Object Recognition

Proposal

Alex Conroy
Jordan McCallum
Nebojsa Ajdarevic

student numbers:
08324841
09155881
10434348

CAB420 Assignment 2

Queensland University of Technology (QUT)
GPO Box 2434, Brisbane QLD 4001, Australia

May 3, 2020

Team Members

Team Member	Student Number
Alex Conroy	08324841
Jordan McCallum	09155881
Nebojsa Ajdarevic	10434348

Project Title

Object Recognition

Motivations and Objective

To create and evaluate multiple object recognition systems which can identify which object category is the best match for a given test image and determine which is the best suited approach for this dataset.

Dataset

The dataset being used for this project is the Caltech 256 dataset. It contains images of 256 object categories taken at varying orientations, varying lighting conditions, and with different backgrounds.

Evaluation Protocol

To evaluate the best method to identify object categories the report will evaluate three different model types in three different categories.

The three evaluation categories are:

1. Model Accuracy
2. Computational Efficiency
3. Computational Time

These three categories form the three pillars of an excellent Machine Learning model, and as such will all be evaluated to determine which model is best suited.

The three model types that will be compared in these categories are:

1. Convolution Neural Network (CNN)
2. Multiclass Support Vector Machine (SVM)
3. Generative Adversarial Networks (GAN)

These represent a wide variety of potential methods of modelling; supervised, semi-supervised, categorical, and regression.

The first model to be used will be a Convolution Neural Network (CNN). Specifically, the model will use EfficientNet, this model released approximately 12 months ago by Google. The model that has set records for both accuracy and computational efficiency. EfficientNet uses multi-dimensional scaling to avoid diminished returns on computational efficiency. The model also employs compounds

scaling to attempt a synergy of dimension scaling, again this is done in the name of efficiency. These measures have helped the model gain a reported 5x reduction in parameters required to achieve the same accuracy rating as comparable CNNs. Lastly EfficientNet has displayed similar magnitude gains in time when using heatmaps to capture items in a picture. Overall making this model choice a great implementation of a CNN model type.

The second model was first introduced in 2014 is the Generative Adversarial Networks (GAN). GANs belong to the generative modelling family, meaning that they produce new content. This is accomplished by generating complex random variable with respect to a probability distribution before being indirectly trained by a neural network and reshaped by the target distribution. The indirect training is accomplished by tricking another network that is trained at the same time to distinguish “generated” data from “true” data. Both can be employed with forward and backward propagation.

The last model to be a multiclass Support Vector Machine and combined bag of features. This approach is designed to emphasize better recognition and classification accuracy. The approach will be modelled after the paper ‘Object Recognition Using SVM Based Bag of Combined Features’ by Fozia Mehboob, Muhammad Abbas, and Abdul Rauf (2019). The paper outlines the use of Root Scale invariant feature transforms (SIFT) and speeded-up robust features (SURF) to produce state of the art performance benchmarks in object recognition on another Caltech dataset.

Reference

Mehboob F., Abbas M., Rauf A. (2019) Object Recognition Using SVM Based Bag of Combined Features. In: Arai K., Kapoor S., Bhatia R. (eds) Intelligent Computing. SAI 2018. Advances in Intelligent Systems and Computing, vol 858. Springer, Cham