Thermographic Image Analysis Method in Detection of Canine Bone Cancer (Osteosarcoma)

by Maryamsadat Amini, Bachelor of Science

A Thesis Submitted in Partial Fulfillment of the Requirements for the Master of Science Degree

Department of Electrical and Computer Engineering in the Graduate School
Southern Illinois University Edwardsville
Edwardsville, IL

July, 2012
ABSTRACT

THERMOGRAPHIC IMAGE ANALYSIS METHOD IN DETECTION
OF CANINE BONE CANCER (OSTEOSARCOMA)

by

MARYAMSADAT AMINI

Chairperson: Professor Scott E Umbaugh

Introduction: Canine bone cancer is a common type of cancer that grows fast and may be fatal. It usually appears in the limbs which is called "appendicular osteosarcoma." Diagnostic imaging methods such as X-rays, computed tomography (CT scan), and magnetic resonance imaging (MRI) are more common methods in bone cancer detection than invasive physical examination such as biopsy. These imaging methods have some disadvantages; including high expenses, high dose of radiation, and keeping the patient (canine) motionless during the imaging procedures. The research study is to investigate whether thermographic imaging can be used as an alternative diagnostic imaging method for canine bone cancer detection.

Objectives: The main purpose of the research study is to determine the diagnostic usage of thermographic imaging in canine bone cancer. It is also determined which limb region produces the highest classification success rate. In addition, several experiments are performed to investigate whether the hair type of the dogs affects using the thermographic images for bone cancer detection. In order to facilitate the mask creation in the image segmentation stage, an automatic mask creation algorithm is developed in this study.

Results: The best classification success rate in canine bone cancer detection is 80.77% produced by full-limb region with nearest neighbor classification method and normRGB-lum color normalization method. The average of the best correct classification rates for all the different limb region categories with different classification methods is 70.83%. Experimental results from the canine hair type with the thermal images, the long-hair category performance in bone cancer detection is better than the short-hair category performance a maximum of 2.00%, which is not determined to be significant. In addition, the
automatic mask creation algorithm developed in the study produces automatic masks similar to the manual masks with the approximate success rate of 40%. **Methods:** For the canine bone cancer detection, the thermal images are divided into different limb regions, and four color normalization methods are applied on each limb region category. Based on the different conditions, different experiments are implemented. To classify the thermograms of each experiment, several classification methods such as K-nearest neighbor, nearest neighbor, linear discriminant analysis with equal prior probability, linear discriminant analysis with proportional prior probability, and nearest centroid with equal prior probability. The classification methods are applied to classify the objects based on the extracted features analysis. **Conclusion:** It is possible to detect canine bone cancer with the overall classification success rate of 70.83%, which is not enough for a reliable diagnosis. So, further experiments are required to improve the classification rate. The most effective classification method for bone cancer detection is nearest neighbor or K-nearest neighbor with K = 1. In the investigation, the limb region of full-limb generates the highest classification rate. Also, the hair type of the dog minimally affects the thermal image analysis.
ACKNOWLEDGEMENT

First and foremost I offer my sincerest gratitude to my supervisor, Dr. Scott Umbaugh, who has supported me throughout my master study and thesis with his patience and knowledge. I am grateful to him for everything. I could not have imagined having a better advisor.

Besides my advisor, I am extremely thankful to Patrick Solt, Jakia Afruz, and Peng Liu for their help. Also I would like to thank Long Island Veterinary Specialists Dr. Dominic J. Marino and Dr. Catherine A. Loughin for providing fund and thermographic images to do this research.

Last but not the least, I would like to thank my family: my parents Mostafa Amini and Hamideh Hosseini, for giving me life and supporting me financially and spiritually throughout my life.
# TABLE OF CONTENTS

ABSTRACT .......................................................................................................................... I  
ACKNOWLEDGEMENT ........................................................................................................ III  
TABLE OF CONTENTS ........................................................................................................ IV  
LIST OF FIGURES ................................................................................................................ VI  
LIST OF TABLES .................................................................................................................. VIII  
1 CHAPTER 1 INTRODUCTION ............................................................................................. 1  
  1.1 Objectives of the Thesis .............................................................................................. 3  
  1.2 The Thesis Structure ............................................................................................... 3  
2 CHAPTER 2 LITERATURE REVIEW .................................................................................... 5  
  2.1 Current Imaging Methods of Bone Cancer Detection ................................................ 5  
    2.1.1 X-rays ................................................................................................................... 5  
    2.1.2 Computed Tomography (CT scan) ....................................................................... 5  
    2.1.3 Magnetic Resonance Imaging (MRI) ..................................................................... 6  
  2.2 Thermal Imaging (Infrared Thermographic Imaging) .................................................. 6  
    2.2.1 Thermograms ....................................................................................................... 7  
    2.2.2 Thermographic Camera ....................................................................................... 8  
  2.3 Related Cancerous Tumor Detection ......................................................................... 8  
  2.4 Conclusion .................................................................................................................. 11  
3 CHAPTER 3 EXPERIMENTAL MATERIALS AND TOOLS .................................................. 12  
  3.1 Digital Infrared Thermal Imaging System .................................................................. 12  
  3.2 The Thermographic Images ...................................................................................... 13  
  3.3 Border Masks ............................................................................................................ 20  
  3.4 Programs and Algorithms .......................................................................................... 22  
    3.4.1 CVIPtools v5.3 (Computer Vision and Image Processing Tools) ....................... 22  
    3.4.2 CVIP-ATAT (CVIP Algorithm Test and Analysis Tool) ...................................... 22  
    3.4.3 CVIP-FEPC (CVIP Feature Extraction and Pattern Classification) .................. 23  
    3.4.4 Color Normalization Algorithm ......................................................................... 23  
    3.4.5 Partek Discovery Suite ....................................................................................... 24  
    3.4.6 Microsoft Excel ................................................................................................... 24
LIST OF FIGURES

FIGURE Page

2.1: Thermographic image of a dog................................................................. 7


3.2: Standard positioning of a dog for full-limb thermography [Loughin et al.; 2007] ...14

3.3: Standard positioning of a dog for the thermography of the various joint ROIs in the forelimb and hind limb [Loughin et al.; 2007] ................................................................................................................. 15

3.4: Non-cancer thermographic images of a hind limb in which the regions of interests are shown by a red border (a) full-limb, (b) hip, (c) knee, and (d) wrist regions .........................16

3.5: Cancer thermographic images of a forelimb in which the regions of interests are shown by a red border (a) full-limb, (b) shoulder, (c) elbow, and (d) wrist regions ......................17

3.6: (a) The hip region (shown with a red border) of two dogs diagnosed with cancer, (a) with long-hair, and (b) with short-hair ......................................................................................... 19

3.7: (b), (d), (f), and (h) are the manual masks created from (a), (c), (e), and (g) original images, respectively .................................................. 21

4.1: The first category of experiments using original images .................................. 25

4.2: The second category of experiments using color normalized images ............26

4.3: Cross structuring element ........................................................................... 29

4.4: Comparison of automatic and manual masks of one specific original image ....30

4.5: A Flowchart of the automatic mask creation algorithm, and the resultant images corresponding to each step ................................................................. 31

4.6: Four different directions with their corresponding angles ................................ 35

5.1: Masks comparison of full-limb region ........................................................... 44

5.2: Masks comparison of hip/shoulder region ..................................................... 46

5.3: Masks comparison of knee/elbow region ..................................................... 47
5.4: Masks comparison of wrist region

5.5: Best sub-algorithm selected in all the limb regions (based on the XOR and subtraction-energy error metrics)

5.6: CVIP-FEPC best classification success rates for the bone cancer detection with different limb regions images. Note: the best results are obtained by either original or color normalized images.

5.7: Partek best classification success rates for the bone cancer detection with different limb regions images. Note: the best results are obtained by either original or color normalized images.

