THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF ELECTRICAL ENGINEERING

AUTOMATED CLASSIFICATION OF HISTOPATHOLOGICAL IMAGES

JAMES J. GILLESPIE
Spring 2012

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Electrical Engineering
with honors in Electrical Engineering

Reviewed and approved* by the following:

Vishal Monga
Assistant Professor of Electrical Engineering
Thesis Supervisor

Julio V. Urbina
Assistant Professor of Electrical Engineering
Honors Advisor

*Signatures are on file in the Schreyer Honors College

Abstract

Automated image classification has become extremely desirable in the medical field for faster and more accurate diagnosis of disease. These systems have also gained significant attention because they help doctors in evaluating large volumes of medical imagery. An automated classification system has been developed that categorizes histopathological images provided by the Animal Diagnostic Laboratory (ADL) at the Pennsylvania State University. The classification system has been developed for four different mammalian organs: liver, lung, kidney, and spleen. The classifier automatically labels each histopathological image as healthy, inflammatory, or necrotic (liver only). This system uses supervised classification, which consists of two key stages. First, there is a feature extraction stage where discriminative image features are obtained. Next, these features are fed into a decision engine that performs the class assignment. The primary focus of the research proposed in this paper is the development of a robust feature extraction algorithm for histopathological image classification. Because of its widespread use as a state of the art classifier, Support Vector Machines are used as the decision engine in this research. After training the system with 20 images per class per organ, the system is tested using 30 images per class per organ, and produces classification accuracy up to 97%. These results prove the validity in continuing to improve the feature extraction and investigate different classifiers to provide an even more robust automated classification system for histopathological images.

Table of Contents

Chapter 1: Introduction

1.1 Image Classification: An Overview

Automated image classification is an important component of digital signal processing that has applications across various fields of study. The purpose of automated image classification is to provide a robust system that can accurately classify unknown images into two or more classes. For example, image classification has been used by the military to automatically classify targets into a military vehicle class or a civilian vehicle [1]. Automated image classification has also been used for security reasons such as face recognition [2] and fingerprint detection [3].

There are multiple advantages to building and implementing an image classification system. First, it provides an automated solution that would usually need human evaluation to make a decision on the class of an input image. Human evaluation of images can often be a costly option for classifying images, and sometimes it is not an option due to time constraints. Second, an automated image classification system can usually make a decision faster than human evaluation. This advantage makes image classification systems extremely valuable in time sensitive decision making such as automated target recognition used by the military. Finally, these systems can often be more accurate than human evaluation. In the medical field, automated image classification systems have been developed that are more accurate than

classification from a doctor [4]. This is due to the complex nature of medical imagery. The advantages of image classification systems make it an extremely valuable solution for a variety of applications.

While automated image classification does have its advantages, there are many challenges in developing a robust, accurate system. In order to develop such a system, a training, or example set of images is necessary for each class of images. For example, in automated target recognition, the system must have a set of images from the military vehicle class and the civilian vehicle class. Another challenge in developing an image classification system is developing an algorithm that is insensitive to distortions in the imagery such as noise or blurring. In real world scenarios, images are more often than not corrupted by some sort of distortion. Additionally, it is challenging to convert what is interpreted by the human eye into a mathematical representation. It is easy to recognize the image of one human face from another human face; however, it is challenging to develop a system that extracts these distinguishable features from a discrete valued two dimensional signal that is an image.

1.2 Classification of Medical Images

Automated image classification has become increasingly desirable in the medical field for a variety of reasons. An examination of these benefits will help to illustrate the relevance and significance of the research explained in this paper.

For a variety of diseases, early detection is one of the only ways to prevent the disease from becoming life threatening. In order to detect a disease early, many people undergo screenings very early in their life as a precautionary measure. Doctors are left with large quantities of medical imagery to examine and evaluate as a result of these precautionary screenings. Examination of medical imagery is a very time consuming process, and often doctors spend a majority of time evaluating benign tissue. In the case of prostate cancer, it is estimated that 80% of prostate biopsies are benign [5]; breast biopsies are also found

to be benign 70-80% of the time [6]. Pathologists therefore spend large amounts of time evaluating healthy or benign tissue. Automated image classification becomes desirable in order reduce the amount of time doctors spend evaluating healthy tissue to allow them to concentrate more on evaluating malignant tissue and treating patients. Not only can an automated image classification system save doctors and patients time, but it can often be more accurate than evaluation by a pathologist in detecting early signs of disease. Due to the low contrast nature of digital mammography, the early signs of breast cancer can be extremely difficult to detect, even for a trained specialist [6].

While automated image classification in the medical field has numerous advantages, there are some serious trade-offs. Due to real world circumstances, it is extremely difficult to develop a classification system that reaches 100% accuracy. In the medical field, inaccurate results can lead to inaccurate diagnosis of disease. In one instance, the error could be made by incorrectly diagnosing a healthy person as possibly diseased. This would lead to an unnecessary biopsy that would cost the doctor and patient time and money. The error also could be that a person with a disease was misdiagnosed as healthy because of a faulty image classification system [6]. The results of this misdiagnosis could have extremely dire consequences.

1.3 The Image Classification Process

The image classification process is a very methodical process and follows the diagram shown in Figure 1.

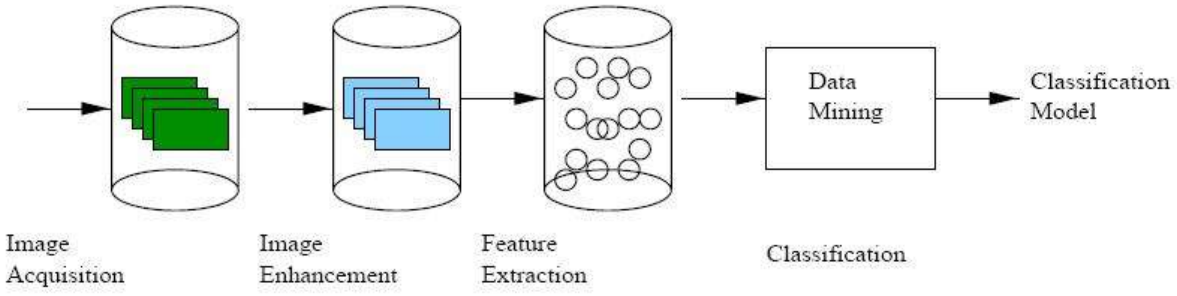


Figure 1: Image Classification Process (Courtesy [6])

The first step of the process is the image acquisition stage. This step involves collecting a diverse set of images for the classification system. The next step, image enhancement, involves preparing the image for the feature extraction stage. Image enhancement can involve cropping, segmenting, or normalizing each image to provide uniformity amongst an image set.

Feature extraction is the process of representing significant attributes of the image in a vector. This step is very crucial in the design of an accurate image classification system because the classifier makes its decision based on the feature vectors it is provided. Chapter 2 provides example images from each organ type, and explains how the important features are extracted.

The final step in the image classification process is the actual classification of the images. There are several different types of classifiers that have been developed and heavily researched. For this experiment, the classifier is chosen to be a Support Vector Machine [7], which has earned significant recognition as a powerful decision engine. Feature

vectors from training images are first provided to the Support Vector Machine; once it learns the trends from the training images, it can classify a test image. An important factor in the performance of a decision engine is having a large and diverse set of training images. Support Vector Machines will be discussed in detail in Chapter 3.

Chapter 2: Feature Extraction

The image classification system described in this paper has been designed to classify images provided by The Pennsylvania State University Animal Diagnostics Laboratory (ADL). The ADL database is made up of photomicrographs of liver, lung, kidney, and spleen cells. The lung, kidney, and spleen images are separated into two different categories: healthy and inflammatory. The liver images are separated into three categories: healthy, inflammatory, and necrotic.

