Title: Identifying aircraft from above

Name: Kai Roper-Blackman

Registration Number: 1602999

Supervisor(s): Adrian Clarke, Sebastian Halder

Second assessor: Luca Citi

Degree Course: BSc Computer Science (G400)

Contents

[Acknowledgements 4](#_Toc5894791)

[Abstract 5](#_Toc5894792)

[List of Symbols 6](#_Toc5894793)

[Project Aims and Objectives 7](#_Toc5894794)

[Primary Objectives 7](#_Toc5894795)

[Additional goals 7](#_Toc5894796)

[Methods 8](#_Toc5894797)

[Internal and External Libraries 8](#_Toc5894798)

[OpenCV 8](#_Toc5894799)

[Scikit-Learn 8](#_Toc5894800)

[Scikit-Image 8](#_Toc5894801)

[FPDF 8](#_Toc5894802)

[NumPy 8](#_Toc5894803)

[Matplotlib 8](#_Toc5894804)

[Tkinter 8](#_Toc5894805)

[PIL 9](#_Toc5894806)

[Data set 10](#_Toc5894807)

[Images 10](#_Toc5894808)

[Parsing 11](#_Toc5894809)

[Pre-Processing 11](#_Toc5894810)

[Histogram of oriented gradients 11](#_Toc5894811)

[Feature vector 12](#_Toc5894812)

[Machine learning 14](#_Toc5894813)

[Supervised learning 14](#_Toc5894814)

[Support vector machines 14](#_Toc5894815)

[Cross validation 16](#_Toc5894816)

[Technical Achievement 18](#_Toc5894817)

[Normal standalone classification 18](#_Toc5894818)

[Large image search 18](#_Toc5894819)

[Project Planning 19](#_Toc5894820)

[Momentum 19](#_Toc5894821)

[Adapting to change 19](#_Toc5894822)

[Identifying and dealing with risks 19](#_Toc5894823)

[Achievement 19](#_Toc5894824)

[Performance 19](#_Toc5894825)

[What have I learnt? 19](#_Toc5894826)

[Conclusions 20](#_Toc5894827)

[Discussion 20](#_Toc5894828)

[References 21](#_Toc5894829)

[Tables, Graphs, Figures and Equations 21](#_Toc5894830)

[Appendices 21](#_Toc5894831)

[Sustainability 21](#_Toc5894832)

[Legal 21](#_Toc5894833)

[Ethical 21](#_Toc5894834)

[Intellectual property 21](#_Toc5894835)

# Acknowledgements

A special thanks to my parents Annie Roper and Seon Blackman for supporting me throughout University.

Thank you to all the academic staff in CSEE, specifically Adrian Clarke, Sebastian Halder and Luca Citi for feedback and guidance throughout this project.

# Abstract

We all know what an aircraft looks like, but does a computer? A seemingly simple task that can be carried out by individuals at age two, poses a complex problem to modern technology. Machine learning is a relatively new field with little research but already boasts claim to many applications such as driverless cars and face recognition systems. The development of object recognition is the center of many companies’ business models and objectives, making aircraft identification such an interesting topic to research.

Existing images of ground and aircraft are pre-processed using Histogram of Gradients to create feature descriptors. Feature descriptors describe the orientation of a gradient within an image subsection. Support Vector Machines are passed feature descriptors with labels for training. Once training is completed, the support vector machine accepts a test set and returns predictions. Large image search takes a large image and looks within a smaller area for aircraft. Search area parameters are provided by the user.

The results obtained from cross validation show an accuracy of 100% when identifying standalone aircraft. However, when searching for aircraft in larger images, accuracy drastically decreases to around 55% as some aircraft are overlooked. After optimization, the system used to identify aircraft can be applied to other identification problems with possible military and commercial uses.

# List of Symbols

|  |  |
| --- | --- |
| **Acronym** | **Meaning** |
| SVM | Support Vector Machine |
| PNG | Portable Network Graphics |
| px | Pixels |
| HOG | Histogram of Oriented Gradients |
| PIL | Python Image library |
| GUI | Graphical User Interface |
| ML | Machine Learning |
| AI | Artificial Intelligence |
|  |  |
|  |  |

# Project Aims and Objectives

The original aims of this project have changed since the initial report. This was because of a change of approach to the given problem. The main objective is to allow a computer to differentiate between **aircraft** and **ground** images. This can be broken into several smaller objectives.