5.8: Color normalization method evaluation based on the average of the best results

5.9: Feature evaluation based on CVIP-FEPC results with K-nearest neighbor K = 4

5.10: Best classification rates for the bone cancer detection with different limb regions images of long-hair category. Note: the best results are obtained by either original or color normalized image.

5.11: Best classification rates for the bone cancer detection with different limb regions images of short-hair category. Note: the best results are obtained by either original or color normalized images.

5.12: The average of the best classification rates and also the average of the all classification rates for both long-hair and short-hair categories.
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1: The number of thermographic images in each limb region category</td>
<td>17</td>
</tr>
<tr>
<td>3.2: The number of thermographic images in each hair type category</td>
<td>18</td>
</tr>
<tr>
<td>3.3: The number of thermographic images in each limb region based on</td>
<td>19</td>
</tr>
<tr>
<td>the hair type category</td>
<td></td>
</tr>
<tr>
<td>5.1: Best automatic mask creation sub-algorithm applied on the</td>
<td>43</td>
</tr>
<tr>
<td>cancer thermograms of full-limb region</td>
<td></td>
</tr>
<tr>
<td>5.2: Best automatic mask creation sub-algorithm applied on the</td>
<td>45</td>
</tr>
<tr>
<td>cancer thermograms of hip/shoulder region</td>
<td></td>
</tr>
<tr>
<td>5.3: Best automatic mask creation sub-algorithm applied on the</td>
<td>47</td>
</tr>
<tr>
<td>cancer thermograms of knee/elbow region</td>
<td></td>
</tr>
<tr>
<td>5.4: Best automatic mask creation sub-algorithm applied on the</td>
<td>48</td>
</tr>
<tr>
<td>cancer thermograms of wrist region</td>
<td></td>
</tr>
<tr>
<td>5.5: Classification results for original and color normalized images</td>
<td>52</td>
</tr>
<tr>
<td>of all limb regions</td>
<td></td>
</tr>
<tr>
<td>5.6: Best feature sets and data normalization methods of the</td>
<td>53</td>
</tr>
<tr>
<td>experiments using original and color normalized images- all limb</td>
<td></td>
</tr>
<tr>
<td>regions</td>
<td></td>
</tr>
<tr>
<td>5.7: Best classification success rates by using six classification</td>
<td>54</td>
</tr>
<tr>
<td>methods of Partek- all limb regions</td>
<td></td>
</tr>
<tr>
<td>5.8: Classification results for original and color normalized images</td>
<td>56</td>
</tr>
<tr>
<td>of full-limb region</td>
<td></td>
</tr>
<tr>
<td>5.9: Best feature sets and data normalization methods of the</td>
<td>57</td>
</tr>
<tr>
<td>experiments using original and color normalized images- full-limb</td>
<td></td>
</tr>
<tr>
<td>5.10: Classification results for original and color normalized</td>
<td>58</td>
</tr>
<tr>
<td>images of hip/shoulder region</td>
<td></td>
</tr>
<tr>
<td>5.11: Best feature sets and data normalization methods of the</td>
<td>59</td>
</tr>
<tr>
<td>experiments using original and color normalized images- hip/shoulder</td>
<td></td>
</tr>
<tr>
<td>5.12: Classification results for original and color normalized</td>
<td>60</td>
</tr>
<tr>
<td>images of knee/elbow region</td>
<td></td>
</tr>
</tbody>
</table>
5.13: Best feature sets and data normalization methods of the experiments using original and color normalized images - knee/elbow .................................................................61

5.14: Classification results for original and color normalized images of wrist region ........62

5.15: Best feature sets and data normalization methods of the experiments using original and color normalized images - wrist ........................................................................63

5.16: Best classification success rates by using six classification methods of Partek- full-limb ................................................................................................................64

5.17: Best classification success rates by using six classification methods of Partek- hip/shoulder ...........................................................................................................65

5.18: Best classification success rates by using six classification methods of Partek- knee/elbow ...........................................................................................................65

5.19: Best classification success rates by using six classification methods of Partek- wrist..67

5.20: Classification results of long-hair category for original and color normalized images of all limb regions ........................................................................................................74

5.21: Classification results of long-hair category for original and color normalized images of full-limb region ....................................................................................................75

5.22: Classification results of long-hair category for original and color normalized images of hip/shoulder region ....................................................................................................76

5.23: Classification results of long-hair category for original and color normalized images of knee/elbow region ....................................................................................................77

5.24: Classification results of long-hair category for original and color normalized images of wrist region ...........................................................................................................78

5.25: Classification results of short-hair category for original and color normalized images of all limb regions ........................................................................................................79

5.26: Classification results of short-hair category for original and color normalized images of full-limb region ........................................................................................................80

5.27: Classification results of short-hair category for original and color normalized images of hip/shoulder region ........................................................................................................81
5.28: Classification results of short-hair category for original and color normalized images of knee/elbow region ........................................................................................................82

5.29: Classification results of short-hair category for original and color normalized images of wrist region ........................................................................................................................................83
CHAPTER 1
INTRODUCTION

Thermography is a noninvasive diagnostic imaging technique that represents the cutaneous temperature distribution in the form of a color image. The surface temperature measured is originated from local dermal microcirculation, which is under control of the sympathetic nervous system. As heat is not conducted from internal portions of body to the surface, the surface temperature is not affected by heat of deeper parts. However, various conditions of disease or injury can influence local dermal microcirculation and temperature. The correlation between temperature recordings and the presence of a disease or injury is the clinical basis for thermography. Clinically, thermography can be used as a diagnostic tool to improve interpretation of physical examination results, to lead therapeutic procedure, and to evaluate body response to treatment. The diagnostic role of thermography imaging system can be utilized as adjunct test to detect malignant diseases, such as bone cancer [Loughin et al.; 2007].

Canine bone cancer (or osteosarcoma; osteo = bone, sarcoma = cancer) is a common type of cancer that grows fast and may be fatal. Larger breeds are more prone to bone cancer, however it may occur in smaller breeds as well. Bone cancer most frequently occurs in middle-aged and old dogs, and usually in the legs. Osteosarcomas usually appear in the limbs which are called "appendicular osteosarcomas." They develop from within and become painful as they grow outward, which destroys the bone from the inside out. An obvious swelling occurs as the tumor grows and the lameness is expected within one to three months. Tumorous bone is not as strong as normal bone and tends to break with slight injury. These types of breaks, which never heal, are called “pathologic fractures” and may confirm the
diagnosis of the bone tumor. Bone cancer can be detected by radiography (x-rays) in most cases. However, sometimes there are ambiguities and a biopsy can provide incontrovertible proof of the diagnosis. A biopsy is the removal of a sample of tissue or cells from a living subject for further examination. This medical test is performed by a surgeon or an interventional radiologist. In order to determine the presence of a disease, the removed tissue is analyzed under a microscope by a pathologist, and can also be examined chemically [Vetinfo; 2012].

Unfortunately, osteosarcoma is a fast growing tumor. In the most cases, when the tumor is detected by current diagnostic method such as X-rays, and biopsy, the malignancy has already metastasized to the adjacent tissues and organs. This has encouraged many researchers to look for alternative diagnostic imaging systems such as thermography. The investigation here is to determine if thermography imaging technique enables the detection of the abnormality even before its visual detection by providing physiological information, which ultimately accelerates the treatment procedure.

In this research study, image processing and data analysis methods are applied to determine if the thermographic features can be utilized for canine bone cancer diagnosis. The investigation includes three main steps:

- Image segmentation and mask creation
- Feature extraction
- Data normalization and pattern classification

The first step provides only the region of interest (ROI) of the thermal images by mask creation to avoid storage and analysis of unnecessary information. All of the image masks are created manually which is a time-consuming and inefficient task. Therefore, an automatic mask creation algorithm is developed and applied in the research. In the second step, several
features such as histogram and texture features, which are used for analysis and classification, are extracted. In the last step, the extracted information are normalized and classified into two classes of cancer and non-cancer. Also, the classification correctness metrics (or the success rates) such as sensitivity and specificity are evaluated for each of the implemented pattern classification method. The success rates indicate whether the thermal images correlate with biopsy results.

