Summary

Autofluorescence endoscopy (AFE) can be used to identify areas of the esophagus, which are undertaking a transformation into a condition known as Barrett’s esophagus. This transformation can lead to neoplastic lesions, which can become cancerous. The number of patients affected by neoplastic lesions linked to Barrett’s esophagus has been on the increase for several years [1]. Previous research [14] has been done to see if a classifier can be developed to distinguish between regions of normal mucosa (tissue) and areas of high-grade dysphasia (abnormality of tissue development). The software developed for this task requires user inputted expert knowledge to choose regions within images of the esophagus that represent areas of both ‘normal’ and ‘abnormal’ tissue for use in the classification algorithm. The aim of this project is to investigate if computer vision techniques can be used to gather the ‘normal’ tissue areas from the image automatically with no user interaction and still achieve similar accuracy levels as the original user submitted regions. This report presents the development and findings of three different techniques that have been created to perform this ‘normal’ tissue selection. Their development, implementation, and results are presented in this report. Also included is a section that outlines possible improvements to the project that could be undertaken.
Acknowledgements

I would like to thank all of the people who have given me help and assistance during this project. I would like to thank my supervisor Derek Magee for his guidance during the project of which I am sure made the project possible. I would also like to give thanks to Dr Venkataraman Subramanian for his research into Autofluorescence endoscopy and the data provided by him and Leeds General Infirmary Gastroenterology department. Also Vilius Narbutas for his work on the ‘Tumor Detection’ Program.
# Contents

Chapter 1 – Introduction

Introduction to Problem ................................................................. 1
Aims and Objectives ........................................................................ 1
Relevance to Computer Science ...................................................... 1

Chapter 2 – Background to problem

Barrett’s Esophagus ........................................................................ 2
Autofluorescence Endoscopy ............................................................. 3
Endoscopy ....................................................................................... 3
Autofluorescence ............................................................................ 3
Autofluorescence Endoscopy ............................................................. 4
Computer-aided Diagnosis ............................................................... 4
Dr V. Subramanian’s Research ......................................................... 5
Possible Solutions to problem ......................................................... 6

Chapter 3 – Project Development ..................................................... 7
Previous Work ................................................................................ 7
Planning Tools ................................................................................. 7
Methodology ................................................................................... 8
Project Aim and Objective Development ........................................ 9
Changes to Original Project Plan ...................................................... 11
Development Tools ......................................................................... 12
Language ....................................................................................... 12
Software ....................................................................................... 12
Data source .................................................................................... 12

Chapter 3 – Methods .................................................................... 13
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1</td>
<td>13</td>
</tr>
<tr>
<td>Method 2</td>
<td>14</td>
</tr>
<tr>
<td>Method 3</td>
<td>15</td>
</tr>
<tr>
<td>Chapter 4 - Evaluation</td>
<td>17</td>
</tr>
<tr>
<td>Evaluation Criteria</td>
<td>17</td>
</tr>
<tr>
<td>Data Input</td>
<td>17</td>
</tr>
<tr>
<td>Results</td>
<td>17</td>
</tr>
<tr>
<td>Chapter 5 – Conclusion</td>
<td>20</td>
</tr>
<tr>
<td>Findings</td>
<td>20</td>
</tr>
<tr>
<td>Limitations</td>
<td>20</td>
</tr>
<tr>
<td>Possible Future Development</td>
<td>21</td>
</tr>
<tr>
<td>Bibliography</td>
<td>23</td>
</tr>
<tr>
<td>Appendix A</td>
<td>28</td>
</tr>
<tr>
<td>Appendix B</td>
<td>30</td>
</tr>
<tr>
<td>Appendix C - How ethical issues are addressed</td>
<td>31</td>
</tr>
<tr>
<td>Appendix D – Project Gantt chart</td>
<td>32</td>
</tr>
</tbody>
</table>
Chapter 1 – Introduction

Introduction to Problem

Barrett’s esophagus (BE) is a condition that affects the upper digestive tract. It can be found in as many as 15% of cases where the patient complains of heartburn [2]. The condition causes the tissue lining of the esophagus to change to cells normally found lower down in the gastrointestinal tract. Links have also been detected between BE and Adenocarcinoma which can be a fatal form of cancer, [3] estimates the risk of developing Adenocarcinoma in patients diagnosed with BE at 10%. Diagnosis relies on an invasive endoscopy examination undertaken by a professional Gastroenterologist. This is then followed by biopsy to confirm the results of endoscopy procedure. This can be both time consuming and expensive [4]. Autofluorescence is a technique used to improve the lesion detection rates of endoscopic screening but is associated with a high false positive rate [5]. As a result, computer vision and machine learning techniques have been applied in [4] to create a system which can be used to automate the diagnosis of BE and reduce the false positive rate. This system still relies heavily on the experience and training of the Gastroenterologist.

Aims and Objectives

The aim of this project is to be able to automate the selection of regular esophagus lining mucosa (tissue) from images of patients with possible BE with the aim of maintaining the levels of accuracy for classification reported in [4]. To achieve this goal, the project has these main objectives:

1. To automate the selection of ‘normal’ mucosa within Autofluorescence endoscopy images.
2. To produce multiple techniques for mucosa selection.
3. To compare the accuracy of new and existing techniques at classification on an unseen test set.

Relevance to Computer Science

Computer science is a very large field of study that includes research areas such as Artificial intelligence and computer graphics. The fields that are relevant to this project are Computer Vision and Machine Learning. This project will be dealing with image manipulation and analysis as well as working with machine learning algorithms. Medical diagnosis relies on identification of features that represent specific conditions therefore if these features can be detected and quantified by a system then diagnosis could be performed by that same system.
Chapter 2 – Background to problem

Barrett’s Esophagus

Barrett’s esophagus (BE) is a condition that is estimated to affect 1.6 to 6.8 percent of people [6]. This condition describes the transformation of regular esophagus mucosa to tissue that is normally found further along in the gastrointestinal tract. The gastrointestinal tract includes the mouth, throat, stomach and lower intestine. The esophagus is the tube connecting the mouth and stomach.

While the main causes of Barrett’s esophagus is still unknown it has been linked to Gastroesophageal reflux where the repeated exposure to stomach acid damages the cells of the lower esophagus [7]. This condition is considered very dangerous because it can be a pre-condition to a form of cancer known as esophageal Adenocarcinoma which is estimated to kill 15,210 patients in 2013 [8].

The different levels of transformation that the mucosa can go through have been classified into 5 levels of different dysplasia (abnormal growth) stages [9] ranging from non-dysplasia to invasive neoplasia (abnormal mass of tissue). The treatment recommended to the patient will be based on which category the tissue samples are graded. Multiple pathologists will verify the classification and diagnosis.