## Primary Objectives

1. To gather a set of aircraft and ground images to create a training and test set
2. To normalise the training and test set
3. Change the orientation of aircraft to ensure the face north
4. Resize the images so they can be pre-processed accurately
5. To pre-process the training set and label accordingly.
6. To Train an SVM with the Training set
7. To test the SVM with the test set
8. Further optimise the SVM by tuning decision boundaries
9. Cross-validate the training set
10. Display results of classification to user.

Assuming all primary objectives are completed successfully, additional objectives are to be attempted

## Additional goals

1. To attempt to recognise aircraft in a large image using user defined search criteria and show aircraft locations to user after search is completed.
2. To generate heat maps of large images to identify areas with a probability of containing aircraft
3. To attempt to recognise non-commercial aircraft such as private aircraft and helicopters.

Once all goals have been completed, a software program with realistic uses should be established. The training set can be varied to identify additional vehicles or identify completely different objects.

# Methods

## Internal and External Libraries

During the project, I have used a variety of libraries to aid the completion of the projects aims and objectives. Writing code to support all functionality provided by product would be very difficult to to achieve in the tight timescale given.

The language I have chosen for this project is python as it has a large range of image processing libraries available and offers a wide range of features such as object orientation. Internal libraries come as standard with the python package whereas external libraries are developed by third parties. All third-party libraries used during this project are used in accordance to licensing laws and agreements laid out in there terms and conditions.

### OpenCV

The library used to read images from PNG format to NumPy arrays was OpenCV. OpenCV has tools for image manipulation and other image related functions. OpenCV was used to rotate images before training.

### Scikit-Learn

Scikit-Learn is a well-documented Machine learning library that offers a wide range of tools for machine learning, data mining and data analysis. Scikit-learn has developed a Support Vector Machine that can be easily utilised for aircraft identification. The SVM forms the basis of my product and classifies objects as it sees fit.

### Scikit-Image

Scikit-Image is a collection of image processing algorithms for the Python programming language. This library contains the HOG algorithm used in image pre-processing to allow the SVM to distinguish between objects in images.

### FPDF

FPDF is a library that allows interaction between Python and pdf files. This allows the results of classification to be saved to and viewed by the user.

### NumPy

NumPy is the fundamental package for scientific computing in Python [1]. It comes with a variety of features and functions such as mathematical, logical and shape manipulation. This makes it the perfect library to store images as arrays. Other libraries used during this project utlise NumPy often returning NumPy arrays after specific functions. During the training stages, images are rotated using NumPy’s rotate function.

### Matplotlib

Matplotlib is a 2D plotting library that produces publication quality figures. During this project, I used its features to display results of classification to the user and allow them to interact with them using the libraries GUI. This allows them to zoom in and move Images and plots around as they wished. The user can also choose to save the output if they wish for later viewing.

### Tkinter

Tkinter is the Python standard for GUI development. The library allows the development of complex windowed GUI’s. It allows the placement of buttons, labels, text boxes, images, radio buttons and drop-down menus. Tkinter also has GUI’s for functions such as file selection.

### PIL

PIL is python’s image library that adds support for opening and saving images in multiple image formats. The library contains functions to easily manipulate images and draw shapes/ text over them. The shape functions have been utilised in this project.

## Data set

For a computer to recognise an object, first you must teach it what the object looks like. There are many different learning techniques and methods to do this. The technique I chose to implement was supervised learning. To carry out supervised learning, a training set must be created. The training set must show an accurate representation of the object or the model will fail to understand what it is identifying.

### Images

The data set currently consists of 250 images of aircraft and 250 images of ground. Images were obtained from google earth by taking screenshots of airports and proceeding to crop aircraft and ground. Cropped images are then saved as individual images. Cropped images of aircraft have little empty space around them so that the SVM can focus primarily on the SVM. Examples of training images are shown in Figure 1.

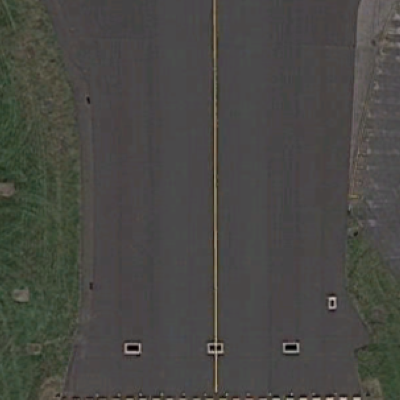


Figure : Image of ground (left) image of aircraft (right)

Originally, I ensured all images of aircraft were rotated so that the aircraft faced north. All images were also of the same dimensions (400x400 px) ensuring training was fair and balanced and are of the PNG format for simplicity.