1.1 Objectives of the Thesis

The main purpose of the research study is to determine the diagnostic usage of thermographic imaging in canine bone cancer. The objectives of the research can be categorized as follows:

- Develop an automatic mask creation algorithm to save time in the image segmentation step.
- Identify the usage of thermographic images for bone cancer detection by using all regions of the dogs’ limbs in two cancer and non-cancer classes.
- Determine which canine limb region specifically generates the highest classification success rate in the bone cancer detection.
- Investigate whether the dog’s hair type can affect on the classification results by classifying the images in two classes of short-hair and long-hair.

1.2 The Thesis Structure

Chapter 2 provides a literature review of the previous related research studies, overall explanation of conventional imaging methods and specifically the imaging system applied in the research.
Chapter 3 presents experimental materials and tools used in this study, such as thermographic images, their corresponding masks, software and programs.

Chapter 4 explains all the implemented methods and algorithms including automatic mask creation, feature extraction, and pattern classification besides data analysis.

Chapter 5 encompasses the total experimental results and a comprehensive discussion.

Chapter 6 provides a comprehensive summary with a conclusion.

Chapter 7 presents the future work which will further develop the research study.
CHAPTER 2
LITERATURE REVIEW

2.1 Current Imaging Methods of Bone Cancer Detection

Diagnostic imaging methods such as X-rays, computed tomography (CT scan), and magnetic resonance imaging (MRI) are more common methods in bone cancer detection than invasive physical examination like biopsy. The following is a brief review of these imaging diagnostic techniques, respectively [National Institutes of Health; 2008].

2.1.1 X-rays

X-rays are ionized forms of high energy electromagnetic radiation that penetrate living tissue to create gray images. Depending on the different bone location, four types of x-rays are used for bone cancer diagnosis such as joints x-ray, hands x-ray, and extremities x-ray. In addition, a chest x-ray is used to determine whether lungs are affected by bone cancer metastasis. Although x-ray tests are cheaper and easier than similar imaging tests, radiation exposure and fewer bone details are the concerns patients and doctors have with x-rays.

2.1.2 Computed Tomography (CT scan)

X-rays are emitted into the body region being studied by CT scanner. The doughnut-shaped machine takes a picture of a thin slice of the area in each rotation. Its implementation in two steps makes the CT scan the most detailed internal imaging system available to physicians.

In the basic CT scan, the region of interest is scanned without a contrast agent.
For more detailed requirements, a contrast agent (by “contrast material”, containing iodine) is implemented on the region of interest before the scan. This step is usually used to make a clear and detailed CT picture of organs and structures.

An allergic reaction to iodine is the most common side effect of the CT scans, beside its relatively high dose of radiation.

2.1.3 Magnetic Resonance Imaging (MRI)

MRI uses powerful magnetic and radio frequency fields to reveal a complete image of the region of interest in the body. The energy of radio waves is absorbed by tissues and then represented by extremely detailed images to allow cancer diagnosis. Although, MRI does not use radiation and its contrasting agent is less likely to have an allergic reaction, but its high cost is one of the most important concerns. In addition, the patients, canines in this research, should remain motionless for the imaging test. In this case, the patient can choose to be sedated for scanning which has a slight risk associated with using the sedation medication.

Regarding all the difficulties which are already mentioned (e.g., high dose of radiation), there is a motivation to find an alternative diagnostic method using thermal imaging.

2.2 Thermal Imaging (Infrared Thermographic Imaging)

Thermal Imaging is an imaging method based on detection of infrared radiation by thermal imaging cameras. The detected radiation in the infrared range of the electromagnetic spectrum (9-14 μm) produces images called thermograms. According to the Planck’s law of black-body radiation law [Planck et al.; 1914], all the objects above absolute zero emit infrared radiation, which allows thermography to make them visible, day or night. Since infrared radiation increases with temperature, variations of temperature are represented in
thermograms. Therefore, the temperature distribution of humans and other warm-blooded animals, in this research dogs, can be visible through thermal imaging cameras.

2.2.1 Thermograms

Thermal images, or thermograms, represent the amount of infrared energy emitted, transmitted, and reflected by an object. Since, in a real environment, there are many infrared energy sources, it is difficult to gain a precise temperature of an object. However, a thermal imaging camera can perform an algorithm to interpret that data and generate an approximation of the object’s temperature. This fact can be described simply by this formula [Maldague et al.; 2001]:

\[ \text{Incident Energy} = \text{Emitted Energy} + \text{Transmitted Energy} + \text{Reflected Energy} \]

Where incident energy is the energy of thermograms taken by thermal imaging cameras, however emitted energy is the desired energy to be measured for an accurate temperature data of an object. The energy that crosses through the subject is transmitted energy, and reflected energy is the amount of energy that the surface of the object reflects. Here is an example of a thermal image used in this research in Figure 2.1.

![Thermographic image of a dog](image)

Figure 2.1: Thermographic image of a dog
2.2.2 Thermographic Camera

A thermographic camera or infrared camera is a device that measures infrared radiation in the form of a color image, in which each color represents a different temperature. This kind of camera works similar to a common camera. An ordinary camera takes pictures using visible light with wavelength range of 450-750 nanometers, however a thermographic camera operates in wavelengths as long as 14,000 nm (14µm).

Common clinical uses of an infrared camera are:
- Early detection of breast cancer
- Monitoring changes in overall health
- Monitoring healing process
- Diseases and virus monitoring
- Fever screening

2.3 Related Cancerous Tumor Detection

Various research studies have been dedicated in the thermography area to detect malignancy in different tissues such as the breast [Qi et al.; 2001] and skin [Srinivas et al.; 2003]. Any cancerous tumor increases blood supply and angiogenesis, growth of new blood vessels, and metabolism as compared to normal cells, which causes a temperature gradient rise [Peter et al.; 1989]. The fact that all cancer cells have the same physiological reaction, results in a unified thermal image processing method for cancer diagnosis. The method has three main steps including image segmentation, feature extraction, and pattern classification.

An automatic approach to asymmetry analysis in breast thermograms was developed utilizing automatic segmentation and pattern classification [Qi et al.; 2001]. In order to segment the left and right breasts uniquely, the Hough transform is used to extract the four
feature curves. As feature extraction is essential in pattern recognition and diagnosis in following, different features are extracted from the segments. Joint entropy and higher-order statistics such as variance, skewness, and kurtosis are shown to be more useful features than lower-order statistics for the asymmetry discrimination [Qi et al.; 2002].

Three breast cancer diagnosis methods using thermal images are analyzed on different numbers of patients. The three methods are; (1) the mean temperature of each breast is calculated and compared. If the difference is greater than 0.5 degree C, it is an abnormal condition. (2) In this method each breast is divided into four quadrants. If the mean temperature of one quadrant is 0.5 to 1 degree C higher than the same quadrant of the opposite breast, a score of 0.5 is given to that quadrant. Similarly, scores are calculated for all four quadrants and the scores are added. The score is evaluated on the scale of 0 to 4. If the score is greater than 1, it is considered abnormal. And (3) in this method the addition of mean differences of the quadrants comparing left and right breasts and absolute differences greater than 1 are considered abnormal.

According to all the described methods, the third method has better results than the other two. The analysis on 13 patients had better results with the third method, which has 7 true negatives and 3 false positives of 10 normal patients and 1 false negative and 2 true positives for abnormal patients. If the threshold of normalcy is increased to more than 1.5 degree C, it is possible to find all normal and abnormal conditions correctly in all the patients [Frize et al.; 2002].