2.1 Brief Discussion of Medical Conditions

Inflammatory and necrotic cell tissue in cattle is often a sign of a contagious disease. In order to prevent the spread of the disease amongst an entire farm worth of cattle, fast diagnosis is essential. The classification system developed in this research will facilitate this diagnostic process. There are a variety of causes of inflammation in the liver, lung, kidney, and spleen organs. When an infection is present, white blood cells are sent to the infection site in order to repair the damaged tissue. The Animal Diagnostic Laboratory studies the type of white blood cells present in order to diagnose the infection. The type of white blood cell can help indicate the duration and cause of inflammation. The four types of white blood cells that doctors at the laboratory discussed were eosinophils, neutrophils, macrophages, and lymphocytes. Inflammation can often be caused by an allergic reaction, bacteria, or parasites; the presence of eosinophils often indicates this type of infection. Tuberculosis is a common disease that causes inflammation due to higher bacteria. Acute infections are identified by the presence of neutrophils, while macrophages and lymphocytes indicate a chronic infection.

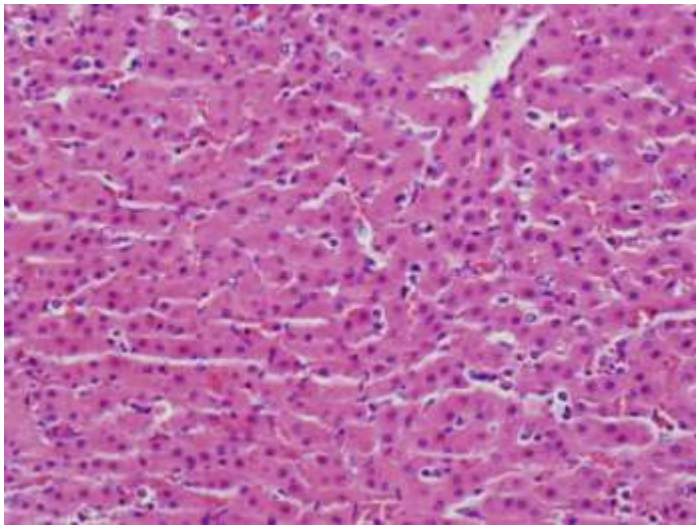
Necrosis is the scientific term for cell death, which also has a variety of causes. Lack of blood supply to an organ can often cause necrosis, as well as toxins from bacteria. The bacteria that cause necrosis can often be ingested, which is significant when a large amount of animals have access to the same food. While white blood cells help determine the cause of inflammation, it is the location of necrosis in the organ that helps to diagnose the cause of necrosis.

The automated image classification system developed in this thesis is by no means a complete solution for the diagnosis of inflammation or necrosis. It is, however, a significant step towards faster diagnosis of these two common but serious conditions that will benefit researchers at the Animal Diagnostic Laboratory at Penn State, and elsewhere.

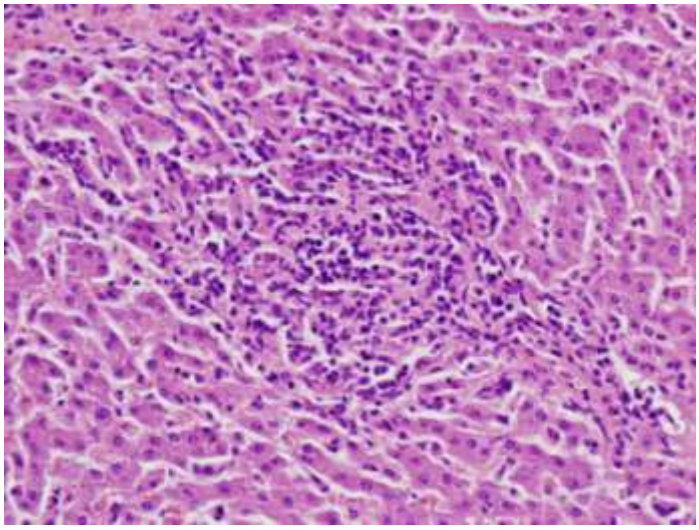
2.2 Evaluation of Organ Features

Before designing a feature extraction algorithm, it was important to evaluate the characteristics that distinguish healthy tissue from inflammatory or necrotic tissue. Each organ had slightly different characteristics; therefore, different features had to be extracted for each organ.

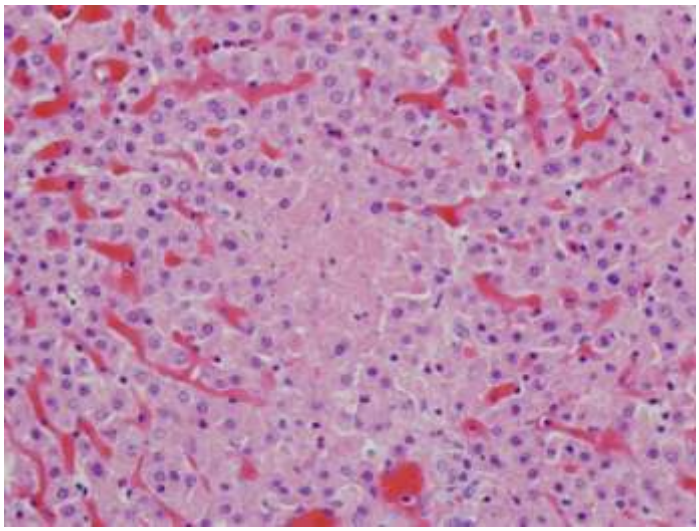
2.2.1 Liver Photomicrographs



(a)



(b)

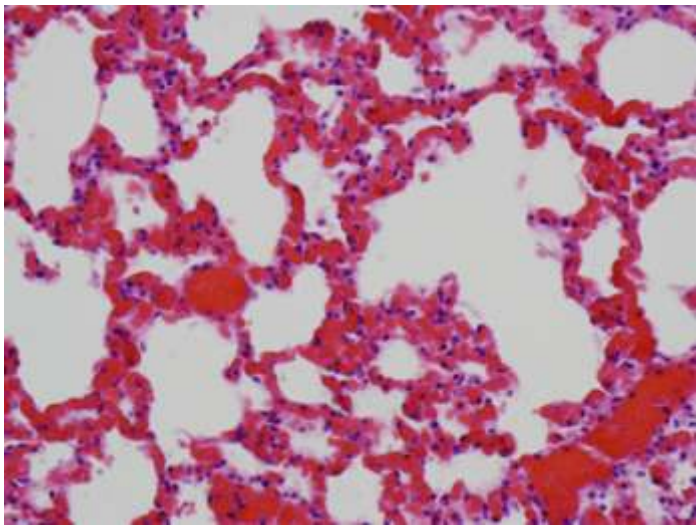


(c)

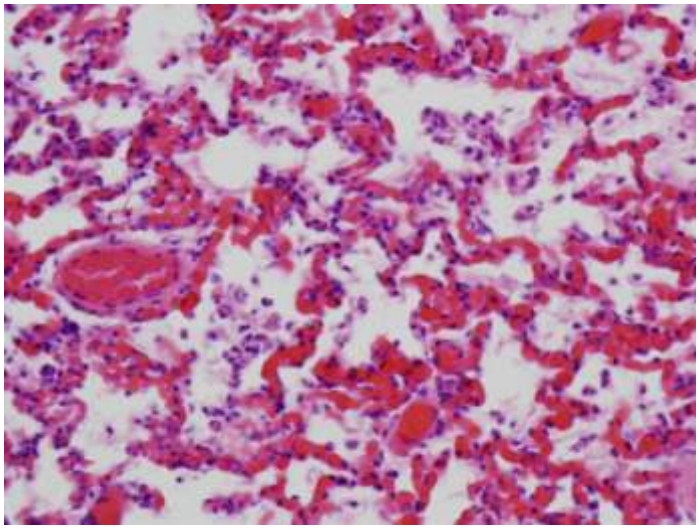
Figure 2: Photomicrographs of (a) healthy, (b) inflamed, and (c) necrotic liver tissue.