However, after development of the image search classification method, it quickly became apparent that aircraft in real life can exist in many different sizes and rotations relative to image image search criteria. At present, the user defines the search area size and movement settings meaning the SVM may not always have a perfect view of an aircraft. The aircraft may appear off centre meaning the SVM will struggle to classify it. To overcome this problem, I needed to generate a realistic training set. To do this I gathered images of some aircraft at random rotations and of varying sizes therefore filling different amounts of the image. This gave the SVM a more realistic understanding of how an aircraft appears in real life.

To again simplify the realistic data set, I ensured each photo contained no more than one aircraft. If an image contained more than one aircraft, it would interpret an aircraft as an object of two or made of two aircraft combined. Figure 2 shows an example of this occurrence.



Figure : Image not included in data showing two aircraft

Images of aircraft with complicated surroundings such as complex jet ways, terminals and vehicles in many cases were also disregarded. This is because it shifts the centre of attention of the SVM. The SVM will start to think an aircraft is an aircraft when it has an assortment of surrounding vehicles, buildings and jet ways in proximity. The idea behind the

### Parsing

An image in computer science is regarded as an array of values. Each field contains a value which is displayed onscreen as a channels colour intensity. To train a machine learning model, the training set is provided as an array of image data. The image data created by reading images from the file system using the OpenCV library. The OpenCV function to read images, returns a NumPy array of raw image data. The image data is then added to a large array containing other training data. The process is repeated for every image in the training set until the training data is contained in a single array. Concurrent to this process, a separate label array is filled with 1’s and 0’s. 1 = Aircraft, 0 = Ground. This array tells the SVM which image belongs to which classifier.

To add an element of further realism, the method that generates the training set for the large image search, manipulates images at random to remove any element of bias that may be incurred by the default training set. This is done before training. Images are rotated randomly by intervals of 90 degrees and then pre-processed (see next section). They are then added to the large training array.

## Pre-Processing

Pre-processing takes an image and extracts key information about it. There are several pre-processing methods that are available and commonly used in machine learning. Pre-processing is used to reduce the load on the machine learning model and only supply it with relevant information. When images aren’t pre-processed, the machine learning model is often slow and inaccurate. This is because of the excess data provided to it. Image pre-processing creates a feature vector that describes an image. The feature vector is then provided to the machine learning model for training and evaluation.

### Histogram of oriented gradients

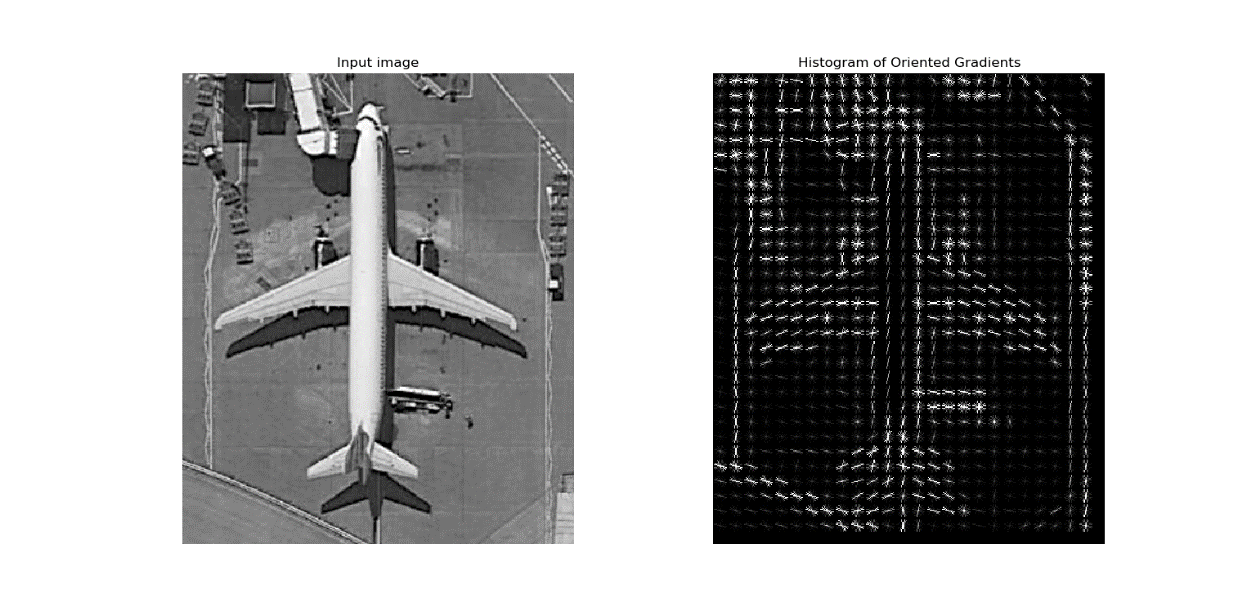
Histogram of oriented gradients is the current pre-processing method used in the parsing stage of the aircraft detection program. HOG counts the occurrences in gradient orientation in localized cells. The cells are normalised in blocks using contrast normalisation. A cell and blocks size can vary. Blocks can consist of any amount of cells. In this case, the blocks cells are normalised to ensure the results have low variance.

