INTRODUCTION TO IMAGE PROCESSING AND COMPUTER VISION

**Project 2: Plant species recognition**

Szymon Majorek

[s.majorek@student.mini.pw.edu.pl](mailto:s.majorek@student.mini.pw.edu.pl)

Table of contents

1. Introduction
   1. General introduction
   2. Data set presentation
2. Part 1 – Feature extraction
   1. Short Introduction
   2. Features and potential drawbacks
3. Part 2 – testing obtained data
   1. Short introduction
   2. Cross validation
   3. 80-20 split
4. Part 3 – results
   1. Cross validation
   2. 80-20 split
5. Conclusions

**1. Introduction**

**1) General Introduction**

Machine learning is a study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instruction, relying on patterns and inference instead. In other words, we can teach program to do something for us, without us giving explicit input for our decision. In our case we will be using Random forests model.

Image recognition on the other hand is the ability to identify and detect objects or features of digital images or videos. It is a method for capturing, processing, examining and sympathizing images. We can use machine learning for image recognition, as we will in this project.

The goal of this project is to use machine learning in order to recognize different species of plants based on images of their leaves. We aim to teach our program to recognize different plant species by teaching it how do their leaves look like.

**2) Data set presentation**

**Our data set is quite simple actually. We have 6 types of leaves with irregular number of images of them. The types are as follows*: Acer Circinatum, Acer Glabrum, Acer Macrophyllum, Acer Negundo, Quercus Garryana* and *Quercus Kelloggii*:

Figure 5 - Acer Negundo

Figure 6 - Acer Macrophyllum

Figure 4 - Quercus Kelloggii

Figure 1 - Acer Circinatum

Figure 3 - Quercus Garryana

Figure 2 - Acer Glabrum

When looking at those images it’s quite easy to say that they are all different species. For example the one on fig 3 is sort of compact but with rounded tips, which is different from fig. 4 which has pointy tips instead. In fact all of them have pointy tips apart from the 3rd one however they have some slight differences which are easy to notice a first glance. They even have different colours but that might just be a coincidence that I randomly chose those examples which actually have visibly different colours. Now that is all obvious by looking at the pictures, however in terms of our project, we need to think a little differently. We know what are leaves, we know how they look like and where to look for differences between leaves of different plants (shape, number of pointy ends, whether the ends are actually pointy or more rounded etc.). One noticeable thing that us, humans don’t usually look for when trying to distinguish different leaves is their colour because we tend to think of leaves as generally green (well with small variations depending on seasons).

**2. Feature extraction**

**1) Short introduction**