It is fairly easy to see in Figure 2 the differences between the healthy, inflamed, and necrotic liver tissue. Healthy liver tissue has uniform nuclei, the bluish dots, across the entire image. The inflamed liver tissue has a large cluster of dark blue nuclei located in the center of the image. The distinguishing feature of necrotic liver tissue is the pale region in the center that lacks any nuclei or cell separation.

2.2.2 Lung Photomicrographs



(a)



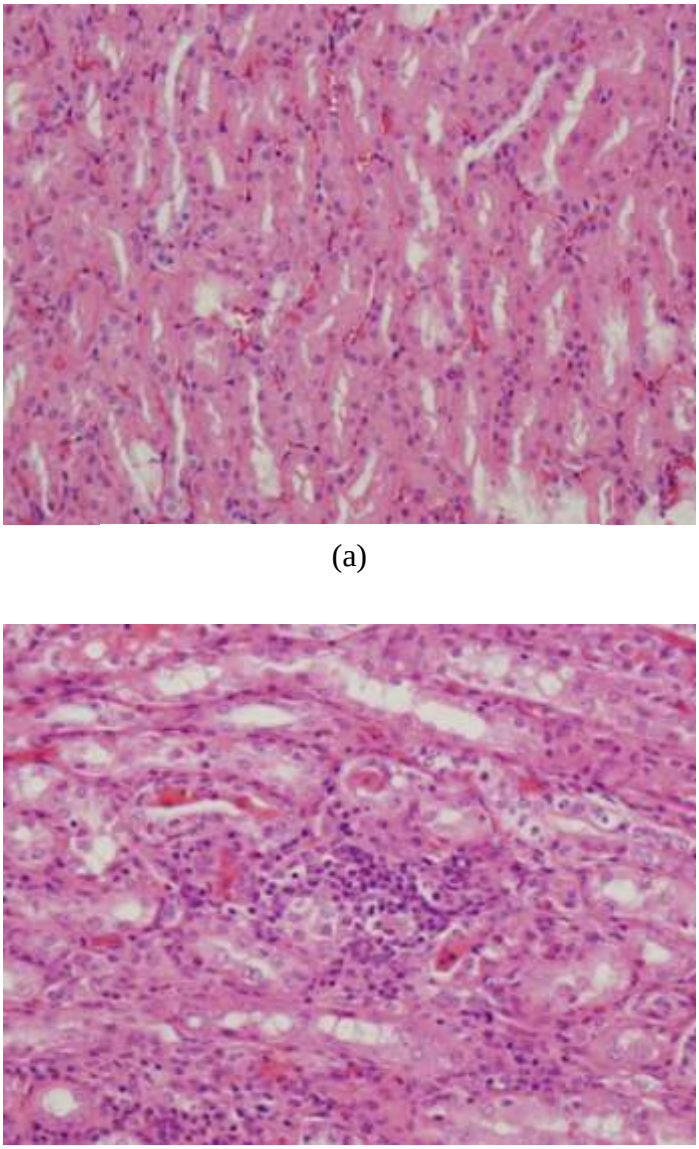
(b)

Figure 3: Photomicrographs of (a) healthy and (b) inflamed lung tissue.

Figure 3 depicts the distinguishing features between healthy lung cells and inflamed lung cells. The healthy lung image is characterized by large clear openings of the alveoli, or lung cells. In the inflamed lung cell photomicrograph, the alveoli do not have clear openings; they are filled with bluish-purple inflammatory cells. A large part of the feature

extraction for the lung focused on identifying the large, clear openings versus openings that were filled with inflamed cells.

2.2.3 Kidney Photomicrographs



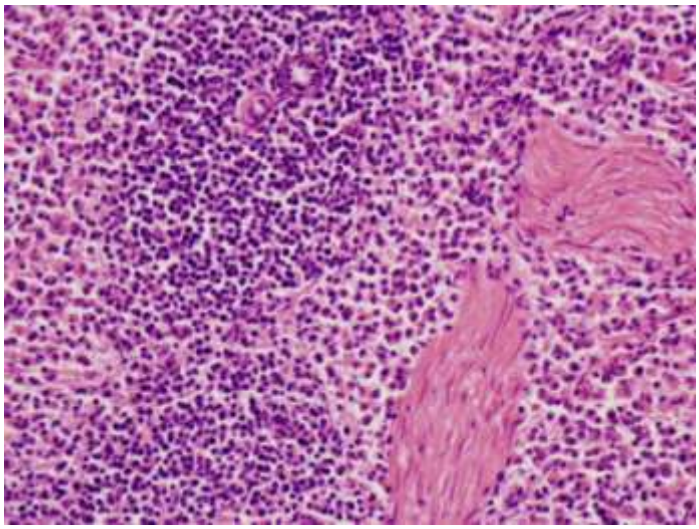
(a)

(b)

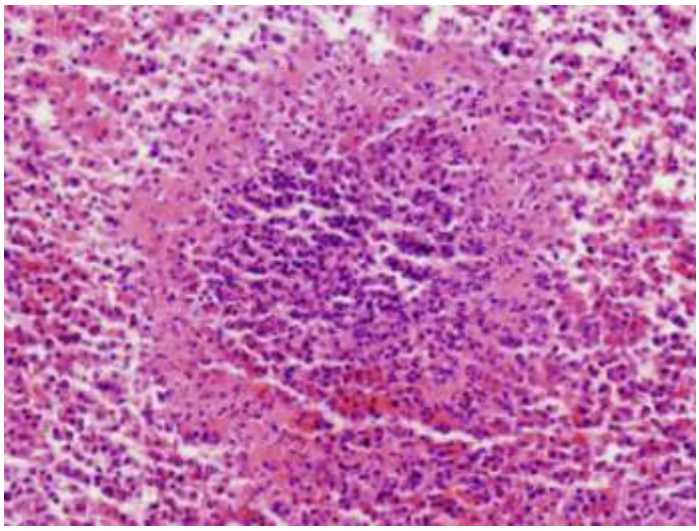
Figure 4: Photomicrographs of (a) healthy and (b) inflamed kidney tissue.

Similar to the liver cell, healthy kidney cells display uniform nuclei over the entire region of the photomicrograph. In Figure 4 (b), inflamed cells are located in the center of the image and are characterized by the clustering of dark blue nuclei.

2.2.4 Spleen Photomicrographs



(a)



(b)

Figure 5: Photomicrographs of (a) healthy and (b) inflamed spleen tissue.

Figure 5 (a) is a photomicrograph of healthy spleen cells. Most of the image displays uniform nuclei similar to the kidney and liver cells. However, as shown in Figure 5 (a), there are certain areas of the spleen, characterized by large, paler portions of pink, that

are not indicative of disease. This characteristic of spleen cells makes classification more difficult. Figure 5 (b) displays the characteristics of inflamed spleen cells, which are clusters of dark blue nuclei surrounded by an elliptical area of pale pink. The classification of photomicrographs of the spleen was the most challenging because of the non-uniform characteristics of both healthy and inflamed cells.

2.3 Feature Extraction

As shown in Section 2.2, the characteristics that separate healthy cells from inflammatory cells are similar amongst all of the organs. For this reason, a unique set of features was not necessary for each organ. This section provides detail of the various features that were extracted and the reasoning behind each feature.

2.3.1 Pixel Histograms

The first, most rudimentary feature extracted from all images was a quantized pixel histogram. An image histogram provides information regarding the frequency of pixel values within an image. This feature was selected because color plays an important role in distinguishing healthy cells from diseased cells. Healthy images will have a histogram containing pixel values closer to the white side of the spectrum, while inflammatory images will have a histogram containing pixel values closer to the black side of the spectrum, due to the significantly darker nature of inflamed cells. A histogram quantized to 16 bins was created for each of the red, green, and blue color channels. Figure 6 illustrates the traditional histogram compared to the quantized histogram. The histogram was quantized to reduce the computational complexity of the algorithm as well as to reduce the dimensionality of the feature vector.

