Title: Bird Classification Based on the Efficient Net Architecture

Abstract. This research aims to apply one of the modern Neural Network architectures, Efficient Net, on bird images to predict the categories they belong to.

Keywords: Classification, Deep learning, Efficient Net, Image Processing, Data Manipulation.

Introduction. Animal life has always been in the circle of interest of various scientists. The research's prior goal is to analyse how Efficient Net architecture works on bird classification. There have been several Convolutional Neural Network approaches and improvements for classification problems and Efficient Net differs from other ones due to its high accuracy on ImageNet. Briefly, there are several Efficient Net variants (B0-B7) and at least two of them (B0 and B1) are considered to be applied to the image classification to make comparisons on the results.

Materials and Methods. Initially, the issue depends on features like feather or hair colour, the size of the beak, and wings, to mention a few, which are extracted to make predictions. The dataset taken from the Kaggle contains images of 200 various bird species and Data Normalization as a pre-processing step can be done before feeding it into the model. Moreover, there are approximately 11.8k images in the dataset with various shapes and several techniques, such as normal data splitting, and stratified data splitting, will be used to get better performance The dataset will be divided into the training (80%), validation (10%), and test (10%) parts to build the model and check how the model behaves based on the training data. While training the model, the data will be divided into smaller sets called "mini-batches", and batch size is planned to be 16,32 or 64.

All Efficient Net techniques have a different number of layers ranging from 237 (EfficientNet-B0) to 813 (EfficientNet-B7) and the stages contain various steps including Pooling, Scaling, Convolution and Squeeze Excitation. Moreover, as it is known to build CNN models, feedforward and back-propagation, in which the weights are adjusted, algorithms are done and thus, Adam, the SGD with momentum and RMSProp optimizations will be used as an optimization algorithm.

The application of theory and testing model's efficiency is regarded to be accomplished in the Google Collaboratory environment which gives additional opportunities such as a GPU with an effective level of runtime. Turning to the technologies, Open Computer Vision (OpenCV) and high-level machine learning framework PyTorch from the Torch family, and other Python libraries, including PIL (Pillow), and Numpy, will be used to build the architecture from scratch. In case of any problems, such as overfitting, new ways will be defined to overcome the challenge and build a better model.

Results and Conclusion. The built model's accuracy will be determined related to the evaluation metrics, including accuracy, precision, recall, and f1-score and the outputs of the two models will be definitely compared to each other. In addition, it is possible to make a comparison between the models typed from scratch and the pre-trained model. Mainly, where the model makes an incorrect prediction, and the values of training loss and test loss, are essentially considered to be touched and explained.