In another research paper on the breast cancer diagnosis using thermographic images [Schaefer et al.; 2002], a series of statistical features extracted from the thermal images coupled with fuzzy rule based classification of both left and right breasts. The thermal images are taken from front view and lateral view in some cases. One of the efficient ways to
diagnose the cancer is to compare symmetry of left and right breasts. In case of tumor presence, there will be high flow of blood and changes in vascular pattern, and hence there will be changes in temperature distribution of two breasts. Also, healthy objects are typically symmetric. Therefore, the method is started with breast segmentation from the thermograms, either automatically or manually. The segments are converted to a polar co-ordinates representation for easy calculations. Then a series of statistical features are calculated that are as follows.

1) Basic statistical features include standard deviation temperature, absolute value and median temperature.

2) Moments which describe the center of gravity from geometrical center of breast.

3) Histogram features a normalized histogram of both regions of interest and cross correlation between the histograms.

4) Cross co-occurrence matrix for texture recognition.

5) Mutual information

6) Fourier analysis for Fourier spectrum and differences of absolute values of region of interests.

Each thermal image uses a set of four basic statistical features, four moment features, eight histogram features, eight cross co-occurrence features and two Fourier analysis features. In addition, a laplacian filter for contrast enhancement is applied and another subset of features is calculated from the resulting images. Therefore, the total number of 38 factors for each thermogram is used for the asymmetry description of breast between the two sides. The factors are normalized to the interval [0:1] as comparable units.

Fuzzy rule based classification is used for pattern classification problems. All the feature vectors calculated for each thermogram are given to the fuzzy classifier to distinguish
cancer patients from normal patients. For the classification on the training data, the classification rate is in between 92% and 98%, sensitivity is 83% to 93%, and specificity ranges from 94% to 99%. Whereas on the test set, the maximum classification rate, sensitivity, and specificity are 79.5%, 80%, and 79.5%, respectively. Similar results are obtained for other techniques like mammography, ultrasonography, MRI and DOBI.

Similar methods for determination of melanoma and seborrheic keratosis skin malignancies are implemented [Srinivas et al.; 2003]. Srinivas et al. found the four features including correlation-average, correlation-range, texture-energy-average, and texture-energy-range as the most efficient features in differentiating seborrheic keratosis from melanoma. In general, texture features could identify seborrheic images better than the melanoma success rates.

2.4 Conclusion

According to the mentioned thermographic image analysis successes in the breast and skin malignancy detection, there is an incentive to utilize this imaging method for bone cancer diagnosis. Different physiological characteristics of cancerous tumors distinguish the temperature distribution of abnormal tissue from normal tissue. Based on a specific abnormality and its tissue type the histogram, texture, and color features can be used in the classification process. As features have an important role in the pattern classification process, extraction of efficient features can facilitate and accelerate the diagnostic procedure.
CHAPTER 3

EXPERIMENTAL MATERIALS AND TOOLS

The materials and tools utilized in this research include a thermal imaging system, thermographic images, manually created masks, and six programs. CVIPtools, CVIP-ATAT, CVIP-FEPC, and Color Normalization are the main programs used which are developed at Southern Illinois University at Edwardsville (SIUE). In addition, the Partek Discovery Suite and Microsoft Excel are applied in data retrieval and data analysis phase of the research.

3.1 Digital Infrared Thermal Imaging System

In this research, the digital infrared thermal imaging (DITI) system used is the Meditherm Med2000 IRIS, which is provided by Long Island Veterinary Specialists [LIVS]. It is the only DITI system that is designed for medical applications. Since its design is based only on the clinical environment, it is less expensive than other conventional thermal imaging systems based on speed and expensive optics. Therefore, at less than half the cost of conventional imaging systems, the med2000™ offers accurate measurements, comparable temperature, spatial resolution, simplicity, and longer camera calibration intervals [Meditherm; 2012].

The portable med2000™ consists of two parts: the IR camera and a standard PC or laptop computer. A stand-mounted infrared camera with a focal plane array amorphous silicone microbolometer, and a laptop computer is connected to the camera for real-time data analysis. This system can measure temperatures range of 10° C - 55° C to an accuracy of 0.01° C. Focus adjustment provides small areas down to 75 x 75mm. Thermograms produced by the med2000™ are stored as TIFF images [Meditherm; 2012].
3.2 The Thermographic Images

The thermographic images were taken of each dog, with different views, in the same room with temperature being controlled at 21°C by a centralized air conditioning control system. Since the dogs were kept at the same temperature, it was not necessary to further adapt the dogs to the imaging room temperature. Two technicians wearing latex gloves hold the head and tail for positioning the dogs when imaging, which minimizes thermal artifacts and noise originating from manual contact. In addition, to diminish background artifacts that may be produced by temperature differences in exterior walls, the dogs were placed in front of a uniform interior. The camera was located approximately 1.5 to 4.6 m from the dogs, depending on the region of interest (ROI). The method for imaging includes full left and right lateral limb views of both forelimbs and hind limbs, and cranial and caudal views, which is illustrated in Figure 3.1 and Figure 3.2 [Loughin et al.; 2007].
Figure 3.1: Illustration of the standard positioning of a dog for thermographic study with an infrared camera via lateral views, CR = Cranial, and Ca = Caudal. LFL = Left forelimb. RFL = Right forelimb. LHL = Right hind limb [Loughin et al.; 2007].

Furthermore, images of each limb focused on the joint regions were obtained. There are three joint regions for each limb depicted in Figure 3.3, in which regions 1, 2, and 3 are according to the shoulder/hip, elbow/knee, and wrist respectively.

Figure 3.2: Standard positioning of a dog for full-limb thermography [Loughin et al.; 2007].
Each image was saved with TIFF file type within the software program for further evaluation and review. The program was preset to represent the image temperature range with an 18-shade color map. The color map describes warmer temperatures as white and red and cooler temperatures as blue and black [Loughin et al.; 2007].

In this research study, a total number of 197 thermal images with four different limb regions are used: full-limb region, shoulder/hip region, elbow/knee region, and wrist region. All the images can be also categorized into two classes of dogs with long hair and short hair. The images are taken from 22 dogs with different breeds, genders, and age. The number of images in each region with considering both cancer and non-cancer statuses are 26 for full-limb region, 35 for shoulder/hip region, 84 for elbow/knee region, and 52 for wrist region. Examples of the thermographic images of the four limb regions of a healthy dog (non-cancer status) and of a dog diagnosed with cancer (cancer status) are shown in Figure 3.4 and in Figure 3.5, respectively.
Figure 3.4: Non-cancer thermographic images of a hind limb in which the regions of interests are shown by a red border (a) full-limb, (b) hip, (c) knee, and (d) wrist regions.
In the most of the cases, in this research, only one of the limbs of each dog is affected by bone cancer. Therefore, the opposite limb of the cancerous one is considered as a non-cancer limb. For instance, if the bone cancer is diagnosed only in the right hind of a dog, the left hind images would be used as non-cancer data. According to this fact, approximately half of the total thermographic images are categorized as non-cancer images. The number of images in each category is shown in Table 3.1.
Table 3.1: The number of thermographic images in each limb region category

<table>
<thead>
<tr>
<th>Limb Regions</th>
<th>The Number of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cancer</td>
</tr>
<tr>
<td>Full-limb</td>
<td>13</td>
</tr>
<tr>
<td>Shoulder/Hip</td>
<td>18</td>
</tr>
<tr>
<td>Elbow/Knee</td>
<td>41</td>
</tr>
<tr>
<td>Wrist</td>
<td>27</td>
</tr>
</tbody>
</table>

The canine thermal images can also be classified according to the hair type into long-hair and short-hair classes. The number of images in each class of long-hair and short-hair, considering both cancer and none cancer conditions, is 71, and 126 respectively. The number of thermal images in each class is illustrated in Table 3.2. In addition, detailed information of the number of images in each limb region based on the hair type is depicted in Table 3.3.