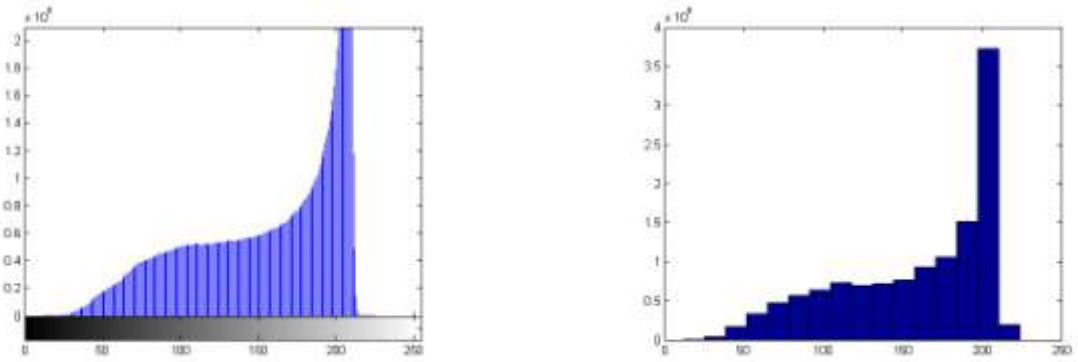


Figure 6: (a) Non-quantized and (b) 16 bin quantized histograms.

2.3.2 Statistical Data

In addition to pixel histograms, a statistical representation of each image was used as another feature. Statistical features have been proven to give good classification results, as noted in [8]. The statistical features extracted from each image were mean, median, maximum, minimum, and standard deviation. The use of statistical features has been proven in a variety of papers as a good feature for image classification; however, from an intuitive level, it is easy to comprehend the importance of these features relating to this specific problem. Pixel values will be lower in an inflamed image due to the darkness of the inflamed cells. This will force the mean of the pixel values lower than that of a normal liver cell image, and will also affect the other statistical measures used.

2.3.3 Gray Level Co-Occurrence Matrix

One of the most useful features utilized in this classification system is the gray level co-occurrence matrix (GLCM). Properties of the gray level co-occurrence matrix have been proven to provide valuable information regarding the texture of an image [9]. The MATLAB command *graycomatrix* was used to generate the gray level co-occurrence

matrix of each image, while *graycoprops* was used to extract the properties of the GLCM The following properties were used as features in this classification scheme [10]:

- Contrast: A measure of the intensity contrast between a pixel and its neighbor over the whole image.
- Correlation: A measure of how correlated a pixel is to its neighbor over the whole image.
- Energy: The sum of squared elements in the GLCM
- Homogeneity: A value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal

2.3.4 Wavelet Decomposition

The final feature extracted from each image was a nine-level two dimensional discrete wavelet approximation. Wavelets have gained significant attention as a useful tool for a variety of image processing applications [11] [12]. The benefit to using wavelet decomposition is the fact that it provides information not normally available from the spatial domain information [13]. Wavelets have been used in applications such as JPEG200 image compression [14] as well as features in image classification systems [15]. Four images are generated as a result of the wavelet decomposition; however, this system utilizes only the LL component. The LL sub-image represents the low spatial frequency components in both the horizontal and vertical directions. A nine-level decomposition was necessary because of the size of the images provided. Each image was 3072 by 4080 pixels. The nine-level decomposition provided an 8 by 6 approximation of the image, which was then used as a feature for the classifier.

2.4 Lung Specific Features

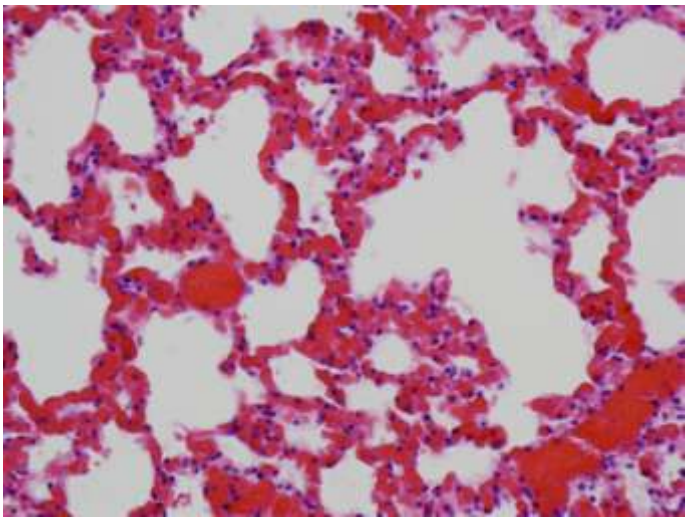
The liver, kidney, and spleen exhibit similar distinguishing characteristics between healthy and diseased tissue. The lung, however, is distinctively different from the rest of the organs. For this reason, two additional features were extracted in order to capture the distinguishing characteristics of the lung images.

2.4.1 Percentage White

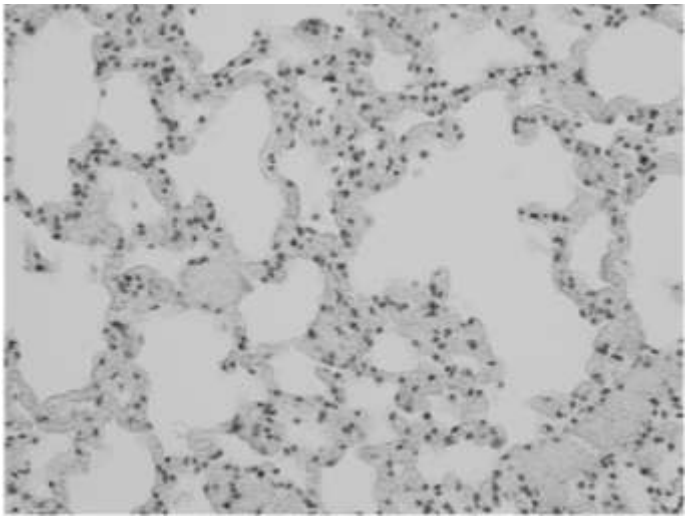
One of the distinguishing features between healthy lung tissue and inflamed lung tissue is the amount of white present in the image. As shown in Section 2.2, healthy lung tissue has large white openings throughout the image, whereas inflamed lung tissue has inflamed cells where the large openings should be. The percentage of white pixels was calculated in the following way:

1. Convert color image to red, green, and blue grayscale images.
2. Convert grayscale images to a binary image using a threshold of 180.
3. Sum up the total number of white pixels and divide by the total number of pixels.

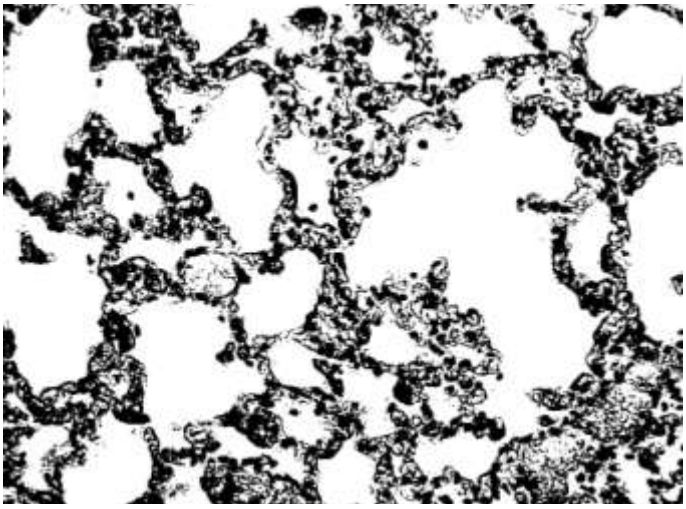
Figure 7 displays an example of the image processing done in order to make this calculation. For a majority of the training set of images, it was observed that the percentage white statistic was reliable in distinguishing healthy from diseased tissue.