Contrast normalisation is a simple image enhancement technique that changes the range of image intensity values. Minimum and maximum values are supplied to a normalisation function which are applied to the image. This process is carried out to remove bias caused by illumination and shadowing in the given image.

After contrast normalisation is completed, the next step is to create a histogram for every cell in the image. Once the histogram is created, the orientation can be calculated. The direction of the gradient is added to a feature vector.

HOG was implemented in this program as it is widely used with SVM’s for object detection. Figure 3 shows a visualisation of the HOG algorithm on an aircraft from the training set.

Figure : Histogram of Oriented Gradients algorithm applied to a training image



The sizes of the cells are used in the HOG algorithm in the aircraft recognition program are 16x16px. The blocks consist of only one cell as this is the default value. If the SVM’s predictions were not accurate enough, the block sized can be increased if necessary to take make full use of contrast normalisation.

### Feature vector

A feature vector is a vector containing information describing an objects important details and characteristics. It can describe features and characteristics of a small area of an image or can describe an entire image. A feature can take many forms such as average density, centroid (centre of image where all medians intercept), pixel colour count, gradient magnitude and orientation.

The feature vector produced by HOG in the aircraft recognition system contains number of blocks per row, number of blocks per column, number of cells per row, number of cells per column and number of orientations and, orientation of every cell processed. Take an image with dimensions of 400x400px:

|  |  |  |
| --- | --- | --- |
| **Image size: 400x 400px (RGB)** | **Calculation** | **Length** |
| **Image length flattened** | (400x400)x3 | 480000 |
| **Feature Vector** | (400x400)/16 | 10000 |
| **Percentage difference** | (480000-10000)/480000\*100 | -97% (470000) |

By applying the HOG algorithm before training, we can not only reduce the amount of information passed to the SVM by roughly 97% but, increase the accuracy of the data passed to it. This plays a large part in computational complexity as the training and classification of a single image is 48 times faster. It is argued the application of HOG can add complexity to the system however, HOG will not increase the complexity by the 48 times it was reduced. Figure 4 shows the difference between an aircraft and ground feature vector.

A screenshot of a computer

Description automatically generated

Figure : Ground feature vector (Red), Aircraft feature vector (Blue)

Explain figure 4 here

## Machine learning

Machine learning is an application of Artificial intelligence that allows computer to automatically learn and improve using statistical models, algorithms and access to data without explicit programming. Machine learning has multiple processes, the first being training. Training data is provided to the Machine learning model so it can observe the data and look for patterns. The more data fed to a machine learning model, the greater the understanding thus increase in accuracy. However, to get accurate and reliable results, the ML model must be passed reliable data otherwise its ability to predict correctly will be greatly flawed.

After the training stage is completed, the ML model can be passed test data. The model can evaluate the test data and return a prediction. In our case, whether the image contains an aircraft or is purely ground. After predictions are made, if correct, can be added to the training set to improve accuracy.

Optimisation is the final stage and can often make or break a ML models accuracy. Decision boundaries and functions created by the model can be altered by the input variables such as C, gamma and kernel. These essentially define the way the model can make decisions. While ML can analyse vast quantities of data and quickly yet accurately make predictions, it requires a large amount of time and resources to train it correctly.