Table 3.2: The number of thermographic images in each hair type category

<table>
<thead>
<tr>
<th>Canine Hair Type</th>
<th>The Number of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cancer</td>
</tr>
<tr>
<td>Long-hair</td>
<td>35</td>
</tr>
<tr>
<td>Short-hair</td>
<td>64</td>
</tr>
</tbody>
</table>
Table 3.3: The number of thermographic images in each limb region based on the hair type category

<table>
<thead>
<tr>
<th>Limb Regions</th>
<th>The Number of Images</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short-hair</td>
<td>Long-hair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cancer</td>
<td>Non-cancer</td>
<td>Total</td>
<td>Cancer</td>
<td>Non-cancer</td>
</tr>
<tr>
<td>Full-limb</td>
<td>8</td>
<td>8</td>
<td>16</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Shoulder/Hip</td>
<td>11</td>
<td>10</td>
<td>21</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Elbow/Knee</td>
<td>25</td>
<td>26</td>
<td>51</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Wrist</td>
<td>20</td>
<td>18</td>
<td>38</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

As thermal images represent objects by their temperature and hair is one of the coolest parts, it is difficult to recognize the canine hair type (long or short) by looking at the thermograms. However, one sample of each long-hair and short-hair thermal images are shown in Figure 3.6.

![Figure 3.6](image)

Figure 3.6: (a) The hip region (shown with a red border) of two dogs diagnosed with cancer, (a) with long-hair, and (b) with short-hair
3.3 Border Masks

According to this fact, that only local dermal microcirculation is influenced by a cancerous tumor, focusing on the location of the tumor plays an important role in cancer detection. In order to keep the data of only a specific region of the interest (or a segment) of an image, image segmentation techniques such as border mask creation can be an efficient solution. The data compression application of the border masks is not negligible, especially in the image processing projects involving with a large amount of image data. In this study, the total number of 197 border masks is created manually for the thermographic images with four different limbs; full-limb, shoulder/hip, elbow/knee, and wrist. The masks are created by using CVIPtools, with Utilities-> Create-> Border mask, which is a time-consuming task and may lead to potential errors. An automatic mask creation algorithm is developed to save time and to increase precision in the image segmentation method. Figure 3.7 shows four different limb regions with their corresponding masks created manually.

(a) Full-limb original image
(b) Full-limb mask image
Figure 3.7: (b), (d), (f), and (h) are the manual masks created from (a), (c), (e), and (g) original images, respectively.
3.4 Programs and Algorithms

The programs applied in this study can be categorized in two classes; image processing, and data analysis. There are four image processing programs; CVIPtools, CVIP-ATAT, CVIP-FEPC, and Color Normalization. In order to analyze and visualize the data obtained from image processing techniques, two programs are utilized; Partek Discovery Suite, and Microsoft Excel.

3.4.1 CVIPtools v5.3 (Computer Vision and Image Processing Tools)

CVIPtools version 5.3 is the current Windows version, which was developed by the Computer Vision and Image Processing (CVIP) Laboratory in the Department of Electrical and Computer Engineering of Southern Illinois University Edwardsville (SIUE). This software provides the capability of computer processing of digital images by various imaging functions [CVIPtools; 2012]. CVIPtools allows for the manual processing of one image at a time and produces an instant result.

In this study, the software is applied for manual mask creation. In addition, it is efficient to test an algorithm on one sample image to see the result in a short time. This application can be used as a guideline for algorithm development such as the automatic mask creation algorithm in this research.

3.4.2 CVIP-ATAT (CVIP Algorithm Test and Analysis Tool)

CVIP-ATAT is also a Windows based application that provides automatic processing of a large number of images, in a single run [CVIP-ATAT; 2012]. In addition, several image comparison techniques such as Root-Mean-Square (RMS) Error, Signal-to-Noise-Ratio (SNR), Subtraction-Energy and the Logical XOR are provided by the software. The image
comparison methods application is primarily to compare the resultant images obtained from an algorithm with the ideal image provided by the user. CVIP-ATAT allows the user to define and test various algorithms which can be created by a sequence of built-in imaging functions. Furthermore, this software provides an ability to compare the defined algorithms in one experimental run, by ranking the average and standard deviation of the image comparison results.

In this study, CVIP-ATAT is applied to develop an automatic mask creation algorithm. The algorithm is created and developed to diminish the spent time in the image segmentation step.

3.4.3 CVIP-FEPC (CVIP Feature Extraction and Pattern Classification)

CVIP-FEPC [CVIP-FEPC; 2010] provides both feature extraction and pattern classification in a single run. This software allows the user to select different combinations of features to be extracted from a large group of images. Automatically, after feature extraction of all images, a combination of pattern classification methods selected by the user is implemented. Finally, the software produces a total correctness classification rate, and also the classification rate for each class, by which the sensitivity, and specificity of each set of experiment can be calculated.

3.4.4 Color Normalization Algorithm

The color normalization algorithm is utilized to normalize colors corresponding to the different temperatures in the thermographic images. Four color normalized spaces including luminance (lum), normalized grey (normGrey), normalized RGB (normRGB), and normalized RGB luminance (normRGB-lum) are considered in the color normalization algorithm [Umbaugh, Solt; Jan 2008].
### 3.4.5 Partek Discovery Suite

The Partek Discovery Suite main purpose is to find data patterns, to solve pattern analysis and classification problems, and to provide user friendly data visualization. The abilities are obtained by integration of modern methods of data analysis and visualization with classical statistics. This software allows the user to choose a wide variety of statistical and numerical functions, transformations, and tests [Partek; 2012].

### 3.4.6 Microsoft Excel

Microsoft Excel is a spreadsheet application with several features such as calculation, and ranking which are applicable for this type of research study.
CHAPTER 4

METHODS AND PROCESSES

This research study utilizes three principal processes to investigate whether thermography can be applied for canine bone cancer diagnosis. The processes, which are listed below, contain diverse image processing and data analysis methods.

- Image segmentation and mask creation
- Feature extraction
- Data normalization and pattern classification

The experiments made by the processes, can be classified into two categories. In the first category, the processing applies to the original images, however in the second category the color normalized images are used as inputs. The two categories of the experiments are represented in Figure 4.1 and Figure 4.2.

![Diagram](image_url)

**Figure 4.1**: The first category of experiments using original images
Figure 4.2: The second category of experiments using color normalized images

As it is illustrated in Figure 4.1, in the first category, the original images are applied directly in the feature extraction process. However, in the second category shown in Figure 4.2, the features of the color normalized images are extracted. The color normalization methods allow us to extract mean temperature of the region of interest by histogram mean feature. Since the temperature corresponding to each color changes image by image, we implement color normalization methods to rely on each color as a specific temperature.

In the both categories, Figure 4.1 and Figure 4.2, the three main processes including mask creation, feature extraction, and pattern classification are implemented by using same methods. This chapter is dedicated to describe the methods used in the three processes.
4.1 Mask Creation

Mask creation is one of the image segmentation methods to partition a digital image into two types of segments with values of ‘0’ (black) and ‘1’ (white, in case of eight bits for each pixel the value is 255). Therefore, masks are black and white, binary, images in which pixels with the value of ‘1’ (white) represent the region of interest (ROI). In the image processing, masks are used to restrict the processing to the only ROI (white pixels) of the input images which decreases the required time and memory in the research experiments. In this study, the bone cancer tumor and metastasized regions are considered as the regions of interest, which are determined for each image by experts in Long Island Veterinary Specialists [LIVS]. By using CVIPtools, the masks of the selected regions are created manually, which is time-consuming and inefficient. This fact makes a motivation to develop an automatic mask creation algorithm by using CVIP-ATAT. In addition, diverse image and algorithm comparison methods are applied to evaluate the developed algorithm using different parameters values, which all are provided by CVIP-ATAT.