(a)



(b)



(c)

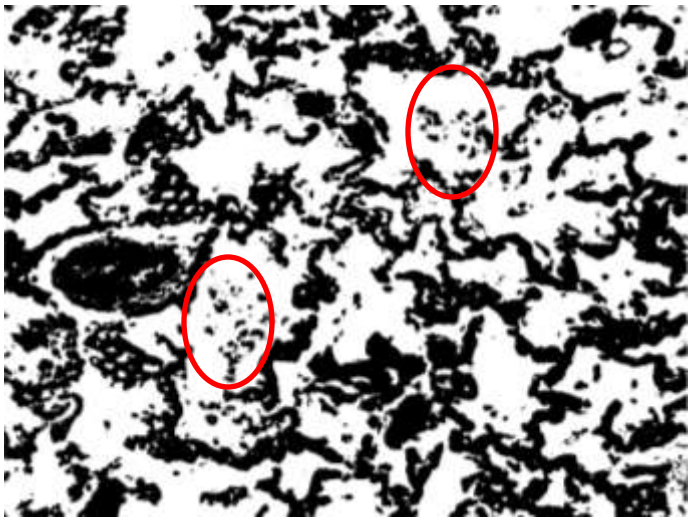
Figure 7: (a) Color image, (b), red channel grayscale image, and (c) binary image of lung cell photomicrograph

2.3.2 Percentage Filled

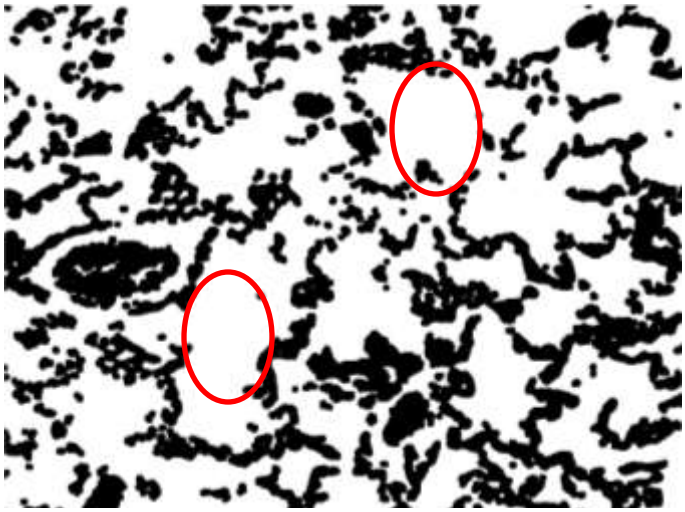
Another feature used for the lung images is the percentage of white that is filled. In the healthy lung tissue, it is easy to identify the open white spaces. The white spaces that *should* be open are still identifiable by eye in the inflamed lung tissue. To the human eye, the main distinguishing feature of inflammatory lung tissue is that the white space is filled with inflamed cells. The percentage filled feature is an attempt to calculate the white openings of either healthy or inflamed tissue, and then figure out if the tissue if filled with inflamed cells. In order to calculate the percentage filled, a classical image processing technique known as morphological closing was used [16]. The following steps were used to calculate the percentage filled:

1. Convert color image to red, green, and blue components.
2. Convert grayscale images (pixel intensity range 0 to 255) to binary images using a threshold of 160.
3. Morphologically close the image.
4. Subtract the closed image from the binary image.
5. Sum up white pixels from image in step 4 and divide by sum of white pixels from image in step 3.

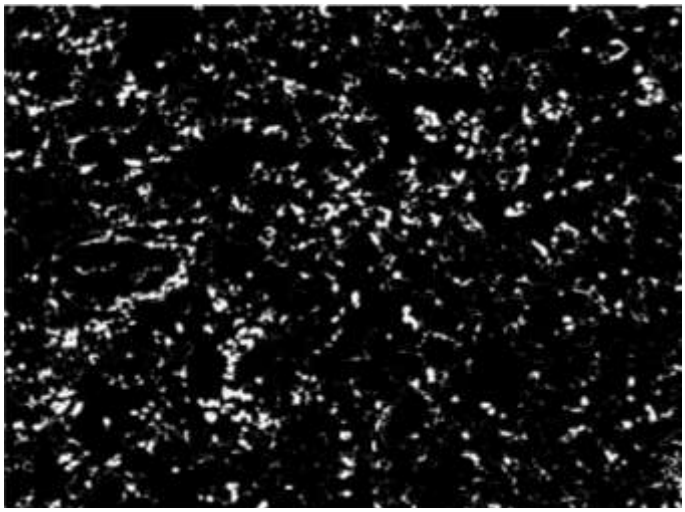
Figure 8 displays the steps taken in image processing in order to calculate the percentage filled feature for an inflammatory lung photomicrograph. The red circles in Figure 8 (a) call attention to inflammatory cells in the lung tissue. In Figure 8 (b), the red circles illustrate that the morphologically closed image no longer has these inflammatory cells. Thus, the white region that should be clear of inflammatory cells is identified, and the percentage filled is then calculated as stated above.



(a)



(b)



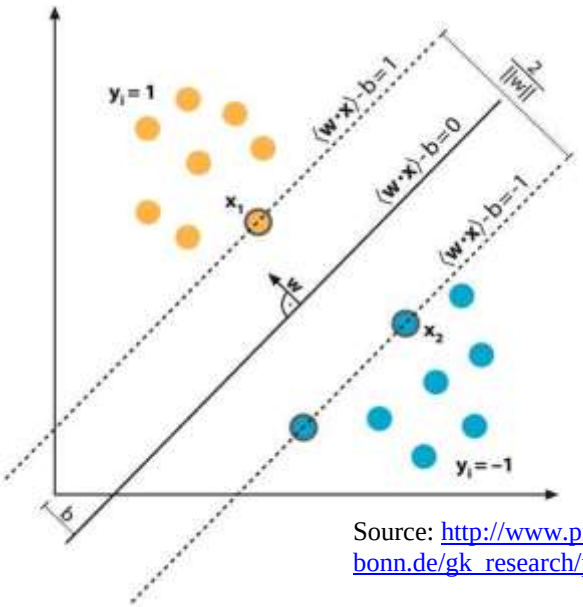
(c)

Figure 8: (a) Binary image, (b) closed binary image, and (c) closed minus binary image

Chapter 3: Decision Engine: SVM

The classification algorithm used in this system was a Support Vector Machine. While there are a variety of classification algorithms, the Support Vector Machine has been proven to be a powerful classifier [7]. The focus of this research was to investigate and develop a strong feature extraction algorithm without experimenting with a variety of classifiers.

After extracting useful features from every training image, the feature vector is fed into the input of the Support Vector Machine. A label vector is also input into the Support Vector Machine, labeling each set of training features as either healthy or diseased. From there, the data is mapped non-linearly to a dimension higher than that of the feature space [17]. Once the data is mapped to a higher dimension, the Support Vector Machine generates a hyperplane that optimally separates the healthy and diseased classes. Figure 9 illustrates in a simplistic manner the Support Vector Machine algorithm:



Source: http://www.precision-crop-protection.uni-bonn.de/gk_research/project.php?project=3_06

Figure 9: Plot of feature vectors and hyperplane

Each red or blue dot represents one feature vector from one image that is mapped to a higher dimension. The Support Vector Machine establishes the hyperplane, the line labeled (w*x)- b = 0 in Figure 9.

After the training phase, the Support Vector Machine is ready to classify new images. The input to the Support Vector Machine is now the feature vector of the test images; a label vector is not an input at this stage, as this would defeat the purpose of testing. The algorithm then continues to map the feature vectors to a higher dimension using the same mapping function as for the training images. Once the feature vectors are mapped, the Support Vector Machine makes a decision on the class of the image based on its spatial relation to the hyperplane developed in the training stage.

In Chapter 2, dimensionality was mentioned as a reason for reducing the number of samples for the quantized histogram. The "curse of dimensionality" is a well-known aspect of any sort of statistical classification problem. The "curse of dimensionality" states that the demand for a large number of samples grows exponentially with the dimensionality of the feature space [17]. Therefore, as more and more features are added, classification rates can decrease unless the number of training samples is increased. Because this classification system is designed for a small number of training and test images, the feature space was kept relatively small in order to avoid the "curse of dimensionality."