The current methods of diagnosis for this condition are as follows. A doctor will first perform an endoscopic examination where a small video camera mounted on a moveable stalk will be inserted into the throat through the mouth. The doctor will then use the images from the camera to examine the lining of the esophagus. If the doctor identifies a region that they consider to be of risk, a biopsy of that area will be arranged. Random tissue samples from the test area will be taken for analysis. The effectiveness of the biopsies currently undertaken using current medical standards is under question, as only 4-6% of the identified area will be sampled [10]. After the area has been sampled a pathologist will examine the resulting specimens under a microscope. To a trained eye the difference between the cell makeup of regular mucosa and tissue consistent with BE is quite noticeable. This can be seen in [Fig 3]. A pathologist will examine the samples to determine the classification.

Once a patient has been identified and diagnosed with BE they will be expected to attend regular annual checkups to monitor the development of the BE lesions. In cases where the dysplasia may become life threatening the patient have to undergo different treatments based on the unique development of their specific lesions. This can include regular surgery to remove lesions or in cases
where surgery is not an option, chemotherapy may be recommended. Not all patients with BE will develop esophageal cancer. Patients diagnosed with esophageal cancer have an 85% mortality rate [11] and most pass on within a year of diagnosis [12]. For these reasons early diagnosis and regular examinations are vital to maintaining the health of patients.

**Autofluorescence Endoscopy**

**Endoscopy**

Endoscopy is the method doctors use to examine the internal structures and cavities within the body. This is achieved by placing a fibre optic video camera inside the body generally through the nose, throat or rectum. The operator can control the orientation of the camera while it is inside the body. The camera even has the ability to curl up and look back at its self. This is particularity useful in situations such as looking at the roof of the stomach where the camera has to come down into the stomach through the esophagus before it can be rotated back up to look at the entrance it has just used to enter the stomach. Modern endoscopes use fibre optic cable to act as an extended flexible lens. The first endoscopes to use fibre optics were developed by Basil Hirschowitz and Larry Curtiss in 1957 [13].

As discussed earlier in the report the endoscope is used to allow the doctor to examine the esophagus and stomach. Most endoscopes work under white light and have a LED or small light on the end of the camera to allow the area under inspection to be visible inside the body where no naturally occurring light exists.

**Autofluorescence**

Autofluorescence is a known phenomenon that occurs in living tissue. When exposed to light the cells in tissue absorb the light and reflect what is not absorbed back to the camera. Using different wavelengths of light it is possible to get diagnostic information from the resulting images. [14] Explains that “Cells contain molecules, which become fluorescent when excited by UV/Vis radiation of suitable wavelength.” The research by [11] shows “Changes occurring in the cell and tissue state during physiological and/or pathological processes result in modifications of the amount and distribution of endogenous fluorophores and chemical–physical properties of their microenvironment. Therefore, analytical techniques based on auto-fluorescence monitoring can be utilized in order to obtain information about morphological and physiological state of cells and tissues.” This means that cells that have undergone changes will emit different colours and it can all be done in real time and without any staining to the tissue, [14, 15] displays this ability.
**Autofluorescence Endoscopy**

Autofluorescence endoscopy (AFE) combines the functionality of Autofluorescence and endoscopy to provide a tool for doctors in the field of Gastroenterology.

AFE images are produced by exposing the area of interest to excitation light. Then a filter is placed in front of the camera lens to block excess reflected excitation light. The resulting light given off by the mucosa can be assigned to the RGB colour scale to produce an image that shows the region of interest (ROI) in detail [16].

[Fig 1] shows an image of a patient’s esophagus in both regular white light mode and AFE mode. Under autofluorescence inspection green areas represent non-dysplasic mucosa and the areas that are magenta coloured represent possible dysplastic mucosa. It is important to note at this point that early studies into the effectiveness of AFE detection by doctors had the effect of increasing lesion detection rates from 53% to 90% compared to inspections on the same image under white light. This level of detection came at the price of a false positive detection rate of 81%[17]. Similar results were achieved in another multi-centre study [18].

AFE has shown to be effective at detecting lesions in cases of BE, laryngeal cancer and gastric cancer [19-22].

**Computer-aided Diagnosis**

Computer aided diagnosis (CAD) is a field of study shared by both computer science and medicine. Research into the ability of computers to provide medical diagnosis began as early as the 1960’s with unsuccessful results [23-26]. Initial development aimed to produce computers that could automate diagnosis without human interaction. This field of research experienced a paradigm shift in the 1980’s from its initial roots in automated diagnosis to a more secondary role offering an automated ‘second opinion’ that professionals could use to accompany their diagnosis [26]. Many of the features that medical personal look for in an image can be learnt by a computer system that combines the disciplines of machine learning and computer vision. Lots of research into the effectiveness of using computer systems in this manner has been undertaken [28-31,4]. CAD has many possible applications. It can be used to assist diagnosis in Detection of Clustered microcalcifications on Mammograms (breast cancer) [32], colonic polyps [33] and pulmonary nodes
The high-risk nature of some of these conditions means that it is important that this technology is utilized because early diagnosis of symptoms can decrease fatality rates. The ‘second opinion’ offered by CAD systems offers unbiased numerical analysis that can be used to improve detection rates and give informative data when used in conjunction with a trained doctor. The use of this technology is most common in the field of radiology.

**Dr V. Subramanian’s Research**

Dr Venkataraman Subramanian is a doctor at the Leeds General Infirmary working in the gastroenterology department. His research [14] forms the foundation of my project. The research done was an investigation to see how effective a CAD system could be at reducing false positive rates of neoplasia in AFE endoscopy examinations. As mentioned earlier in the report AFE is effective at neoplasia detection but comes with a high false positive rate [17-18].

Images of patients with BE and suspected neoplastic regions were examined with AFE and still images were captured of suspected regions. Biopsies were performed to determine the level of transformation the cells had undergone. For a machine learning system to make a classification it must have some quantitative data that can be evaluated. The chosen measures in [4] were autofluorescence intensity (AFEI) and color contrast index (CCI). [4] defines AFEI as the “ratio of average red to green channel grayscale value of lesion compared to background”. CCI is defined by [34] and represents the perceived colour difference by the gastroenterologist [4]. These features were chosen from a list of possible variables because they offered the highest information gain [4].

Dr Subramanian used these images to train a classifier to partition the two sets of images. The two sets are cancer and non-cancer. The term ‘cancer’ is used to describe any area that could be classified as high-grade dysplasia or higher in the Vienna classification [9]. The program used to train the classifier is known as WEKA [36] it is a suite of machine learning algorithms made available under the GNU general public license. The algorithm chosen to create the classifier is called J48, it is a decision tree closely related to Quinlan’s C4.5 algorithm [37]. This sort of classifier is known as a supervised learning algorithm and requires data that has been already classified and is used to set the thresholds of different decision variable within the decision tree. For this research 82 images were chosen containing a range of AFE identified lesions. [4] reports that the trained algorithm achieved 90.2% accuracy on the training data set when used with 10-fold cross validation. We identified earlier that unassisted AFE comes with a high false positive rate <70% [16-17]. This study
shows a drop in this false positive rate from 70% to 16% when evaluated on a test data set of 164 previously unseen images.