4.1.1 Automatic Mask Creation Algorithm

The automatic mask creation algorithm is developed by using CVIP-ATAT, and contains four main steps including green band extraction, histogram thresholding segmentation, binary thresholding, and morphological operations.

4.1.1.1 Green band extraction

In the first step of the algorithm, the green band of the image is extracted. In most of the cases, the green band surrounds the regions that only contain the red band, and it also includes some important information about any abnormalities in the adjacent regions. The red band of the thermal images includes the regions with the high temperature, which likely
are tumors and metastasized tissues. Although the algorithm only keeps the green band and discards the red band, by using morphological operations such as closing and dilation we can recover the red band information. The blue band information is not kept for further processes, since the color band corresponds to the regions with the low temperatures which are not cancerous tissues. The output image of the band extraction is a gray level image which is the converted different green values to the related grey values. This method decreases the complexity (time and memory) of the algorithm by focusing only on the data of the probable region of interest.

4.1.1.2 Histogram Thresholding Segmentation

This method segments the image by using the thresholding of histograms. In this technique, based on the specific features, a set of histograms is created. In each of the histograms, the best peak is selected and two thresholds are chosen on the sides of the peak. On the basis of this thresholding of the histogram, the image is split into regions [Umbaugh; 2010]. This method is applied to decrease the number of gray levels and reduces the data to be processed in next steps.

4.1.1.3 Binary Thresholding

Binary thresholding is the third step of the algorithm. This function is selected to convert the obtained grey level images into the binary images. Since the image masks are binary images including only ‘0’ and ‘1’ values, a binary thresholding is needed for the grey to binary conversion. Based on the histogram values of the images, the threshold value of 7 is used to represent only the necessary data with the values of ‘1’.
4.1.1.4 Morphological Operations

In the computer vision, morphology refers to the description of shapes properties on the image. In this study, morphological operations are used to simplify the objects in the segmented images which make it efficient to search for the region of interest. The morphological operations used in the algorithm are opening, erosion, closing, and dilation. These operations can be applied with different structuring element, which determines how the objects borders will be dilated or eroded. The cross structuring element is used in this research, where the equal parameters values of mask size, cross thickness, and cross size change the shape of the structuring element to a rectangle. The structuring element is selected because of the smooth borders in dilated and eroded results in case of using small mask sizes. A cross structuring element with mask size of three, cross thickness of one, and cross size of three is illustrated in Figure 4.3.

![Cross structuring element](image)

Figure 4.3: Cross structuring element
mask size = 1, cross thickness =1, and cross size = 3

Morphological opening is used to eliminate small regions that are not included in the region of interest, and erosion is used to erode the unwanted objects’ boundaries. Also the closing operation is implemented to fill in object holes, and dilation is applied for expanding objects to connect disjoint parts.
In the Figure 4.4, the mask created automatically by CVIP-ATAT and its corresponding manual mask created by CVIP-tools are shown. In addition, there is a flowchart of all the steps of automatic mask creation algorithm alongside with the resultant images corresponding to each step shown in Figure 4.5.
Figure 4.5: A Flowchart of the automatic mask creation algorithm, and the resultant images corresponding to each step.
4.1.2 Image Comparison

Image comparison methods are necessary to be applied to evaluate the similarity measure of automatic masks created and manual masks. Two image comparison methods implemented are the logical XOR and subtraction energy.

The logical XOR function calculates the XOR error metric of automatic mask and manual mask which can be described as follows,

Note that the XOR error metric of 0 represents the ideal automatic mask.

\[
\text{XOR error metric} = \text{(ideal automatic mask, manual mask)}
\]

The subtraction energy operation also produces an error metric for the image comparison. The error metric is calculated by the formula below. By using this image comparison method, an ideal automatic mask can be defined by an error metric of 1.

\[
\text{Subtraction-energy error metric} = \text{energy (subtraction (manual mask, automatic mask))}
\]

\[
\text{Energy} =
\]

4.1.3 Algorithm Comparison

The algorithm comparison is the last section of the development of the automatic mask creation algorithm. Since, there are different sub-algorithms created by using different parameters of the main algorithm, two comparison methods are utilized to find the best parameters for the best results. Therefore, a ranking of the average and standard deviation of the image comparison results is provided for each sub-algorithm to find the best ones.
4.2 Feature Extraction

The second process of the thermal image analysis is feature extraction and data normalization. In this process, several first-order and second-order features are extracted by the CVIP-FEPC application. As the input images of the CVIP-FEPC can be either color normalized or original thermal images, the features to be extracted differ based on the situation. In the case of using original images, four types of histogram features, five types of texture features, and one type of spectral feature are calculated. The histogram features include histogram standard deviation, skew, energy, and entropy. Also, the applied texture features are texture energy, inertia, correlation, inverse difference, and entropy. However, in experiments using color normalized images, all mentioned features as well as histogram mean feature are extracted. The mentioned number for each feature represents only the number of different types of histogram, texture, and spectral features. However, in the real experiments of CVIP-FEPC, the total number of 43 features is extracted for original images and 46 features for the color normalized images. Since each histogram feature is extracted for three bands of red, green, and blue (RGB), the four types of histogram features would be in total 12 features. In case of using color normalized images, the total number of 15 histogram features are extracted. Also, for each type of texture feature (five types), the range and average of the four directions (will be explained in Section 4.2.2) are calculated which results in total 10 texture features. Moreover, the spectral feature is measured for three rings and three sectors. The spectral features of each sector and each ring are extracted for three color bands (RGB) which makes 18 spectral features. The number of 18 spectral features also should be added to three features of the spectral DC value for three color bands, so in total there are 21 spectral features.
4.2.1 Histogram Feature

The image histogram is a graphical representation of the number of pixels for each grey level value. In another word histogram is “a model of the probability distribution of the gray levels” [Umbaugh; 2010]. The histogram features are statistical-based features which contain information about the grey-level distribution for the image. The histogram features are measured for three color bands of red, green, and blue (RGB) in case of using color images. Mean, standard deviation, skew, energy, and entropy are the features based on the first-order histogram probability.

The mean histogram feature is the average of the grey-level values, and the standard deviation feature describes the contrast of the image. The skew histogram calculates the asymmetry of the mean in the histogram. The energy of histogram is a maximum value of 1 for an image with a constant value and decreases as the number of grey-level values increase in the image. The histogram entropy measures the number of bits needed to code the data of each pixel [Umbaugh; 2010].

4.2.2 Texture Features

Texture features can be measured by using the second-order histogram of the gray levels. The second-order histogram techniques are also referred as gray-level co-occurrence matrix or grey-level dependency matrix techniques which provide information about pairs of pixels and their related gray levels. Two parameters of distance and angle are necessary for texture features application. The distance refers to the pixel distance between the pairs of pixels, and the angle represents the angle between pixels in each pair [Umbaugh; 2010]. In this study, the texture features are extracted based on a six pixel distance in each pair; then the average and the range of the texture features extracted in four directions are used for the
next processes. The pixel distance value in texture features is dependent upon the scale of the regions of interest by pixel. According to the previous thermographic image analysis projects in elbow dysplasia [Umbaugh, Solt; Sep 2009] and Chiari malformation [Umbaugh et al.; Jan 2010, Umbaugh et al.; May 2010, Umbaugh et al.; June 2011] with a similar scale for the regions of interest, the six pixel distance produce efficient texture features values for higher object classification success rates. The directions to be considered for the texture features are horizontal (0° and 180°), vertical (90° and 270°), left diagonal (135° and 315°), and right diagonal (45° and 225°), which are shown in Figure 4.6. Five texture features generated by the methods are provided by CVIP-FEPC, which are energy, inertia, correlation, inverse difference, and entropy.