Chapter 4: Experimental Results

The automated image classification system developed in this paper was subject to two different rounds of testing.

4.1 Initial Testing and Results

For the first round of testing, the Pennsylvania State University Animal Diagnostic Laboratory provided 25 images per class per organ. The classification system was developed and tested multiple times with this initial set of images until favorable results were achieved. A subset of 10 images per class per organ were used as training images for the classification system, and the remaining 15 images per class per organ were used as test images. The 10 training images per organ per class were chosen randomly during each trial in order to test the robustness of the system. The desired classification accuracy, around 85% for initial testing, was not achieved until all of the features discussed in Chapter 2 were added to the feature extraction algorithm. Tables 1-4 show the experimental results using four different random training sets. Each row in these confusion matrices represents the expected classification category, while each column represents the category predicted by the classification system. Therefore, a good classifier will have numbers close to one across the diagonal; these numbers are bolded in the following tables.

Table 1: Initial Classification Results For Liver Cell Images Using Four Random Training Sets

	Healthy	**Inflamed**	**Necrotic**
Healthy	**1.00**	0.00	0.00
Inflamed	0.00	**1.00**	0.00
Necrotic	0.00	0.07	**0.93**

	Healthy	**Inflamed**	**Necrotic**
Healthy	**0.93**	0.00	0.07
Inflamed	0.00	**1.00**	0.00
Necrotic	0.00	0.07	**0.93**

	Healthy	**Inflamed**	**Necrotic**
Healthy	**0.93**	0.00	0.07
Inflamed	0.00	**0.92**	0.08
Necrotic	0.00	0.07	**0.93**

	Healthy	**Inflamed**	**Necrotic**
Healthy	**0.86**	0.14	0.00
Inflamed	0.00	**1.00**	0.00
Necrotic	0.07	0.13	**0.80**

Table 2: Initial Classification Results For Lung Cell Images Using Four Random Training Sets

	Healthy	**Inflamed**
Healthy	**1.00**	0.00
Inflamed	0.07	**0.93**

	Healthy	**Inflamed**
Healthy	**0.86**	0.14
Inflamed	0.00	**1.00**

	Healthy	**Inflamed**
Healthy	**0.86**	0.14
Inflamed	0.07	**0.93**

	Healthy	**Inflamed**
Healthy	**0.86**	0.14
Inflamed	0.07	**0.93**

Table 3: Initial Classification Results For Kidney Cell Images Using Four Random Training Sets

	Healthy	**Inflamed**
Healthy	**1.00**	0.00
Inflamed	0.00	**1.00**

	Healthy	**Inflamed**
Healthy	**1.00**	0.00
Inflamed	0.07	**0.93**

	Healthy	**Inflamed**
Healthy	**0.80**	0.20
Inflamed	0.00	**1.00**

	Healthy	**Inflamed**
Healthy	**1.00**	0.00
Inflamed	0.07	**0.93**

Table 4: Initial Classification Results For Spleen Cell Images Using Four Random Training Sets

	Healthy	**Inflamed**
Healthy	**0.87**	0.13
Inflamed	0.07	**0.93**

	Healthy	**Inflamed**
Healthy	**0.87**	0.13
Inflamed	0.13	**0.87**

	Healthy	**Inflamed**
Healthy	**0.87**	0.13
Inflamed	0.00	**1.00**

	Healthy	**Inflamed**
Healthy	**0.73**	0.27
Inflamed	0.07	**0.93**

Tables 1-4 illustrate numerous characteristics about the proposed classifier. Overall, the classifier reaches accuracies in the 80-90% range for each random training set, indicating that the classifier is robust to changes in the training set. Also, higher classification accuracies are seen for the inflammatory and necrotic cases. In classification of medical imagery, it is highly desirable to have a system that is very accurate in classifying malignant tissue.

4.2 Final Testing and Results

The results from Tables 1-4 indicated that the classification system was performing successfully. Because the system was proven to have high classification accuracy for the initial set of images, an entirely new set of images was obtained from the ADL in order to test the system on a larger set of images. During the final trials, the entire image set consisted of the original images used in the initial testing and a new set of images provided by the ADL. In total, there were 50 images per class per organ. Twenty training images per class were chosen randomly in order to ensure that results were not based off of an extremely good or extremely bad training set. For each organ, the classification system was tested with twenty different training sets. The average classification accuracies and standard deviations for each organ are shown in Table 5. Figures 10-13 illustrate classification accuracies between using a training set of 20 images compared to a training set of 25 images.

Table 5: Final confusion matrices of (a) liver, (b) lung, (c) kidney, and (d) spleen tissue.

	Healthy	**Inflammatory**	**Necrotic**
Healthy	**0.94 ± 0.05**		
Inflammatory		**0.92 ± 0.06**	
Necrotic			**0.86 ± 0.20**

(a)

	Healthy	**Inflammatory**
Healthy	**0.95 ± 0.04**	
Inflammatory		**0.97 ± 0.02**

(b)

	Healthy	**Inflammatory**
Healthy	**0.74 ± 0.07**	
Inflammatory		**0.84 ± 0.20**

(c)

	Healthy	**Inflammatory**
Healthy	**0.89 ± 0.05**	
Inflammatory		**0.89 ± 0.06**

(d)