### Supervised learning

Supervised learning is the process of providing a machine learning model with labelled training data. The ML model produces an inferred function to make predictions about the test data.

A set of N training examples in the form **{(x1, y1),…, (xN, yN)}**, **xi**representing the feature vector provided by HOG pre-processing, **yi** being the label to accompany the data that informs the ML model of its class, the ML model will produce a function in the form **g : X -> Y** where **X** is the input space and **Y** is the output space. The function **g** is referred to as the hypothesis space. Many ML algorithms are probabilistic meaning **g** takes the form of a conditional probability model where **g(x) = P(y|x)**. The model returns its prediction based on the highest probability.

To contrast, unsupervised learning allows a computer to draw its inferences from training data. This allows the ML model to group classes as it sees fit. This can yield unexpected results however after extensive training, become very accurate.

### Support vector machines

A Support vector machine is a supervised learning algorithm that analyses data. The common uses for SVM’s are classification and regression analysis. SVM’s are commonly used in real world problems such as protein classification, image classification, text classification and classification of handwritten characters.

A SVM is provided training data in the form **{(x1, y1),…, (xN, yN)}**, **xi**representing the feature vector provided by HOG pre-processing, **yi** being the data’s label. In our case, **xi** represents images of ground and **yi** is either a 1 or 0 dependant on the human classification of the image. The SVM plots a graph and plots feature vectors as individual points after kernel processing. After all feature vectors in the training set are plotted, a hyperplane is plotted between classes. This graph can exist in 3-Dimensions the hyperplane can exist as any shape. This can be seen as something similar to a line of best fit. An SVM can have as many hyperplanes as it wishes to group classes. The more hyperplanes present, the more complex classification becomes decision boundaries are challenged. Figure 5 shows an example of an SVM after training is completed.



Figure : An example of a support vector machine [2]

To make predictions, the SVM is provided with an unseen test set of feature vectors. This could be from a large image or smaller image. The SVM plots the feature vector and dependant on its position relative to the decision boundary, returns a probability and prediction.

#### Kernel

The kernel is a set of mathematical functions that control the way an SVM classifies data. There are several commonly used kernels such as linear, non-linear, polynomial and Radial Basis Function. The kernel often defines the shape of the hyperplane

This is the mathematical equation of the RBF kernel. RBF is commonly used as the default algorithm for ML models as it needs no prior knowledge about the dataset. Figure 6 shows the RBF kernel equation.

Laplace RBF kernel equation

Figure : RBF kernel equation

Feature vector **xi** and label **yi** is passed to the RBF kernel function. After computation is completed, the kernel returns coordinates to be plotted by the SVM. This equation is computed for every feature vector provided to the SVM during training and testing stage.

#### Hyper parameters and decision boundaries

The SVM can take parameters that help define decision boundaries. For many applications, the default values provided by the algorithm can be used. However, dependant on the type of data and shape, these may need to be tuned.

**C** is penalty parameter of the error term. This controls the distance of the margins from the hyperplane that ignores features that may have some form of error during training. Margins are shown either side of the yellow dotted line in Figure 5

**gamma** is the coefficient used in the RBF kernel along with many others. A high gamma value leads to high bias and low variance where as a low gamma variance leads to a small bias and high variance.

A hyperplanes margin is regarded as soft when it ignores features close to the hyperplane. This often leads to a better fit and overall increase in accuracy in many cases. A margin is hard when a it ignores no features and tries to divide classes perfectly down the middle.

#### Tuning

To optimally tune hyper parameters, a grid search algorithm is used. Grid search takes a set of seemingly random values and passes them to the SVM. In our case, C and gamma values are generated by decrementing from 10 and 1 respectively. The SVM is trained using its own data set and then tested. The grid search algorithm cross checks the prediction from the SVM with the human classifications obtained and data collection. Based on the highest scoring combination of C and gamma with the given test set, the algorithm chooses the optimal input parameters. We now know the values of C and gamma are optimised specifically for the purpose of classifying aircraft and ground images.

*INSER*T TUNING DATA HERE

#### Under and Over fitting

Under and over fitting occurs when the hyperplane divides classes too closely or too loosely. Figure 7 shows different types of fits.



Figure : SVM hyperplane fit examples [3]

When a ML model is over fitted, it makes it difficult for future predictions to be accurate as they must fit a very specific criteria therefore often fail to reach the correct result. Overfitting occurs when a hyperplanes margin is hard.