This research shows that it is possible to achieve an improvement on lesion detection rates compared to professional diagnosis alone. This sort of technology can act as an aid for researches and medical staff. Performing multiple biopsies can be both expensive and time consuming for both staff and patients. Having a prescreening measure such as provided by [4] can help to reduce pressure on all parties.

There are a couple of issues associated with this current research. The current approach relies on a lot of user input as the system components are all comprised of separate applications. This makes the process of classifying an image very time consuming and requires a user who is familiar with the process. Also the selection of the regions of interest within an image requires the expert knowledge of a gastroenterologist. These regions are selected by hand. The automation and consolidation of this system is of great interest to researches.

**Possible Solutions to problem**

The problem of choosing appropriate regions for evaluation within image is a problem that can be solved by methods used in computer vision. In the image there must be some collection of distinguishable features that a gastroenterologist will look for. These same features can be recognized by a computer to assist in many situations.

One feature that can be used to recognize differences in sections of images is image texture. Image texture can be calculated to give us a quantitative value of the spatial arrangement or intensity of colour for a given region within an image. This could be used to detect regions within an image that match or are similar in texture to neoplastic lesions. Due to the fact that dysplasia is not normally uniform in its distribution with mucosa this could be quite difficult to implement effectively [14].

Another possible computer vision tool that would be useful for this problem is image segmentation. Image segmentation takes a digital image and divides the image into bordered subsections based on chosen evaluation criteria. This could be strong colours or based on edge detection. The resulting image will be a collection of these sub images. These sub images could be used for region selection.
Chapter 3 – Project Development

Previous Work
Dr Subramanian’s research [4] and others [31-34] have shown it is possible to combine areas of research that are of interest to computer scientists such as machine learning and computer vision to practical fields of medicine with beneficial results. As a result of the findings of Dr Subramanian the University of Leeds was commissioned to produce a program that could incorporate the classifier from [4] to take video or still images and perform a classification all within a single system. Vilius Narbutas, an undergraduate intern student at the university, produced the system over the summer of 2012.

The program shown in [Fig 2] allows for direct interface with the AFE equipment through the video input library [38]. The program could be used in real time during an endoscopy examination and allowed the examiner the ability to store still images and perform classifications on suspected lesions. The results can be exported as a .csv (comma separated values) file and stores the marked regions of interest as well as the classification prediction and decision feature values. The program also facilitates the loading of previously saved regions of interest and can also implement different classifiers in the .json (JavaScript Object Notation) format that are compatible with the waffles machine learning library [39]. The original classifier created by WEKA was converted to make it waffles compatible. The ‘Tumor Detection’ program is a user-friendly graphical interface that provides all the functionality required to act as an aid in AFE diagnosis of patients with BE. The program still requires the manual markup of 2 areas.

Planning Tools
The project has a finite amount of time that is available for planning, development, implementation, testing and evaluation. For this reason it is required to utilize both time and project management tools and methodologies. Good project management requires the identification of the tasks required to achieve a goal. How these tasks interact with each other in the form of dependencies and critical paths define the length of the project and allows for monitoring of progress throughout. Microsoft project [40] was used to monitor the subtasks of the project and keep a record of progress; it also
provides a platform for visualization of the entire project through the production of Gantt charts. The software was provided free of charge for academic use though Microsoft’s DreamSpark program.

Appendix D contains the full project Gantt chart that contains the identified subtasks of the project along with information such as expected completion time, resource allocation, critical path analysis, task dependencies and task completion percentages.

The project culminates in this report and as such will contain information from all stages of the project lifecycle. To ensure the accuracy of the findings and issues experienced during the project a blog has been maintained to record issues, milestones and ideas [41]. The blog also allows the developer a space to record ‘To Do’ lists that help to manage time and provide simple achievable objectives that maintain productivity. It assists the developer in picking up where they left off during the previous work session e.g. Friday-Monday.

**Methodology**

The bulk of this project was spent on software development. Proper planning of a software development project requires commitment to a methodology of design and implementation. Several different methodologies exist to cater to different project needs and requirements. The methodology chosen for this project is Rapid Application Development (RAD) [42].

RAD was chosen because of the unique needs of the project made it the most suitable methodology. It allows for the rapid design and development of prototype builds which meet more of the project requirements as the number of completed builds increases this fit well with the modular nature of multiple method testing. This methodology is useful when the full project requirements and subsequent design are not initially known as it allows for a fluid design and implement process until a build is created that meets all of the clients/developers requirements. This is a strong contrast to other methodologies such as the waterfall method that require the design to be fully realized and complete before implementation.

The waterfall approach to software development requires the division of the project into dependant sections that cannot be revised once completed and must be completed in a defined order. The image below illustrates a typical waterfall process. This process is documentation oriented with the expected ability to be able to add and remove staff on the project without losing any of the vital knowledge to continue development as new team members should be able to familiarize themselves with the project from the documentation [41]. As there will only be a single developer working on this project the benefit from such practice would be minimal.
This type of approach can help to reduce issues with development as bugs discovered in design are much easier to remove in terms of time and effort than those discovered during testing [43]. This benefit can become negligible when new requirements become available after the design phase, as there is no option to alter the design specification outside of the design phase without beginning the whole waterfall process again. RAD is the most suitable for this situation as a useable product can be presented to the client for each requirement listed and progression to testing and evaluation stages can be done even if all of the project requirements have not been implemented by the end of the allotted project completion date.

**Project Aim and Objective Development**

It was hypothesized that it could be possible to further automate the system in question to reduce the systems reliance on a human operator. My project set out to investigate what possible improvements could be made to the system to achieve this goal. From the outset it was the projects aim to facilitate this automation while still maintaining sufficient accuracy of results that the system would remain a viable tool for diagnosis.

![Project scope mind map](image)

As part of the project development it was important to identify what possible areas of the program and classifier could be improved. The image above shows a mind map of the possible areas of improvement and investigation. This diagram was created early on in the project and illustrates the scope of early development ideas. The two main development directions that were identified from discussion were improvement to the classifier or ROI selection methods.
Possible improvements to the classifier were considered. The prediction of a ROI within an image was solely based on the classification provided by a trained decision tree [14]. An initial idea was to test the effectiveness of different classifiers based on the same training data that was used to train the J48 classifier. This idea was expanded to investigate if it could be possible to implement multiple classifiers to examine the ROI and provide an aggregate of those predictions to give a final overall prediction. It was hoped that by using multiple classifiers it would be possible to remove outlying results that would affect the accuracy of the final prediction. For example if three classifiers were used to evaluate the prediction and two out of three gave a positive prediction for ‘cancer’ the final prediction would be ‘cancer’ while a decision based on the single classifier that gave a negative prediction would affect the accuracy of results.