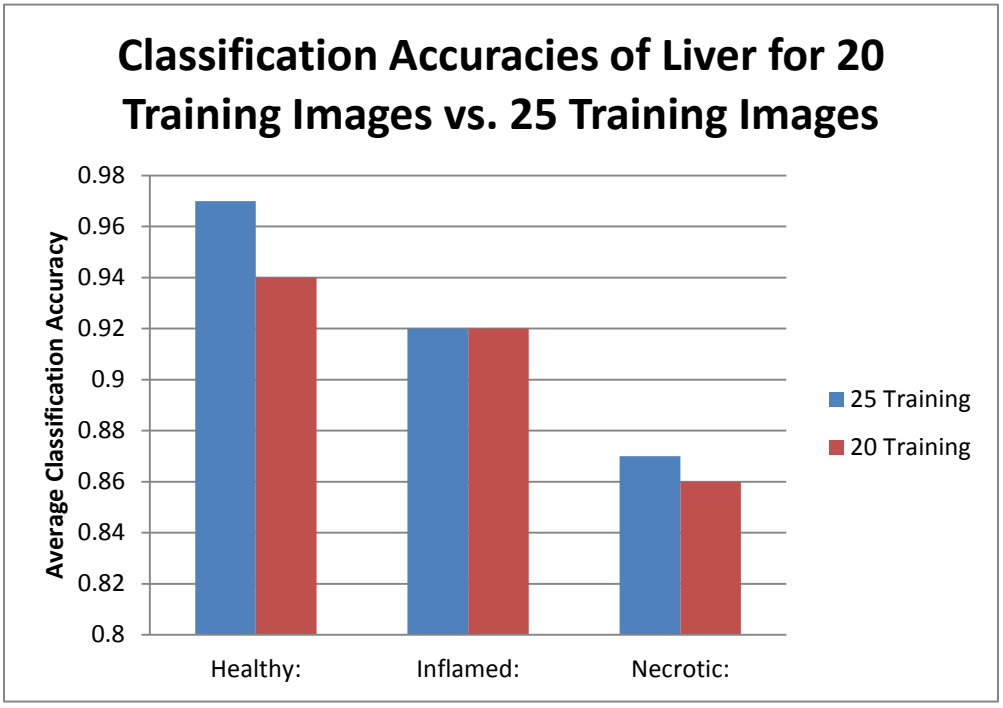


Figure 10: Classification accuracy of liver cell images

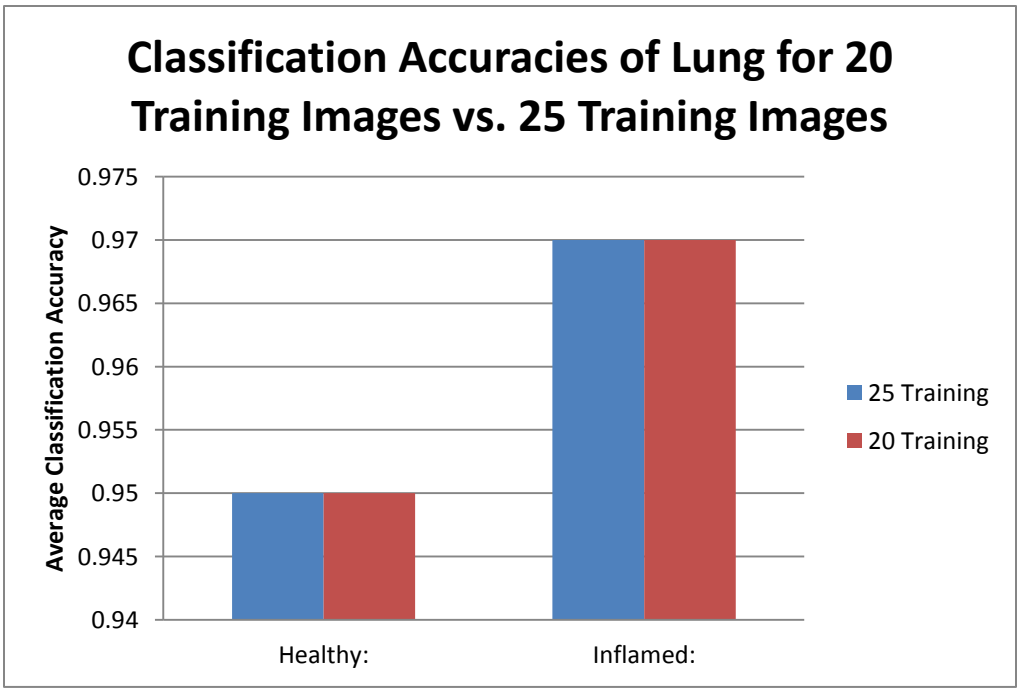


Figure 11: Classification accuracy of lung cell images

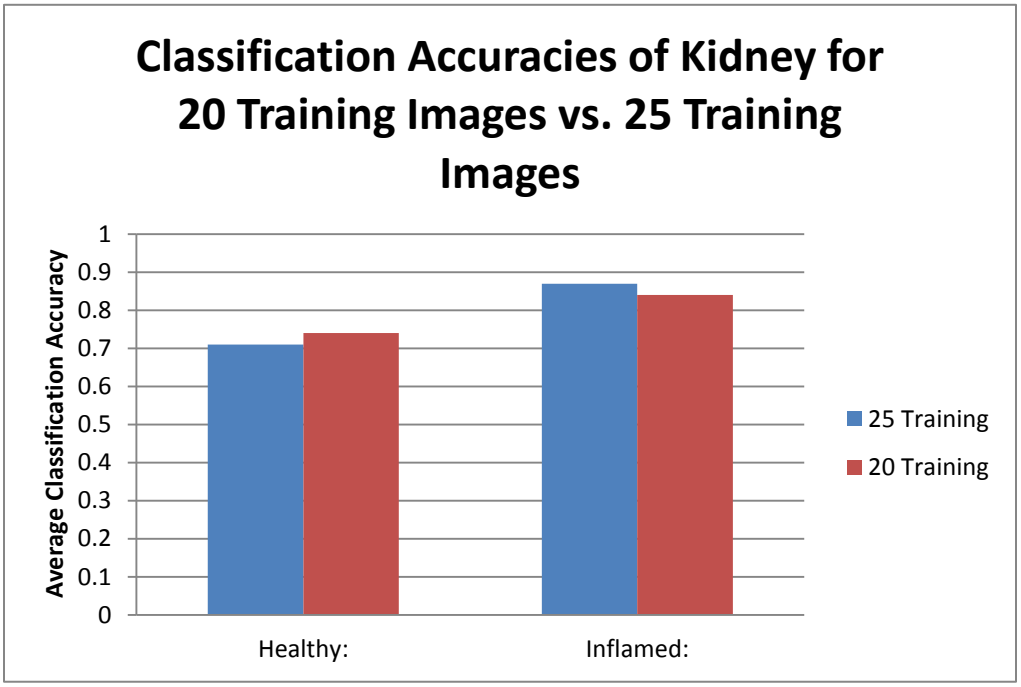


Figure 12: Classification accuracy of kidney cell images

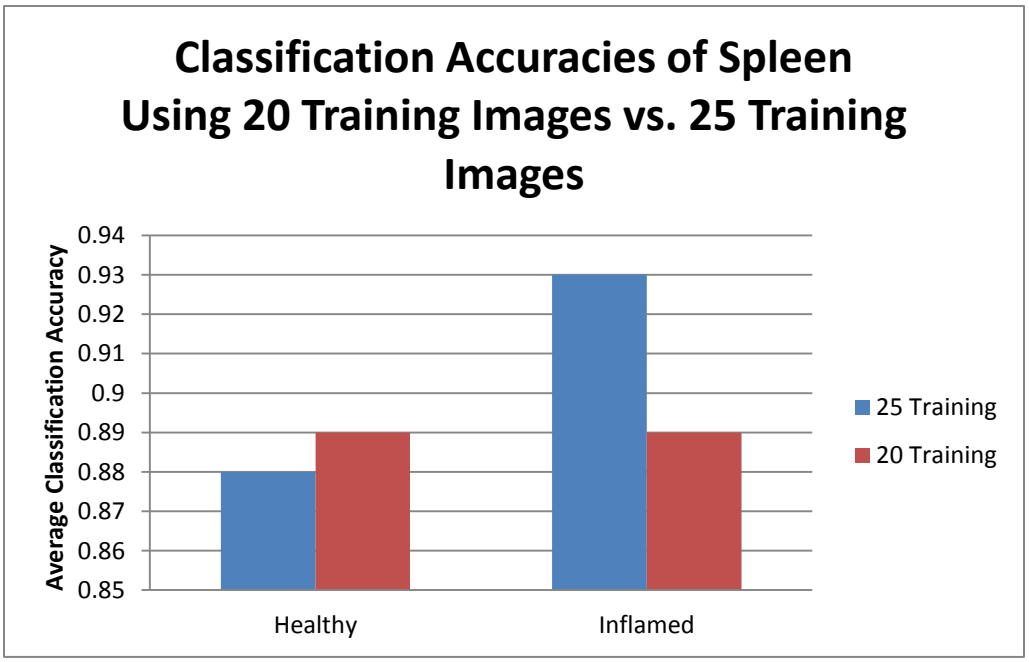


Figure 13: Classification accuracy of spleen cell images

The experimental results in Table 5 are a good representation of the effectiveness of the proposed classification system because of the number of random training trials. The standard deviation was included in this set of results to show how far the accuracies varied from the mean value over the twenty trials. Other than the kidney, the classification accuracies were all above 85%, which is comparable to other classifiers in the medical field [6], [10]. Comparison of the initial set of kidney images and the final set of kidney images showed that there were significant differences in the staining of the tissue, which could contribute to the lower classification rate. Table 5 (a) also illustrates a large deviation in the identification rates of necrotic liver tissue. Again, this was due to large differences in the necrotic tissue from the initial and final sets of images from the ADL.

Another measure of the effectiveness of a classification system is how well it performs for smaller training sets. Figures 10-13 illustrate that the classification accuracies do not drastically decrease when using a training set of 20 images compared to a training set of 25 images. In some cases, the classification accuracy is slightly higher for the training set of 20 images. This discrepancy is due to the fact that the training sets were randomly selected. Nonetheless, the insensitivity to size in the training set illustrates the robustness of this classification system.