![Figure 4.6: Four different directions with their corresponding angles](image)

The energy calculates the distribution across the grey level values which represents smoothness of the image. The inertia provides information about the contrast, while the correlation measures the similarities between pixels. The inverse difference calculates the
homogeneity, and the entropy provides the information content which is inversely related to
the energy [Umbaugh; 2010].

4.2.3 Spectral Feature

Spectral features or frequency/sequency-domain based features are based on the power
metric. The power can be calculated by the below formula which is the magnitude of the
spectral components squared,

\[
\text{Power} = |T(u, v)|^2
\]

Where \( T(u, v) \) refers to any of the transforms, which the Fourier transform is typically
used.

The standard spectral features are to extract the power of several spectral regions such
as rings, sectors, or boxes. The spectral features can provide texture information of the
image. The regions with high power are related to the low frequency which represents coarse
textures. As frequency gets higher, the power reduces and texture will be finer. In this
research, the spectral feature is extracted from ring and sector regions with the parameter of
three which is the number of the rings and sectors applied in the spectral domain based image
[Umbaugh; 2010].

In this research study, the spectral features of three sectors and three rings are extracted.
Since, color thermal images are applied in the research, each of the sector or ring spectral
features are calculated for three color bands (RGB). Also three spectral DC values for three
color bands are measured as spectral features.
4.3 Data Normalization and Pattern Classification

The extracted features are the subjects for the data normalization methods to be exchanged to the comparable units for the pattern classification process. The normalized features values are used for the similarity and/or distance measurement to classify image objects into either cancer or non-cancer classes.

4.3.1 Data Normalization

The applied data normalization methods are standard normal density normalization and softmax scaling normalization.

4.3.1.1 Standard Normal Density Normalization

One of the common statistical-based methods for data normalization is standard normal density normalization. In this method, each vector component subtracts the mean and divides by the standard deviation. This can be explained as follows, assumed a set of $k$ features vectors [Umbaugh; 2010],

$$F_j = \{F_1, F_2, \ldots, F_k\}, \text{ with } n \text{ features in each vector.}$$

$$F_j$$

for $j = 1, 2, \ldots, k$

Means $m_i = -$ for $i = 1, 2, \ldots, n$

Standard deviation $= -$ for $i = 1, 2, \ldots, n$

Now, each feature component subtracts mean and divides by the standard deviation:

$$f_{ij\text{SND}} = -$$
The resultant distribution on each vector component is called standard normal density (SND) which would result values in [0, 1] interval.

4.3.1.2 Softmax Scaling Normalization

The softmax scaling is one of the nonlinear data normalization methods, that may be desired in cases without even data distribution about the mean. In this method, the data is compressed into the range of 0-1, and changes the spread and/or shape of the data distribution. This method needs two steps, given \( m_i \) as mean, \( f_{ij} \) as each feature component, as standard deviation, and \( r \) as a user defined factor [Umbaugh; 2010]:

\[
\text{STEP1 } y = \frac{f_{ij}}{m_i}
\]

\[
\text{STEP2 } f_{ijSMC} = \frac{y}{r} \quad \text{for all } i, j
\]

The first step is similar to the standard normal density normalization, but with a factor of \( r \) which is defined by the user to determine the range of feature values \( (f_{ij}) \). For small values of \( y \) with respect to \( f_{ij} \), the process is relatively linear, and for \( y \) values farther away from the mean, the data is compressed exponentially [Umbaugh; 2010].

4.3.2 Pattern Classification

In the pattern classification process, the image objects should be divided into a training set and a test set. The training set is used to develop the classification algorithm, and the test set is used to test the algorithm. Both the training and test sets should include all types of images in the application domain, otherwise the success rate measured by test set is not a good predictor for the application. The testing method used in this research study is leave-one-out method. In this method, all but one of the image samples are used in training set, and then it is tested on the one that was left out. This method is implemented as many times as
there are samples. The number of sample images that are classified correctly represents the success rate for the testing. There are two programs used for the pattern classification which are CVIP-FEPC and Partek, in both the leave-one-out testing method is applied [Umbaugh; 2010].

4.3.2.1 CVIP-FEPC Classification Methods

There are three classification methods available in CVIP-FEPC including Nearest Neighbor, \( K \)-Nearest Neighbor, and Nearest Centroid. Among them, \( K \)-Nearest Neighbor with a \( K \) value of 4 is the method used in this investigation.

The \( K \)-Nearest Neighbor is one of the simplest algorithms for classification of a sample from the test set. In this method, the sample of interest in the test set is compared to every sample in the training set by using, in this study, the Euclidean distance measure. Then, the unknown feature vector is assigned to the class that occurs most often in the set of \( K \)-Neighbors.

The Euclidean distance is measured by the square-root of the sum of the least squares of differences between vector components. This can be explained as follows [Umbaugh; 2010],

Given two feature vectors \( A \) and \( B \),

\[
A = \quad \text{and} \quad B =
\]

Then the Euclidean distance is calculated by

\[
D_E(A, B) = \quad \text{to} \quad \text{to}
\]
4.3.2.2 Partek Discovery Suite Classification Methods

There are a number of classification techniques available in the Partek Discovery Suite software. However, in this research study, six classification methods are applied which are provided by the Partek software. The classification methods are linear discriminant analysis with equal prior probability (LDA), linear discriminant analysis with proportional prior probability (LDAP), nearest centroid with equal prior probability (NCE), and also K-nearest neighbor (KNN) with K=1, 3, and 5.

The linear discriminant analysis is a method used in statistics and pattern classification to find a linear combination of features which characterizes or categorizes two or more classes of objects. The linear combination is created by using covariance matrix and samples mean of a known set of data called training set. The proportional prior probability option is selected when the classes population size are unequal, otherwise the equal prior probability may be selected.

The nearest centroid method is used to compare the unknown samples to the centroid from the samples in the training set (known samples). The centroids for each class can be measured by averaging of each vector component (feature) in the training set. The comparison may be done by either distance or similarity measure. In this study, the Euclidean distance measure is applied as a comparison metric [Umbaugh; 2010].

4.3.2.3 Pattern Classification Evaluation Metrics

There are two success metrics of sensitivity and specificity which are often used in biomedical image analysis with two classes of diseased and healthy. These measures can also be used in any object classification with a binary result. There are four definitions for this medical classification:
- True Positive (TP): sick person classified as sick correctly.
- False Positive (FP): healthy person classified as sick mistakenly.
- True Negative (TN): healthy person classified as healthy correctly.
- False Negative (FN): sick person classified as healthy mistakenly.

Then the sensitivity and specificity are defined as follows:

Sensitivity = 

Specificity = 
CHAPTER 5
RESULTS AND DISCUSSIONS

The results of this research study can be categorized into three main sections which follow the research objectives. In the first section, the results of the automatic mask creation algorithm are presented and discussed. The second section is for the detection of canine bone cancer by using thermographic images. Finally, in the last section, the investigation related to the effect of the canine hair type in bone cancer diagnosis is discussed based on the outputs.

5.1 Automatic Mask Creation Algorithm

The algorithm of the automatic mask creation (see Figure 4.5) is developed based on the detection of the tumor and metastasized regions by green band extraction. Therefore, the algorithm is tested only on the cancer images which are taken from different affected limb regions. The images of each limb region are subject to the algorithm, separately, which uses four different algorithm implementation on the four limb regions; full-limb, hip/shoulder, knee/elbow, and wrist.