The other direction of development involved investigation into the collation of the ROIs used by the classifiers. In the current implementation of the ‘Tumor Detection’ program two separate ROIs need to be identified by hand by the user within the chosen image. These two ROIs need to correctly include the area that is to be investigated (This is normally a magenta region within the image) and a region which represents regular non-dysplasic mucosa ( normally green in colour under AFE). The selection of both these regions requires a trained AFE practitioner to identify ideal regions to represent both requirements. If the images are being taken from live video this can have the effect of increasing the time of the endoscopy examination as doctors will have to pause the procedure to highlight these areas on the program before being able to continue the procedure. Endoscopy is an invasive technique and reducing examination length while maintaining the quality of its findings could contribute to increased patient happiness.

The automation of ROI selection was chosen as the objective for my project because it was felt that this would give the best reduction in user interaction and if suitable methods were developed for the ROI selection it will still be possible to achieve rate of accuracy similar to those provided by manual hand selection.

This approach was then subdivided into to possible development schemes. First the automation of detection of regions, which could be, identified as possible dysphasic mucosa. This would allow the user to only have to provide a suitable image to the system and the system could then decide which areas within the image would be most suited to investigation. The second would be the automation of region selection of tissue determined to represent ‘normal’ mucosa. Implementation of this feature would allow the user to concentrate on providing a suitable region for investigation reducing user interaction and expert knowledge requirements. For the rest of the report the investigative region will be referred to as the ‘foreground’ region and the non-dysplasic mucosa region as
‘background’. It was decided because of time constraints that the automation of background region selection would be developed further as the main objective of the project.

### Changes to Original Project Plan

The original plan gave a timescale that reflected the entire length of available time for the project. As the project developed this had to be revised because of development environment issues. The project suffered setbacks because the libraries used to develop the ‘Tumor Detection’ were not documented or provided with the program source code. This meant a lot more time was needed to set up an appropriate development environment. This meant that the original project plan did not reflect the final plan timescale that was implemented and as such some aspects of the project had to be revised. One of these revisions was the removal of methods that would use image texture as part of the background selection. With the removal of unachievable targets within the timeframe a new project plan was required. This project plan correctly represented the final project that was undertaken. Another aspect of the project that had to alter during development was algorithm prototyping.

During the project it was discussed if the algorithms to be developed could be prototyped in Matlab code before implementation in C++ for integration with ‘Tumor Detection. This would form the basis for a feasibility study for each method. This also had to be removed in the revised project plan. The time left for the project to be complete and the use of two different languages prevented this aspect of the project from being achievable. As a result of this it was decided to instead continue with the RAD methodology and do all of the prototyping and implementation in C++.

Another revision to the project that occurred was the segmentation of quantified accuracy levels production. Originally this would have been done as part of single program. The program would have been able to take in histology results for each image and compare the results of each method on the trained classifier. All output would have been produced and displayed within ‘Tumor Detection’. This was not possible to implement in the final project as once again the revised project plan did not have enough available time or resources to deviate from the identified critical path of the project.

Due to compatibility issues with the libraries used in this project a revised classifier file was required to be used with ‘Tumor Detection’. The new classifier only differed from the original in its format as it was discovered during the project that it was incompatible with the version of waffles being used.
Development Tools

Language
As described in [Chapter 2] there are two languages that were discussed that could be used in the implementation of the background selection methods C++ and Matlab.

Matlab is a scientific programming language that specializes in numerical computing. It is popular with academic looking to perform matrix manipulations and algorithm prototyping, it also features interfaces for C++. Matlab was developed by MathWorks and is currently released up to version 8.1.

C++ was naturally chosen as the development language this is because it was the language used to write the ‘Tumor Detection’ program and as a result was the natural choice for continued work being carried out in this project. The first commercially available release of C++ was in 1985. It is one of the most popular programming languages in the world [44]. C++ features classes to allow object orientated programming but also allows for programming in different styles as it was designed to be a multi-paradigm language [45]. C++ interfaces exist for all of the libraries that are required for the functional use of ‘Tumor Detection’.

Software
‘Tumor Detection’ was created within an integrated development environment (IDE). IDE’s are software suites designed to assist in the creation of computer programs. They typically provide a large amount of features considered to be useful to a computer programmer. These features can include compliers, debuggers and visualization of graphical user interfaces (GUIs). The IDE used for the project was Visual studio 2010 [46]. This was chosen because it was the same IDE used in creating ‘Tumor Detection’ and the source code provided used Visual studio compatible file formats. It was assumed using the same IDE would reduce issues involved with project migration, as attached settings files included with the original project files would speed up set up of an appropriate development environment for the project.

Data source
The University of Leeds and Leeds General Infirmary Gastroenterology department provided all data for the project. Data was either made available though the project files used in the ‘Tumor Detection’ project or directly by Dr Venkataraman Subramanian. This includes 82 AFE images and 41 AFE images that had been marked by a gastroenterologist using the ‘Tumor Detection’ program. The markings for these images known as masks are stored in .csv files which accompany each image. These 41 marked images also included histological analysis results performed by professional pathologists at Leeds General Infirmary on the marked regions.
Chapter 3 – Methods

Each method in this section uses an algorithm to select different regions to be used as background, in an image, for classification. This is achieved by creating a mask, this is known as ‘dc’, that is a scale copy of the full size AFE image. Based on the colour of each pixel in the mask the corresponding pixel in the real image is added to the ROI. These areas were marked on the image as seen in [Fig 3]. Blue pixels are used to represent an area of suspected neoplasia. The yellow marked pixels are used as areas of non-dysplasic mucosa. Any pixel that retains its pixel values from the original image is considered too unsuitable and excluded from any further use within the algorithm. The images provided came bordered in black pixels with the AFE region in the center of the image.

Method 1

Method 1 will be referred to within this report as ‘Global’. Global is the first method that was added to ‘Tumor Detection’. It is the simplest of the background selection algorithms. The aim of this method was to see if useful background data could be extracted from the area of the image not part of foreground. Consider a set A and B where A is the AFE region within the image. Set B is a set of pixels within foreground. Background in this method will be the set A-B. This is the aim of the method, to include all pixels not part of foreground within the AFE region as the background. The background area should not include any pixels that will affect the accuracy of results. This is achieved in several consecutive steps.

1. Importation or manual selection of foreground region.
2. A nested loop is used to traverse each pixel in dc.
3. Each pixel must pass through a list of requirements to be assigned to background. The requirements remove pixels that are black (0,0,0), gray (126,126,126), blue (0,0,255), have a combined pixel value total greater than 630 or lower than 60. These thresholds were chosen to minimize the effect of outlier pixels that would damage the algorithms accuracy. A pixel coordinate threshold is used to reject pixels which correspond to information marking within the image but outside of the AFE area.
4. Foreground region is then Flood filled with blue pixels to give the final foreground ROI.
5. Once all suitable pixels have been assigned the appropriate colour mask each pixel is added to its ROI to be used for feature variable calculation.

