COMP 6721 (Applied Artificial Intelligence) - Project Proposal

Breed Identification System (Group B)

1. Problem Statement

Human interaction with animals has grown significantly, in last few hundred years with advancement of technology. With scientific advancement curiosity has also grown prompting humans to learn more about different species residing in earth and at the same time save them from getting endangered or extinct through human intervention. Within the same species there can be numerous breeds, which is a distinction worth knowing.

Breed Identification System is designed to classify the breed-class for a given type of animal. Practical applications of our system include pet stores, vet clinics, endangered species survey, potential ecological background study of breeds tied up to their location.

Major challenge for this problem is availability of images for various breeds of one animal and another important obstacle is limited resources available on internet for reference. There also exists the problem of two breeds having identical features which makes it tough to build a reliable solution.

Goal of the project is to explore and give detailed analysis of how different CNN architectures fares against the chosen datasets of animals. We will do a thorough comparison and point out appropriate model through which high accuracy can be achieved. With the rapid growth we are seeing in transfer learning capabilities and the high similarity between problem type we believe any improvements we make in this project can be ported to other animal breed identification without too much trouble.

2. Dataset Description

Animal	Number of Images	Image sizes	No. of breeds
Dog [1]	20.6 k	{400 - 500} X {310 - 345}	120
Cat [2]	127 k	{300 - 330} X {250 - 270}	67
Fish [3]	9 k	590 X 445	9

The table above shows some statistical information about the datasets we have chosen. Images across the dataset are all in color (3 channels) with varying sizes. The cat and dog datasets suffer the class imbalance problem where the dog dataset have certain breeds with twice as many images as other breeds. The cat dataset contains 127k images but to fit the problem description we'll prune the dataset to bring it under 100k.

3. Methodologies

We will start our preprocessing by rescaling/down-sampling the images to a fixed resolution as our CNN models will require a constant input dimensionality. Tools like TSNE will be used to visualize the data and class activation maps. Further, image data augmentation can be looked into in order to improve model performance and reduce generalization error.

We will begin our model experimentation by making use of some of the SOTA models for this type of classification which includes ResNet (ResNet18, ResNet34, ResNet50) [4] and MobileNet (mobileNetV2, mobileNetV3) models. We plan on exploring different techniques like direct transfer learning, training these models from scratch (bearing computation cost in mind). We also propose to do a brief study of newer methods used in multi-class classification [6].

Based on the results of our first cut models, we then plan on conducting an extensive ablation study on a model of choice where we try to determine what's the best combination of hyperparameters i.e., learning rate, loss function, epochs, batch-size, optimizers etc. for the model-dataset which maximizes our evaluation metrics.

We are going to compare them by using the N*N confusion matrix) and depending on the class imbalance F1 score or ROC AUC. We are also open to include new metrics that we might come across while reading papers [7]. Through our experimentations we hope to determine a multitude of factors including (but not just limited to) the impact of class-size, dataset-size on accuracy, different network's depth vs dataset, the cross link between a dataset-model combination and which section of the image is being used to make a prediction (GradCAM), etc. thereby helping us identify which model performs best for a given dataset.

Considering computation resources/cost into picture, we might be using some external tools, e.g., Kaggle or Google Colab (or Colab Pro) to train our models.