Under fitting occurs when the ML model fails to gain an understanding of the underlying structure. This can be a result of selecting the wrong kernel with the shape of the data provided at training.

### Cross validation

k-Fold cross validation is a resampling technique used to evaluate a ML models accuracy and performance on a limited data set. It is often used to gain an insight into expected results with an unseen data set. The technique has a single parameter referred to as k. Parameter k defines the number of groups the total data set will be split into. If k = 10, the process will be called 10 fold cross validation. Cross validation is a popular method because it is simple to understand and usually results in less biased and optimal estimate of the model as opposed to a simple train/test split. It also shows how accurate the training data is. If the cross validations core of the training data is low, then the model may struggle to evaluate unseen test data.

The steps for cross validation are as follows:

1. Randomize the datasets order
2. Split dataset into k groups
3. For each group create

3a. Retain group for test

3b. Train model using remaining groups and evaluate accuracy with test set

3c. Retain the evaluation score and discard model

1. Summarize the skill of model using k evaluation scores

In most cases, the scores are averaged to create the most realistic score for the training set. To evaluate the performance of the aircraft recognition program, I cross validated the standalone and realistic data set.

#### Standalone dataset

|  |  |  |  |
| --- | --- | --- | --- |
| **10 fold CV of standalone dataset** | **Lowest Fold** | **Highest Fold** | **Average** |
| **400 x 400 px** | 0.94 | 1 | 0.99 |

Results of 10-fold cross validation of standalone dataset show that the accuracy of the ML model and data set is incredibly high at 99%. The variance between lowest and highest fold is 6% meaning the model is incredibly accurate. 8 folds achieved 100% during this test.

#### Realistic dataset

|  |  |  |  |
| --- | --- | --- | --- |
| **10 fold CV of realistic dataset** | **Lowest Fold** | **Highest Fold** | **Average** |
| **800 x 800 px** | 0.37 | 0.99 | 0.74 |
| **400 x 400 px** | 0.66 | 0.96 | 0.75 |
| **200 x 200 px** | 0.66 | 0.67 | 0.67 |
| **100 x 100 px** | 0.60 | 0.61 | 0.60 |

Results of 10-fold cross validation of realistic dataset show that when image size is increased, the average cross validation score increases however the variance between highest and lowest fold also increases. As variance was originally higher, images had to be deleted that didn’t confirm to dataset rules.

# Technical Achievement

## Graphical user interface

The graphical user interface is the starting point of the program. The user is able to select classification method and options for the image search classification process. Figure 8 shows the windowed GUI that the user is shown the following execution.

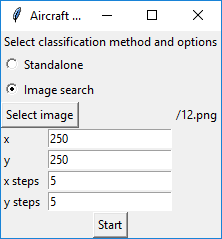


Figure : GUI shown to user after program execution

The user must first select classification method. Standalone classification requires no other input from the user. The user can now press Start to begin the training and classification process.

If the user wishes to search for aircraft in a larger image, the image to search must be selected. Clicking the select image button allows the user to navigate to the location of the image they wish to classify and select it. Figure 9 shows the window used to select images from files system

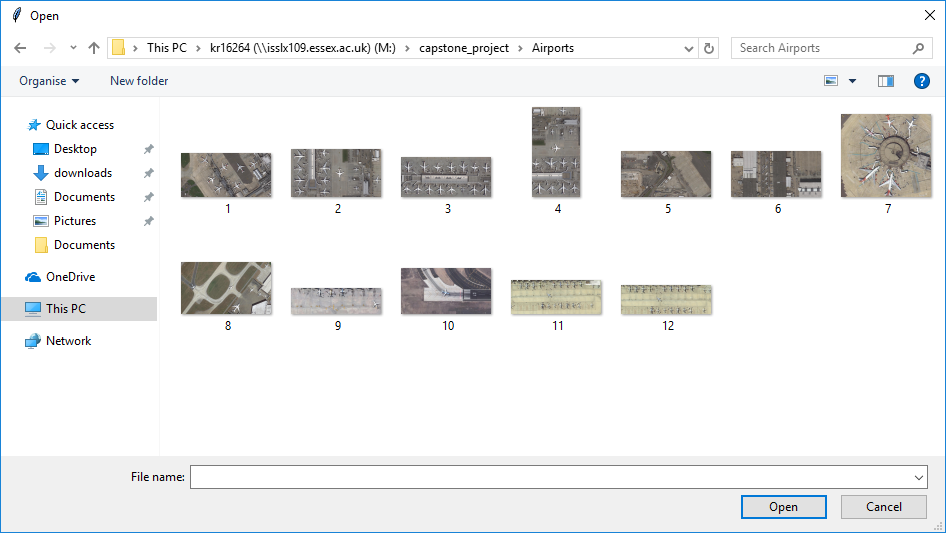


Figure : File selection dialog

After image to search is selected, the user must specify;