An example of Global in operation is displayed in [Fig 4].
Method 2

Method 2 will be known within this report as ‘Box’. Box implements the inclusion of a bounding box (BB) of the foreground area. The aim of this method is to be able to include the tissue directly surrounding the foreground ROI in the image. This is based on the assumption that surrounding tissue will represent non-dysplastic mucosa, as suspected dysplastic mucosa will have been included in the foreground selection. The BB used comprises of a series of 4 lines that correctly partition the image into two sets. Set A which represents pixels outside the box and set B for pixels within the box. The box method only performs calculation on set B. The box method creates foreground and background ROIs from the pixels in set B. The bounding box used is enlarged by 5 pixels in each direction. This allows for all pixels surrounding the foreground area to be considered for inclusion in the background. The algorithm performs as followed.

1. Importation or manual selection of foreground ROI.
2. Calculate boundaries of foreground region.
3. Increase size of boundaries by 5 pixels on both x and y-axis.
4. Draw BB onto mask.
5. Flood fill internal space as yellow pixels (flood fill is bound by blue pixels).
6. Each yellow marked pixel is now checked for black, gray, blue and total pixel values as defined in Global. Any pixel that fails this check is marked red.
7. The BB is now marked as red to prevent inclusion in foreground calculations.
8. Foreground is flood filled blue to complete the ROI masking.
9. Each pixel within BB is stored according to its mask colour.

An example of Box in operation is displayed in [Fig 5].
Method 3

Within this report method 3 is referred to as ‘Mode’. Mode is an attempt to identify non-dysplasic mucosa for use in background selection from the mode colour present in the image. This is achieved by selecting the mode colour of the image and marking all occurrences of that colour in the image. For each marked point an 8-connected neighborhood region is created around each pixel. As long as these pixels pass the standard pixel checks they will be included in the final ROI. The algorithm performs as followed. Mode is the most complex of the background selection algorithms implemented.

1. A copy of the AFE image is created.
2. The number of colours in the copied image is reduced by a method known as colour quantization. This is achieved by a k-means clustering algorithm [47].
3. Every pixel in the new image is traversed and the standard checks of black, gray, blue, >630 and <60 are performed. If the pixel passes the colour is recorded in a colour bin object along with the position.
4. The colour with the most occurrences is selected. This provides a list of all the positions of pixel within this colour bin.
5. Each colour bin is then checked against pixel requirements.
6. Each position within the colour bin is marked yellow on the image mask.
7. An 8-connected neighborhood region is also checked around each pixel for suitable candidates in background selection. As long as one of these pixels has not already been added and it has passed the standard pixel checks it will be marked yellow on the image mask.
8. Before feature variable extraction the foreground area is flood filled to remove any marked pixels within the foreground area.

An example of Mode in operation is displayed in [Fig 6].

Fig 6 - AFE (left) and Mode (right)
Chapter 4 - Evaluation

Evaluation Criteria

The 41 images that have been graded according to the Vienna classification [9] will be used for evaluation. Each image in turn will be given a prediction by the classifier defined in [14]. One prediction per method will be completed for each image including the default manual selection performed by a gastroenterologist. The resulting prediction was then collated into a Microsoft Excel spreadsheet and compared to the histology results for each image. The overall accuracy of predictions made by the classifier for each background selection method will be generated. This included the overall percentage accuracy of each method, the sensitivity and specificity.

Data Input

To ensure the accuracy of the project results, it is important to keep as many of the variables the same between tests and to ensure that only the function the project is investigating is being altered. For this reason for each image being tested, the foreground region of interest will remain the same across each method. To facilitate this ability a function has been added to ‘Tumor Detection’ to allow for the importation of the foreground selected by the gastroenterologist in each image. The importation involves string manipulation of attached .csv files to extract the required data that represents the foreground region. The .csv files contain all of the data relating to a prediction. With this feature it will be possible to evaluate the effect of different background selection methods accurately.