Chapter 5: Conclusion

5.1 Interpreting the Results

The automated image classification system proposed in this paper utilizes a unique set of features and a Support Vector Machine decision engine to classify the image database provided by the Pennsylvania State University Animal Diagnostic Laboratory. Utilizing an initial set of 25 images per class per organ, the classification algorithm achieved outstanding results. However, as the final set of images was added to the collection, the classification rates were slightly lower for the kidney and liver. There are likely two causes for the lower accuracies. First, the second set of images was taken from a different breed of cattle than the first set of images. Many of the tissues had distinctly different characteristics. Second, the final set of images was stained differently than the first set of images. Researchers at the ADL reason that the staining of histopathological images varies from day to day. A potential solution to this problem is discussed below.

5.2 Future Work

The classification system proposed in this paper is an excellent preliminary classifier for an initial image database provided by the Pennsylvania State University Animal Diagnostic Laboratory. The experimental results exemplify that automated classification of histopathological images is a worthwhile field of interest. The continuing collaboration between the ADL and the Department of Electrical Engineering will facilitate an even more complex and robust classification system.

From the image processing side of the project, there are many possibilities for future improvements to this system. First, because of the distinct staining differences between

the initial set of images and the final set of images, a color normalization pre-processing stage would be beneficial [18]. This step will prevent classification errors due to differences amongst staining in the histopathological images. Next, the exploration and utilization of different features is an extremely important in designing a more accurate classification system. While classification of histopathological images is a relatively new venture, there are a wide variety of features that are applicable in this situation. Nuclei segmentation statistics [10], adaptive discriminant wavelet packet transforms [19], and color texture analysis using self-organizing maps [10] are all possible features that have provided strong classification rates. The investigation of completely new features is also important in developing a unique classification system for this application.

Not only is it important to investigate new features for the further development of this system, but it is critical to explore decision engines other than Support Vector Machines. Meta-classification using Support Vector Machines can provide more accurate classification rates [13]. The use of neural networks and boosting algorithms may also provide better results.

From the ADL side of this project, more images will be required to further test the robustness of the system. Many automated image classification systems utilize databases with hundreds images available for training and testing [20]. A large set of images provided from the ADL will allow further development of a more accurate and robust system.

5.3 Concluding Remarks

Automated classification has become increasingly important in the medical field for diagnosis of cancers, especially breast and prostate cancer. While classification of

histopathological images has been approached, the classification of liver, lung, kidney, and spleen tissue from cattle is a unique venture. The automated classification system proposed in this paper is a functioning prototype of a system that will help researchers at the Animal Diagnostic Laboratory at the Pennsylvania State University diagnose inflammatory and necrotic tissue in a faster, more accurate manner.

References

[1] Q. Zhao and J. Principe, "Support vector machines for SAR automatic target recognition," *IEEE Trans. Aerospace and Electronic Systems*, vol. 37, pp. 643–654, April 2001.

[2] G. Burel and D. Carel, "Detection and localization of faces on digital images," *Pattern Recognition Letters,* 15: 963-967, 1994.

[3] P. Baldi and Y. Chauvin, "Neural Networks for Fingerprint Recognition," *Neural Computation*, 5(3):402-418, 1993.

[4] D. Brazokovic and M. Neskovic. "Mammogram screening using multiresolution-based image segmentation," *International Journal of Pattern Recognition and Artificial Intelligence,* 7(6): 1437-1460, 1993.

[5] Gurcan et. al. "Histopathological Image Analysis: A Review," *IEEE Reviews in Biomedical Engineering,* vol. 2, pp. 147-171, 2009.

[6] Antonie et. al. "Application of Data Mining Techniques for Medical Image Classification," *Proc. Of the Second International Workshop on Multimedia Data Mining,* August 2001.

[7] V. N. Vapnik, *The nature of statistical learning theory*. New York, USA: Springer, 1995.

[8] R. Argrawal et. al. "Mining association rules between sets of items in large databases," *Proc. 1993 ACM –SIGMOD Int. Conf. Management of Data,* pp. 207-216, May 1993.

[9] L. E. Boucheron, "Object- and Spatial Level Quantitative Analysis of Multispectral Histopathology Images for Detection and Characterization of Cancer," PhD. Dissertation, Univ. of California Santa Barbara, Santa Barbara, CA, 2008.

[10] O. Sertel et. al. "Histopathological Image Analysis Using Model Based Intermediate Representations and Color Texture: Follicular Lymphoma Grading," *J. Signal Processing Syst.,* to be published.

[11] S. Mallat, *A Wavelet Tour of Signal Processing.* Academic Press, 1999.

[12] G. Strang and T. Nguyen, *Wavelets and Filter Banks.* Wellesley College, 1996.

[13] U. Srinivas, V. Monga, R. Raj. "Meta-classifiers for exploiting feature dependencies in automatic target recognition," *Proc. IEEE Radar Conference,* pp. 147-151, 2011.

[14] DS Taubman et. al. "JPEG2000: Image compression fundamentals, standards, and practice," *Journal of Electronic Imaging,* 2002.

[15] T. Wang and N. Karayiannis, "Detection of microcalcification in digital mammograms using wavelets," *IEEE Trans. Medical Imaging,* 17(4):498-509, 1998.

[16] I. Pitas, "Morphological Shape Decomposition," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, pp. 38-25, Jan. 1990.

[17] O. Duda, P. Hart, D. Stork, *Pattern Classification.* Wiley-Interscience, 2000.

[18] L. Yang, P. Meer, and D. Foran, "Unsupervised segmentation based on robust estimation and color active contour models," *IEEE Trans. Inf. Technol. Biomed. ,* vol. 9, no. 3, pp. 475-486, Sep. 2005.

[19] H. Qureshi et. al. "Adaptive discriminant wavelet packet transformation and local binary patterns for meningioma subtype classification," *Medical Image Computing and Computer Assisted Intervention – MICCAI 2008,* pp. 196-204.

[20] L. Shamir et. al. “IICBU 2008 – A proposed benchmark suite for biological image analysis,” *Medical and Biological Engineering and Computing,* Vol. 46, No. 9, pp. 943-947, Sep. 2007.

James Gillespie

Campus Address
138 S. Atherton Street, Apt. 616
State College, PA 16801

Jjg318@psu.edu
215-264-2878

Permanent Address
4735 Redwood Court
Doylestown, PA 18902

OBJECTIVE Obtain a full time position in electrical engineering after graduation

EDUCATION

Bachelor of Science in Electrical Engineering (May 2012)
The Pennsylvania State University, University Park, PA
Schreyer's Honors College

Relevant Coursework:

Electronic Circuit Design	Digital Design	Control Systems
Communication Systems	Signal Processing	Linear Systems

WORK EXPERIENCE

Manufacturing Engineer
Kraft Foods, Avon, NY — Summer 2011

- Developed a graphical interface for industrial process to improve monitoring of the process
- Wrote, tested, and implemented code for a control panel
- Collaborated with different departments within the plant for several projects

Stockroom Worker
Harold Beck & Sons, Newtown, PA — Summer 2010

- Effectively assembled work orders in a timely fashion
- Maintained multiple responsibilities within stockroom while also working in office

PROJECTS

Automated Classification of Histopathological Images

- Collaborated with the Animal Research Laboratory of The Pennsylvania State University
- Developed code in MATLAB to extract important features of images
- Implemented and tested classification scheme to provide a robust system

Aerial Vehicle Remote Control Design

- Reverse engineered original remote control to determine optical communication scheme
- Implementing new software to enhance usability of a new remote control
- Adapting vehicle in order to add accelerometer to allow closed loop feedback control

Non-mechanical Greenhouse Design for Orphanage in Kenya

- Led group of three other students, and teamed with various faculty members
- Coordinated group meetings, consultations with professors, and administered deadlines
- Analyzed difficulties from previous years to improve our design

COMPUTER SKILLS

Excel	Solidworks	C++ Programming
Word	MultiSim	MATLAB

ACTIVITIES & AWARDS

Recipient of Lamartine Hood, Jr. Scholarship Fall 2011
Member of Tau Beta Pi National Engineering Honors Society
Dean's List (2008-2011)