COMP9417: Machine Learning

2024 T1 – Group Project: Squad404

George Luo Yifei Tian Luke Wang

Z Z Z

Xiaolu Liu Lijun Guo

# Introduction

Image classification has been an age-old machine learning problem which has had an undeniable impact in numerous fields such as medical imaging, autonomous vehicles and biometrics just to name a few. With constant advancements in technology, we have gained access to increasingly powerful hardware with immense data processing capabilities. This has given rise to highly effective models from complex CNNs such as VGG-19, seen in [Simonyan et al., 2014] to the innovative transformer-based models such as ViT explored by [Dosovitskiy et al., 2021] and beyond. However, these models require substantial data for effective training, presenting challenges in data-constrained environments.

This project aims to explore methods to tackle few shot learning problems in such cases where the dataset is very limited in order to create a more accessible approach to image classification. We have been challenged to create a model to tackle a binary image classification task that can generalize over a diverse collection of datasets given only 5-10 images per class. We have approached this by incorporating convolutional neural networks[INSERT CNN PAPER] in their ability to extract rich and descriptive features, as well as prototypical networks in their ability to generalize well over limited datasets[INSERT PROTOTYPICAL PAPER].

## Exploratory Data Analysis

As part of the Berrijam Jam 2024, we have been provided with 3 sample datasets to train our model on. These datasets included:

* Is Epic Intro: A collection of 10 spectrogram images representing the first 30 seconds of a song to determine whether the song has an epic feeling.
* Needs Respray: A collection of 12 images of pavement to determine if there are active weeds growing indicating the need for a chemical respray.
* Is Gen AI: A collection of 20 scenery images to determine which are real and which have been generated by AI.

Each of the datasets are balanced and most are oriented in landscape with the exception of some images from the is Gen AI dataset.

Citations

VGG19: <https://arxiv.org/abs/1409.1556>

ViT: <https://arxiv.org/abs/2010.11929>