Results

<table>
<thead>
<tr>
<th>Image No</th>
<th>Histology</th>
<th>Manual</th>
<th>AFI</th>
<th>CCI</th>
<th>Global</th>
<th>AFI</th>
<th>CCI</th>
<th>Box</th>
<th>AFI</th>
<th>CCI</th>
<th>Mode</th>
<th>AFI</th>
<th>CCI</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>yes</td>
<td>1.47</td>
<td>8.32</td>
<td>no</td>
<td>1.32</td>
<td>6.11</td>
<td>no</td>
<td>1.3</td>
<td>7.75</td>
<td>yes</td>
<td>1.8</td>
<td>10</td>
<td>no</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>1.5</td>
<td>10.6</td>
<td>no</td>
<td>1.22</td>
<td>10.9</td>
<td>no</td>
<td>0.9</td>
<td>18.7</td>
<td>no</td>
<td>0.7</td>
<td>48</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>yes</td>
<td>1.65</td>
<td>17.2</td>
<td>yes</td>
<td>1.72</td>
<td>22.8</td>
<td>yes</td>
<td>1.4</td>
<td>16.1</td>
<td>yes</td>
<td>1.5</td>
<td>18</td>
<td>yes</td>
</tr>
<tr>
<td>4</td>
<td>yes</td>
<td>yes</td>
<td>1.64</td>
<td>25.8</td>
<td>yes</td>
<td>1.37</td>
<td>10.1</td>
<td>no</td>
<td>1.1</td>
<td>9.79</td>
<td>yes</td>
<td>1.7</td>
<td>42</td>
<td>yes</td>
</tr>
<tr>
<td>5</td>
<td>no</td>
<td>no</td>
<td>1.22</td>
<td>32.6</td>
<td>no</td>
<td>1.06</td>
<td>5.82</td>
<td>no</td>
<td>1.2</td>
<td>18.1</td>
<td>no</td>
<td>0.7</td>
<td>10</td>
<td>no</td>
</tr>
<tr>
<td>6</td>
<td>no</td>
<td>no</td>
<td>1.32</td>
<td>11.4</td>
<td>no</td>
<td>1.23</td>
<td>8.57</td>
<td>no</td>
<td>1.1</td>
<td>1.17</td>
<td>yes</td>
<td>1.4</td>
<td>22</td>
<td>no</td>
</tr>
<tr>
<td>7</td>
<td>no</td>
<td>no</td>
<td>1.37</td>
<td>8.16</td>
<td>no</td>
<td>1</td>
<td>8.05</td>
<td>no</td>
<td>1.3</td>
<td>9.43</td>
<td>yes</td>
<td>2.6</td>
<td>15</td>
<td>no</td>
</tr>
<tr>
<td>8</td>
<td>yes</td>
<td>yes</td>
<td>1.82</td>
<td>10.3</td>
<td>no</td>
<td>1.21</td>
<td>5.23</td>
<td>no</td>
<td>1</td>
<td>3.16</td>
<td>no</td>
<td>1.1</td>
<td>23</td>
<td>no</td>
</tr>
<tr>
<td>9</td>
<td>no</td>
<td>no</td>
<td>1.39</td>
<td>7.82</td>
<td>no</td>
<td>0.95</td>
<td>4.34</td>
<td>no</td>
<td>0.9</td>
<td>4.54</td>
<td>no</td>
<td>0.6</td>
<td>29</td>
<td>no</td>
</tr>
<tr>
<td>10</td>
<td>no</td>
<td>yes</td>
<td>1.48</td>
<td>7.29</td>
<td>yes</td>
<td>1.38</td>
<td>14.5</td>
<td>no</td>
<td>1.2</td>
<td>7.11</td>
<td>no</td>
<td>1</td>
<td>29</td>
<td>no</td>
</tr>
<tr>
<td>11</td>
<td>yes</td>
<td>yes</td>
<td>1.44</td>
<td>3.99</td>
<td>yes</td>
<td>1.73</td>
<td>11.7</td>
<td>yes</td>
<td>1.6</td>
<td>14.4</td>
<td>yes</td>
<td>1.6</td>
<td>20</td>
<td>yes</td>
</tr>
<tr>
<td>12</td>
<td>no</td>
<td>no</td>
<td>1.8</td>
<td>11.7</td>
<td>no</td>
<td>0.88</td>
<td>1.89</td>
<td>no</td>
<td>0.8</td>
<td>5.1</td>
<td>no</td>
<td>0.9</td>
<td>36</td>
<td>no</td>
</tr>
<tr>
<td>13</td>
<td>no</td>
<td>no</td>
<td>1.31</td>
<td>28</td>
<td>no</td>
<td>1.24</td>
<td>4.33</td>
<td>no</td>
<td>1.1</td>
<td>16.8</td>
<td>yes</td>
<td>1.4</td>
<td>66</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>no</td>
<td>no</td>
<td>1.1</td>
<td>5.09</td>
<td>no</td>
<td>1.06</td>
<td>7.73</td>
<td>no</td>
<td>1.1</td>
<td>16.4</td>
<td>no</td>
<td>1.2</td>
<td>11</td>
<td>no</td>
</tr>
<tr>
<td>15</td>
<td>yes</td>
<td>yes</td>
<td>2.2</td>
<td>25.9</td>
<td>yes</td>
<td>1.92</td>
<td>26.6</td>
<td>yes</td>
<td>1.8</td>
<td>22.8</td>
<td>yes</td>
<td>2.2</td>
<td>29</td>
<td>yes</td>
</tr>
<tr>
<td>16</td>
<td>no</td>
<td>no</td>
<td>1.35</td>
<td>5.23</td>
<td>no</td>
<td>1.2</td>
<td>8.23</td>
<td>no</td>
<td>1.2</td>
<td>3.09</td>
<td>yes</td>
<td>1.4</td>
<td>17</td>
<td>no</td>
</tr>
<tr>
<td>17</td>
<td>no</td>
<td>no</td>
<td>1.29</td>
<td>2.51</td>
<td>no</td>
<td>1.17</td>
<td>3.03</td>
<td>no</td>
<td>1.2</td>
<td>19.4</td>
<td>no</td>
<td>1.2</td>
<td>16</td>
<td>no</td>
</tr>
<tr>
<td>18</td>
<td>yes</td>
<td>yes</td>
<td>1.37</td>
<td>13.8</td>
<td>no</td>
<td>1.31</td>
<td>13.8</td>
<td>yes</td>
<td>1.3</td>
<td>13.8</td>
<td>no</td>
<td>1.2</td>
<td>41</td>
<td>no</td>
</tr>
<tr>
<td>19</td>
<td>no</td>
<td>no</td>
<td>0.87</td>
<td>8.09</td>
<td>no</td>
<td>0.66</td>
<td>3.47</td>
<td>no</td>
<td>1.7</td>
<td>7.05</td>
<td>no</td>
<td>0.5</td>
<td>37</td>
<td>no</td>
</tr>
<tr>
<td>20</td>
<td>no</td>
<td>no</td>
<td>1.4</td>
<td>5.85</td>
<td>no</td>
<td>1.22</td>
<td>4.33</td>
<td>no</td>
<td>1.1</td>
<td>6.28</td>
<td>yes</td>
<td>1.7</td>
<td>8.3</td>
<td>no</td>
</tr>
<tr>
<td>21</td>
<td>no</td>
<td>no</td>
<td>1.42</td>
<td>5.98</td>
<td>no</td>
<td>1.33</td>
<td>7.46</td>
<td>no</td>
<td>1.2</td>
<td>17.7</td>
<td>yes</td>
<td>1.5</td>
<td>113</td>
<td>no</td>
</tr>
<tr>
<td>22</td>
<td>yes</td>
<td>yes</td>
<td>1.39</td>
<td>71.4</td>
<td>no</td>
<td>1.18</td>
<td>16.7</td>
<td>no</td>
<td>1.2</td>
<td>10.6</td>
<td>yes</td>
<td>1.7</td>
<td>99</td>
<td>no</td>
</tr>
<tr>
<td>23</td>
<td>no</td>
<td>no</td>
<td>1.26</td>
<td>40.8</td>
<td>no</td>
<td>1.13</td>
<td>15.2</td>
<td>no</td>
<td>1.1</td>
<td>11.4</td>
<td>no</td>
<td>0.9</td>
<td>10</td>
<td>no</td>
</tr>
<tr>
<td>24</td>
<td>yes</td>
<td>yes</td>
<td>1.34</td>
<td>13.7</td>
<td>yes</td>
<td>1.43</td>
<td>12.6</td>
<td>yes</td>
<td>1.2</td>
<td>15.3</td>
<td>yes</td>
<td>1.4</td>
<td>13</td>
<td>yes</td>
</tr>
<tr>
<td>25</td>
<td>no</td>
<td>no</td>
<td>1.44</td>
<td>3.21</td>
<td>no</td>
<td>0.88</td>
<td>10.1</td>
<td>no</td>
<td>1.1</td>
<td>6.51</td>
<td>no</td>
<td>1.1</td>
<td>6.5</td>
<td>no</td>
</tr>
<tr>
<td>26</td>
<td>no</td>
<td>no</td>
<td>0.95</td>
<td>5.24</td>
<td>no</td>
<td>0.99</td>
<td>2.1</td>
<td>no</td>
<td>1.1</td>
<td>5.81</td>
<td>no</td>
<td>1.1</td>
<td>86</td>
<td>no</td>
</tr>
<tr>
<td>27</td>
<td>no</td>
<td>yes</td>
<td>1.78</td>
<td>44</td>
<td>yes</td>
<td>1.66</td>
<td>10.3</td>
<td>no</td>
<td>1.4</td>
<td>3.88</td>
<td>yes</td>
<td>2.2</td>
<td>9.9</td>
<td>yes</td>
</tr>
<tr>
<td>28</td>
<td>yes</td>
<td>yes</td>
<td>1.87</td>
<td>13.9</td>
<td>yes</td>
<td>1.68</td>
<td>13.3</td>
<td>no</td>
<td>1.4</td>
<td>7.37</td>
<td>yes</td>
<td>2.2</td>
<td>31</td>
<td>yes</td>
</tr>
<tr>
<td>29</td>
<td>yes</td>
<td>yes</td>
<td>1.67</td>
<td>67.3</td>
<td>yes</td>
<td>1.41</td>
<td>36.4</td>
<td>no</td>
<td>1.2</td>
<td>22.5</td>
<td>yes</td>
<td>2.2</td>
<td>120</td>
<td>yes</td>
</tr>
<tr>
<td>30</td>
<td>no</td>
<td>yes</td>
<td>1.97</td>
<td>6.47</td>
<td>yes</td>
<td>1.01</td>
<td>10.6</td>
<td>no</td>
<td>1.1</td>
<td>14.8</td>
<td>no</td>
<td>1.2</td>
<td>9</td>
<td>no</td>
</tr>
<tr>
<td>31</td>
<td>no</td>
<td>yes</td>
<td>1.85</td>
<td>6.46</td>
<td>yes</td>
<td>1.6</td>
<td>15.3</td>
<td>yes</td>
<td>1.6</td>
<td>5.43</td>
<td>no</td>
<td>1.4</td>
<td>16</td>
<td>yes</td>
</tr>
<tr>
<td>32</td>
<td>no</td>
<td>no</td>
<td>1.05</td>
<td>25.3</td>
<td>no</td>
<td>1.03</td>
<td>15.1</td>
<td>no</td>
<td>0.9</td>
<td>1.82</td>
<td>no</td>
<td>0.9</td>
<td>29</td>
<td>no</td>
</tr>
<tr>
<td>33</td>
<td>no</td>
<td>no</td>
<td>1.29</td>
<td>5.46</td>
<td>no</td>
<td>1.07</td>
<td>1.44</td>
<td>no</td>
<td>1.1</td>
<td>7.53</td>
<td>no</td>
<td>1.3</td>
<td>18</td>
<td>no</td>
</tr>
<tr>
<td>34</td>
<td>yes</td>
<td>yes</td>
<td>1.37</td>
<td>24.3</td>
<td>no</td>
<td>1.19</td>
<td>7.38</td>
<td>no</td>
<td>1.2</td>
<td>8.25</td>
<td>yes</td>
<td>1.4</td>
<td>17</td>
<td>no</td>
</tr>
<tr>
<td>35</td>
<td>no</td>
<td>no</td>
<td>1.18</td>
<td>7.27</td>
<td>no</td>
<td>0.96</td>
<td>4.58</td>
<td>no</td>
<td>1.1</td>
<td>1.73</td>
<td>no</td>
<td>1.2</td>
<td>8.9</td>
<td>no</td>
</tr>
<tr>
<td>36</td>
<td>yes</td>
<td>yes</td>
<td>1.46</td>
<td>23.6</td>
<td>yes</td>
<td>1.42</td>
<td>13.8</td>
<td>no</td>
<td>1.2</td>
<td>3.73</td>
<td>yes</td>
<td>1.6</td>
<td>40</td>
<td>yes</td>
</tr>
<tr>
<td>37</td>
<td>yes</td>
<td>yes</td>
<td>1.33</td>
<td>15.2</td>
<td>no</td>
<td>1.23</td>
<td>5.37</td>
<td>no</td>
<td>1.2</td>
<td>8.44</td>
<td>yes</td>
<td>1.3</td>
<td>9.8</td>
<td>no</td>
</tr>
<tr>
<td>38</td>
<td>yes</td>
<td>yes</td>
<td>1.59</td>
<td>15.1</td>
<td>yes</td>
<td>1.56</td>
<td>9.9</td>
<td>no</td>
<td>1.3</td>
<td>7.55</td>
<td>yes</td>
<td>1.4</td>
<td>50</td>
<td>yes</td>
</tr>
<tr>
<td>39</td>
<td>no</td>
<td>no</td>
<td>1.03</td>
<td>9.3</td>
<td>no</td>
<td>1.01</td>
<td>7.78</td>
<td>no</td>
<td>1.1</td>
<td>13.8</td>
<td>no</td>
<td>1.1</td>
<td>130</td>
<td>no</td>
</tr>
<tr>
<td>40</td>
<td>no</td>
<td>yes</td>
<td>1.88</td>
<td>15.3</td>
<td>yes</td>
<td>1.62</td>
<td>9.09</td>
<td>yes</td>
<td>1.5</td>
<td>14.5</td>
<td>yes</td>
<td>1.4</td>
<td>31</td>
<td>yes</td>
</tr>
<tr>
<td>41</td>
<td>yes</td>
<td>yes</td>
<td>1.62</td>
<td>11.3</td>
<td>no</td>
<td>1.35</td>
<td>7.27</td>
<td>no</td>
<td>1.2</td>
<td>6.5</td>
<td>no</td>
<td>1.1</td>
<td>16</td>
<td>no</td>
</tr>
</tbody>
</table>