Gantt Chart

Task	Responsible	25-Sep-22	2-Oct-22	9-Oct-22	16-Oct-22	23-Oct-22	30-Oct-22	6-Nov-22	13-Nov-22	20-Nov-22	27-Nov-22	4-Dec-22	Total Hours
		1	2	3	4	5	6	7	8	9	10	11	
Problem Statement Research	Team	4											4
Dataset Search	Nandkumar, Abhishek		4										4
Drafting Proposal	Everyone		2										2
Data Preprocessing	Vishvesh			8	7								15
Resnet50 code on Dog	Nandkumar			7	8	5							20
Resnet50 code on Fish	Vishvesh			6	7	4							17
Resnet50 code on Cat	Tannavi			3	5	7	8						23
Resnet50 ablation+testing	Abhishek			5	5	7	5						22
Data Preprocessing	Vishvesh						5	4					9
MobilenetV2 code on Dog	Tannavi						6	7					13
MobilenetV2 code on Fish	Vishvesh							6	7	4			17
MobilenetV2 code on Cat	Nandkumar						7	8	7				22
MobileNet ablation+testing	Vishvesh							4	6	8			18
Model Research + Progress Report	Tannavi							2	2	5			9
Data Preprocessing	Abhishek							5	4				9
Renet18 code on Dog	Nankumar								3	6	8		17
Renet18 code on Fish	Vishvesh								2	2	6		10
Renet18 code on Cat	Abhishek								5	7	8		20
Resnet18 ablation+ testing	Abhishek									3	4	7	14
Detailed ablation study	Tannavi									2	6	7	15
Eval metrics research+scorecard	Vishvesh									3	5	9	17
Modify optimizers / hyperparameters in all 9 models for enhancement	Nandkumar, Vishvesh										6	8	14
Integration of all models	Nandkumar											5	5
Making Final Report	Vishvesh,Tannvi										1	4	5
Making Final PPT	Abhishek										1	2	3
Total Hours		4	6	29	32	23	31	36	36	40	45	42	324

On initial level, team found a problem statement according to ease of access of datasets with certain conditions like number of classes, varying image sizes and sufficient samples for each class. After finalizing the problem statement and datasets we moved on to drafting proposal and begin preprocessing the datasets. After proposal acceptance, we will start building our implementations for different models and based on our computational cost research we decide which model we are training from scratch, and for which model we are doing the transfer learning. We have given extra time for data pre-processing again in case we discover anything along the way and need it to reflect across all model-dataset combos. Progress-reporting will be done after the completion of 6 model instance out of 9. Then we focus on building the remaining models and doing a comparative analysis of all 9 model instances and results will be curated. Time for modifications in optimizers/ hyper-parameters are allotted in project schedule to enhance results of all models which a certain model-dataset combo will be chosen for conducting a detailed ablation study. After completion of all development tasks, we focus on integrating all 9 instances and its results to an accessible format and location, while documenting the steps. At last, final report and PPT will be prepared and delivered before the allocated deadline.

References:

[1] https://www.kaggle.com/competitions/dog-breed-identification/data

[2] https://www.kaggle.com/datasets/ma7555/cat-breeds-dataset

[3] https://www.kaggle.com/datasets/crowww/a-large-scale-fish-dataset

[4] Howard, Andrew AU - Zhu, Menglong AU - Chen, Bo AU - Kalenichenko, Dmitry AU - Wang, Weijun AU - Weyand, Tobias AU - Andreetto, Marco AU - Adam, Hartwig PY - 2017/04/16 SP - T1 - MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications ER

[5] He, Kaiming AU - Zhang, Xiangyu AU - Ren, Shaoqing AU - Sun, Jian PY - 2016/06/01 SP - 770 EP - 778 T1 - Deep Residual Learning for Image Recognition DO - 10.1109/CVPR.2016.90 ER

[6] Vasudevan, Varun AU - Bassenne, Maxime AU - Islam, Md Tauhidul AU - Xing, Lei PY - 2022/01/29 SP - T1 - Image Classification using Graph Neural Network and Multiscale Wavelet Superpixels ER

[7] Grandini, Margherita et al. "Metrics for Multi-Class Classification: an Overview." *ArXiv* abs/2008.05756 (2020): n. pag.

[8] Shah, Vishal AU - Sajnani, Neha PY - 2020/11/20 SP - 65 EP - 68 T1 - Multi-Class Image Classification using CNN and Tflite VL - 3 DO - 10.47607/ijresm.2020.375 JO - International Journal of Research in Engineering, Science and Management ER