* x- The x dimension of the search area in px
* y- The y dimension of the search area in px
* x steps- The amount of moves along the x axis
* y steps- The amount of moves along the y axis

From this data, the program selects suitable x and y step values to ensure image boundaries are met.

After the user has selected search criteria, the Start button can be pressed to start the image search process.

## Standalone classification

Standalone classification is the recognition of aircraft images where images from a data set are selected at random for training and classification. There are 200 images in the standalone dataset. 100 images of ground and 100 images of aircraft. 5 images of aircraft and 5 images of ground are selected for testing. The remaining 180 images are used to train the ML model.

The images are put through usual process of parsing, HOG pre-processing before being passed to the SVM. After training is complete, test images are passed to the SVM. The SVM returns a set of predictions for the images tested. To show the user the final results, the test images are displayed in a grid with accompanying predictions and probabilities. The probability shows how sure the ML model is that the image belongs to a particular class.

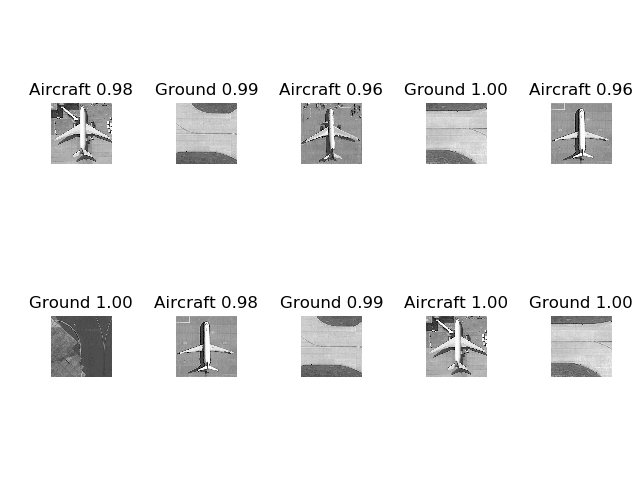


Figure : Results from standalone classification

The results in Figure 10 show that the program correctly identified all ground and aircraft. The accuracy of standalone classification in all test cases is 100%.

## Large image search

Large image search requires a more complex data set. The data set has more samples and the data in the image is more representative of aircraft in a real world setting. The basic principle behind large image search is that the user defines search dimensions to the program and the program iterates the image searching in small portions for aircraft. The execution behind it is slightly more complicated.

SPEAK ABOUT TRAINING HERE

The image is first divided into several smaller images. The image sizes are defined by the users input. If the user has given search settings that mean the images overlap, parts of the large image will be processed more than once. If the user has supplied settings that mean the search boxes do not overlap, some areas will not be classified correctly.

The x and y steps are calculated as such

x\_step = image\_width – x\_search\_size / x\_steps

y\_step = image\_height – y\_search\_size / y\_steps

This ensures that no area of the image is missed.

The smaller images are parsed, HOG pre-processed and passed tso the SVM for evaluation

Heat maps

Search results

Accuracy

Reasons for many boxes

Show best examples

# Technical documentation

Tech documentation from gitlab

# Project Planning

## Momentum

## Adapting to change

## Identifying and dealing with risks

## Achievement

## Performance

## What have I learnt?

# Conclusions

## Discussion

# References

[1] <https://docs.scipy.org/doc/numpy-1.13.0/user/whatisnumpy.html>

[2] <https://medium.com/deep-math-machine-learning-ai/chapter-3-support-vector-machine-with-math-47d6193c82be>

[3] <https://medium.com/greyatom/what-is-underfitting-and-overfitting-in-machine-learning-and-how-to-deal-with-it-6803a989c76>

# Tables, Graphs, Figures and Equations

# Appendices

## Sustainability

## Legal

## Ethical

## Intellectual property