**Fig 7 - Method prediction results.**

<table>
<thead>
<tr>
<th></th>
<th>Manual</th>
<th>Global</th>
<th>Box</th>
<th>Mode</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No of images in test</strong></td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td><strong>Accuracy (%)</strong></td>
<td>87.80488</td>
<td>71.05263</td>
<td>65.78947</td>
<td>68.42105</td>
<td>81.57895</td>
</tr>
<tr>
<td><strong>Sensitivity (%)</strong></td>
<td>100</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>56.25</td>
</tr>
<tr>
<td><strong>Specificity (%)</strong></td>
<td>79.16667</td>
<td>86.36364</td>
<td>95.45455</td>
<td>63.63636</td>
<td>90.90909</td>
</tr>
</tbody>
</table>

**Fig 8 - Table of Test data results.**
Each method has been tested against the set of 41 test images. The results of each test can be seen in [Fig 6]. To make visualization of the data more human readable correct classifications against histology results are marked green and incorrect classification and marked red. Each Method has achieved a classification accuracy percentage <65%. Global performed best (71.1%) and Box scored lowest (65.8%). Mode scored 68.2%. Each method scored lower than the Manual selection method (87.8%). Each method performed well at classifying negative predictions. The specificity of results that define the percentage of actual negative results correctly identified by each algorithm Global, Box and Mode (86.4%, 95.5%, 63.6%) support this claim. Where these methods suffer is sensitivity. This is the percentage of correctly classified positive results. A significant drop in performance for each method can be seen in the results [Fig 8]. None of the methods reported scores >50%.

None of the methods implemented in this report have managed to equal or exceed the classification accuracy levels provided by the manual selection method. It is possible to increase the overall accuracy of the system by combining the results of each method Global, Box and Mode into a single method called ‘Combined’. By taking the most common prediction for each image from the three Methods defined in this report it is possible to increase the levels of my results. The new method Combined scored higher in all three tests. It scores an Accuracy of 81.5%, Sensitivity of 56.2% and Specificity of 90.9%. This Combined method manages to increase accuracy by >10% compared to the best single method and even performs better than the manual selection method at negative prediction. It still suffers from a low sensitivity.

