHODOLOGY**

**Data Preprocessing:** We preprocess the dataset by resizing the images to a fixed size, normalizing pixel values.

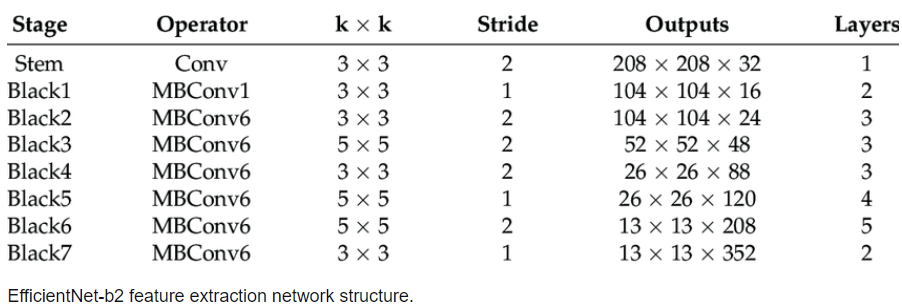
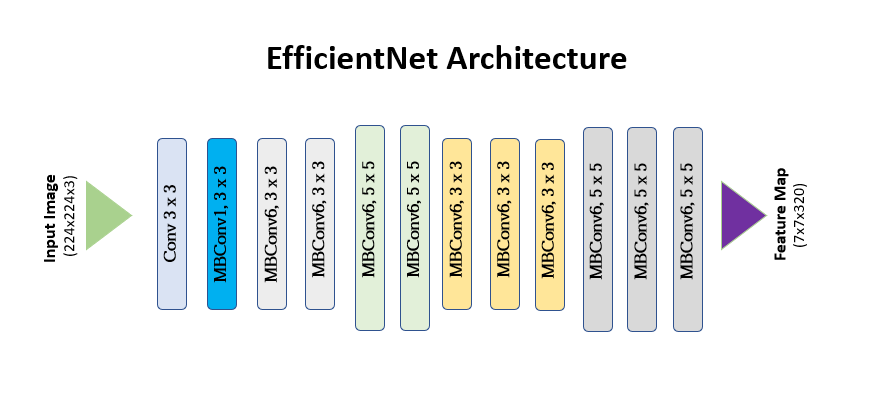
**Specific Requirements**

* python3
* torch
* torchvision
* PIL
* matplotlib
* numpy
* efficientnet\_pytorch

**Model Architecture**

For this project, we use the EfficienNnet-b2 architecture, EfficientNet-B2 is a convolutional neural network architecture that balances model size and computational efficiency using a compound scaling method. It consists of a stack of convolutional layers with depthwise-separable convolutions and squeeze-and-excitation blocks for feature refinement.

**Total no. of Parameters Used: 9142794 (9.14M)**



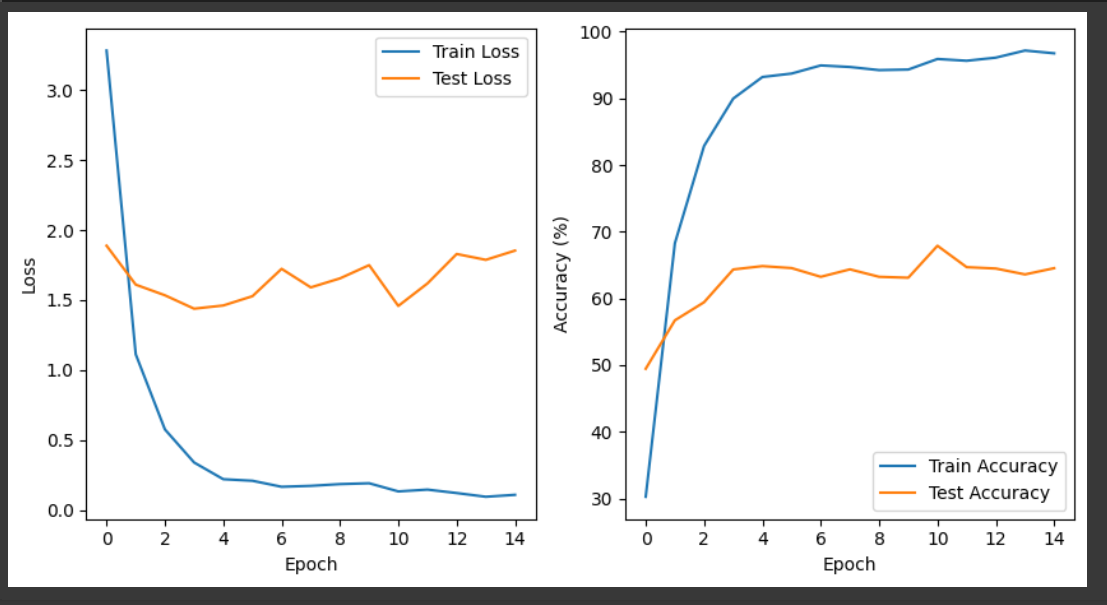
**Experimental setup:**

* **Preprocessing**: Images are resized to 224x224 pixels and normalized to have pixel values in the range [0, 1].
* **Training**: The EfficienNnet-b2 model is trained using Adam optimizer. We use a learning rate scheduler to gradually decrease the learning rate during training.
* **Evaluation**: The trained model is evaluated on the test set using accuracy as the evaluation metric.

**RESULTS**

After training the EfficienNnet-b2 model on the CUB-200-2011 dataset, we achieved the following results:

* **Training Accuracy**: Achieved a training accuracy of 96.76% after 15 epochs.
* **Test Accuracy**: Achieved a test accuracy of 64.53%.



**LINKS**

1. Finalcheckpoint:<https://drive.google.com/file/d/1vqm8nrBaCV7M6p63LdFpK12uJD8_qYpg/view?usp=sharing>
2. Tensorfiles: <https://drive.google.com/drive/folders/1FsXhL-Yp29UIoQrwT9e0bT5bnDqTC9PN?usp=sharing>
3. Y\_test:<https://drive.google.com/file/d/1AlRgntYfutAe2x04qW8SmntcRdRlIazJ/view?usp=sharing>
4. Y\_train:<https://drive.google.com/file/d/13lV2PLCD1XWAZYNkJ4cu6jItIEYfjnLL/view?usp=sharing>

**CONCLUSION**

In this project, we demonstrated the effectiveness of EfficienNnet-b2 architecture for fine-grained classification on the CUB-200-2011 dataset. The model achieved high accuracy and demonstrated good generalization ability. Fine-grained classification tasks are important in various domains such as wildlife monitoring, disease detection in plants, and product recognition in e-commerce. EfficienNnet-b2's efficiency makes it a practical choice for deploying fine-grained classification models in real-world applications.

By continuing to explore and improve fine-grained classification techniques, we can further advance the capabilities of computer vision systems in recognizing subtle visual differences and contributing to various real-world applications.

**REFERENCES**

* Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In Proceedings of the 36th International Conference on Machine Learning (ICML).
* Wah, C., Branson, S., Welinder, P., Perona, P., & Belongie, S. (2011). The Caltech-UCSD Birds-200-2011 Dataset. California Institute of Technology.
* <https://pytorch.org/vision/stable/models.html>