Figure 9 – Visualization of Fig 7.
Chapter 5 – Conclusion

Findings
The results presented in this report have shown that it is possible to further automate aspects of Adenocarcinoma diagnosis in patients suffering with BE. The methods chosen have demonstrated that it is possible to use simple algorithms to reduce the systems dependency on a professional gastroenterologist. The quality of results presented will still provide a useful second opinion and the aim of this project was never to remove the user from making the final decision. The Global method according to the data was the best at classifying the data correctly but it still suffered from a very high sensitivity of 50%. It has also been shown that it is possible to achieve an improvement in all areas of evaluation by combining the three methods in this report to produce a 4th method called ‘combined’.

It is very easy to see from the findings of this report that it is possible to automate aspects of diagnosis normally left up to professionals. While that standard is not currently at the level offered by a human user it is not impossible to see that with further research it would be possible to deliver a system which either rivals or improves upon the performance provided by humans.

Limitations
There were a couple of aspects of this project which put limitations on to the overall ability of the background selection methods and what affect they had on the overall classification of regions of interest within the AFE examination images.

One of the major limitations experienced during this report was the extreme variation in the images provided. There was a lot of variation in the amount of BE in each image. This had a big impact of the performance of each method. For example in [Fig 10] it can be seen that there is a high level of BE present in the image with large parts of the image being displayed in magenta. This means that the selection of an appropriate background region is especially important in this case and it is known from the results that all three methods incorrectly classified the foreground region in this image. It is the nature of the methods being investigated in this report that they are very sensitive to the overall level of BE present in the image. This is caused by the classification criteria of

Figure 10 – AFE with high level of BE (image 2).
the decision tree [14], which as discussed earlier in this report relies on the contrast between the foreground and background regions. As a result the methods seem to perform best on images that have isolated instances of BE within the image.

Another visual aspect of the AFE images that have affected the performance of the background selection methods is the presence of the endoscope within the image. This can have a negative effect on prediction especially in the cases of Mode background selection. As each method has no functionality for the detection of the endoscope within the image in some cases it can be included in the background region selection by the method. This can be a problem because any data gathered from those areas in the image will dilute the values gained from actual tissue in the image.

**Possible Future Development**

The possibilities for further development of this project and the area of AFE are extensive. From the research presented in this report it is not hard to imagine possible improvements and extensions to further the aims of this project and the overall area of research. In this section we will discuss ideas that could be used to further develop the project.

Machine learning in this project has been limited to the use of the classification algorithm provided by [14]. The methods used in this report could be referred to as ‘dumb’ algorithms. This is because they are completely static in their implementation as no facility is made for the methods to improve their performance over time through learning. If a function could be trained on a set of background regions chosen by gastroenterologists, it could be possible to throw out regions that have been automatically chosen by the selection methods but deviate from the trained function’s values by too great a margin. This way it could be possible to rule out the use of certain methods, if it is known the result of a classification based on that method will be unreliable due to an unsuitable background selection.

Machine learning could also be used to improve the selection of foreground ROI. By taking the masks of images that have been pathologically tested if could be possible to train a function to recognize when a chosen region is similar to those that have been evaluated as containing dysplasia. This could be combined with a sliding window or region growing algorithm to investigate the entire image for areas which would be suitable for investigation by the classification algorithm. We could call this function ‘T1’.

Another improvement that could be added that would compliment background region selection is image segmentation. If the image could be segmented into non-overlapping regions then if T1 was used to validate those regions to find a region which is most similar to the ‘ideal’ background region,
it could be possible to select a region within the image which has the best chance of producing an accurate classification when used as part of the whole classification algorithm. This could be expanded to provide the user with a list of regions, all of which are marked on the screen, that the user can then select individually for use. This would allow for the user to use their judgment in region selection with support from the system. This level of support is the aim of CAD systems.

This report has demonstrated the ability to provide useful data for medical staff but all this has to be done using static images and as such has to be used either after the AFE examination or during the examination which could have the effect of lengthening procedures. If the system could be optimized for real-time video, it could be possible to overlay classifications on to the real-time image. In this situation foreground regions could be highlighted and tracked across multiple frames. Classification could also be performed in the background and results displayed in a faded font within the foreground region. A system that offers this functionality could help to reduce procedure times and improve the accuracy of requested biopsies.
Bibliography


[32] Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography


Appendix A

The project began with an ambitious furor which culminated in a project plan which was later revealed to be unachievable. Many of the problems that plagued the project were due to my own inexperience in project management and software development. On reflection the final project that has been delivered in this report met all of my personal goals for this project. At the start of the project I underestimated the time it would take me to setup and become familiar with the development environment that was need to continue working on the project files I had been provided with. I feel as the mistakes made during the project have allowed me to complete this report. Without making the initial underestimation of project task times in the first iteration of the project I would have been unable to correctly manage my time and that of the project to produce this final report. Through the lessons learnt during the project process there are several things that I would change if given the chance to do the project again. The first thing I would have changed would have been the location of my development environment. By choosing to do most of the software development from home, I managed to effectively isolate myself from potential help that could be provided by my supervisors and peers. This meant that problems took longer to solve and assistance with questions and issues were delayed.

Another aspect of the project which affected the overall results of this report was my own programming inexperience. This project was completed using the programming language C++ of which I have no prior experience using. Also the project was originally created and formatted as a visual studio solution in Microsoft Visual Studio. Which once again I had no experience using, this meant that development was extremely slow while I familiarized myself with the language and software. In hindsight I would have made room in the project to learn how to effectively use both Visual studio and C++. This would allow the effective allocation of tasks with the project plan and also mean that I would have the required skills to complete a task before starting it.

One of the most hindering aspects of the project process I experienced was the unavailability of the entire set of project files used by Vilius Narbutas to create ‘Tumor Detection’. My own inexperience caused me to waste time trying to reproduce the same environment Vilius had used. I should have made arrangements to collect all of the files missing files from the start of the project.

In summary I have learnt a lot of valuable lessons from this project. I have not only learnt these lessons but have put them into practice in this report and will continue to do so for the rest of my professional career. I have enjoyed putting the knowledge learnt while studying into a real world
application. I have also enjoyed working in a field of study which still has a lot of room to develop and hopefully will have a positive impact on the lives of patients with Barrett’s esophagus.
Appendix B

Listed are all items externally provided to the project.

- 82 x Anonymised AFE Images.
- 41 x text file containing marked ROI.
- Unmodified Tumor Detection program.
- Unmodified K-means clustering algorithm.
- 41 x Histological biopsy results.
Appendix C - How ethical issues are addressed

This project utilized data provided by the Leeds General Infirmary Gastroenterology Department. In accordance with established standards all images and data were provided free from any meta data information regarding the patients the images represent. All images were provided freely by patients and all effort have been made to ensure their identities have been protected. All data was provided for the project through Dr Venkataraman Subramanian who also provided the data also for use in a previous project undertaken at the University of Leeds school of Computing. Images were also prescreened to ensure that all personal information was removed from within the image.
Appendix D – Project Gantt chart