The main algorithm of the automatic mask creation includes 16 variety of sub-algorithms based on the different parameter values of morphological operations. The best sub-algorithm is selected according to the XOR error metric and subtraction-energy error metric (see Section 4.1.2). Since the distinction of the sub-algorithms is defined only by different parameters of the morphological operations, only the morphological operations for each sub-algorithm are in the table of results. In the tables, letters ‘O’, ‘C’, ‘D’, and ‘E’ represent opening, closing, dilation, and erosion morphological operations, respectively. In front of each of the letters, there are three parameters which, from left to right, are mask size, cross thickness, and cross size. The success rates corresponding to the selected sub-algorithm
are defined based on threshold values of XOR and subtraction-energy error metrics. In the results, the threshold values of 0.8 and 0.85 are used for XOR and subtraction-energy error metrics, respectively, where the created masks with XOR error metrics of 0.8 or less and masks with subtraction-energy error metrics of 0.85 or more are deemed as successful. Therefore, the success rate in the tables is the number of successful automatic masks divided by the total number of manual masks.

The automatic mask creation algorithm (16 sub-algorithms) is tested on the 13 cancer thermal images of full-limb region. The best sub-algorithm out of the 16 sub-algorithms is selected based on the mentioned error metrics, and the results are shown in Table 5.1. According to the results of the Table 5.1, the XOR error metric of the sub-algorithm, 1.65, is not good (a good XOR error should be close to zero). However the average of the subtraction-energy error of 0.85 is satisfying, very close to one, and with success rate of 46.15%.

Table 5.1: Best automatic mask creation sub-algorithm applied on the cancer thermograms of full-limb region

<table>
<thead>
<tr>
<th>Sub-algorithm</th>
<th>XOR Error Metric</th>
<th>Subtr-energy** Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Std. Dev.*</td>
</tr>
<tr>
<td>O(3,3,3), E(5,5,5), E(7,7,7), D(7,5,5), D(5,5,5), C(3,3,3)</td>
<td>1.65</td>
<td>1.44</td>
</tr>
</tbody>
</table>

* The Std. Dev. is the abbreviation of standard deviation.
** The Subtr-energy is the abbreviation of subtraction-energy.

In Figure 5.1, a sample of automatic mask of the full-limb region (a) created by the best sub-algorithm is shown and compared to its corresponding manual mask (b). In addition, the
original image (c) related to the masks is displayed to figure out which parts of the original image are missed or kept in the automatic mask.

(a) Automatic Mask  
(b) Manual Mask  
(c) Original Image

Figure 5.1: Masks comparison of full-limb region

The averages of the XOR and subtraction-energy error of the sample shown in Figure 5.1 are 0.23 and 0.93, respectively. Although the error metrics are very good and show a very similar automatic mask to the manual one, there is still one problem to use the automatic mask in the CVIP_FEPC application. The mask to be used in the CVIP_FEPC should not include more than one object. As shown in Figure 5.1(a), there are two objects in the automatic mask which can be changed to one object mask by a solution will be explained in the future scope chapter.

In the Table 5.2, the results of the algorithm implemented on the 18 hip/shoulder cancer images are illustrated. The automatic masks of the region are good based on the XOR error
average and its success rate which is 55.55%. Nevertheless, the subtraction-energy error average of 0.67 is not satisfying also with success rate of 27.77%.

Table 5.2: Best automatic mask creation sub-algorithm applied on the cancer thermograms of hip/shoulder region

<table>
<thead>
<tr>
<th>Sub-algorithm</th>
<th>XOR Error Metric</th>
<th>Subtr-energy Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>O(3,3,3), E(5,5,5), E(7,7,7), D(7,5,5), D(5,5,5), C(3,3,3)</td>
<td>0.73</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Figure 5.2 is to compare one of the masks created automatically of the hip/shoulder region to its related manual mask. This automatic mask has the average XOR error of 0.2 and the average subtraction-energy error of 0.88. The created mask is one of the best in this specific region which can be used for image segmentation step in CVIP_FEPC.
The results of the algorithm implementation on the 41 knee/elbow region images are depicted in Table 5.3. The results corresponding to both XOR and subtraction-energy error metric are not pleasant. Although the subtraction-energy average of the knee/elbow results is more than hip/shoulder results, the success rate of knee/elbow is less than the hip/shoulder region. The reason can be caused by the data distribution in the experiments. For example, if the data distribution of the knee/elbow subtraction-energy results is not close to a normal distribution, such a success rate may be seen.
Table 5.3: Best automatic mask creation sub-algorithm applied on the cancer thermograms of knee/elbow region

<table>
<thead>
<tr>
<th>Sub-algorithm</th>
<th>XOR Error Metric</th>
<th>Subtr-energy Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>O(3,3,3), E(5,5,5), E(7,7,7), D(7,5,5), D(5,5,5), C(3,3,3)</td>
<td>1.08</td>
<td>0.80</td>
</tr>
</tbody>
</table>

The automatic and manual masks created of one the knee/elbow thermal images are depicted in Figure 5.3. The error metrics related to the automatic mask are 0.49 for XOR error and 0.74 for subtraction-energy error.
The success rates shown in Table 5.4 represent the results of the automatic masks created from 27 wrist region images. The automatic mask creation algorithm produces the highest success rates and the best XOR and subtraction-energy error metrics by using this region’s images. The best results may be caused by well-focused images of the wrist, in which each image is very similar to the region of interest with fewer redundant areas.

Table 5.4: Best automatic mask creation sub-algorithm applied on the cancer thermograms of wrist region

<table>
<thead>
<tr>
<th>Sub-algorithm</th>
<th>XOR Error Metric</th>
<th>Subtr-energy Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>O(3,3,3), E(5,5,5), E(7,7,7), D(7,5,5), D(5,5,5), C(3,3,3)</td>
<td>0.65</td>
<td>0.50</td>
</tr>
</tbody>
</table>

An automatic mask sample of the wrist region besides with its related manual mask and original image are displayed in Figure 5.4.
The average XOR and subtraction-energy error metrics of the automatic mask (Figure 5.4) are 0.14 and 0.85, respectively, which make the automatic mask a good approximation to manual mask.

The best sub-algorithm selected in all the limb regions is the same which is shown completely in Figure 5.5. The sub-algorithm is independent of the different limb region types. However, the success rates change image by image and limb region by limb region.
5.2 Canine Bone Cancer Detection

The experiments for the canine bone cancer detection using thermographic images are divided into two main sections:

Section 1: The detection of bone cancer using thermographic images of all limb regions and classifying the images as *cancer* and *non-cancer*.

Section 2: The bone cancer diagnosis by applying the thermograms of each limb region such as full-limb, hip/shoulder, knee/elbow, and wrist, and then classifying each region images to *cancer* and *non-cancer*.

For each of the sections above, the experiments are done on both the original and color normalized images. As mentioned before, by applying the color normalized images, another histogram feature, namely the histogram mean which provides the average temperature, can be also extracted.
5.2.1 Results and Discussions of Section 1

In this section, the thermal images including all limb regions are applied to detect canine bone cancer. The experiments are performed using CVIP-FEPC and the Partek software. For CVIP-FEPC, in this research, only the classification method of K-nearest neighbor (KNN) is applied; however more classification methods are utilized by the Partek application. In both of the CVIP-FEPC and the Partek software, full leave-one-out is the selected testing method.

5.2.1.1 CVIP-FEPC Experiments

To diagnose canine bone cancer, using thermal images of all limb regions, five different experiments are performed; one experiment includes original images and in the other four experiments we apply four color normalization methods.

In these experiments the K-nearest neighbor with K=4 is used as the classification method; Also, Euclidean distance is used as the distance measure, and standard normal density normalization and softmax scaling normalization with r=1 are applied as the data normalization techniques. In the experiments where the original images are used, the total number of ten types of features including four types of histogram features, five types of texture features, and a type of spectral feature are extracted (see Section 4.2). The texture distance of six pixels is used for the texture features. In the experiments performed by importing color normalized images, all the features mentioned above as well as the histogram mean feature are extracted, resulting in a total of 11 types of features. Therefore, in the experiment with original images, the classification is exploited using all combinations of the 10 types of features, a total of \( = 1024 \) experiments. The same procedure is implemented for the color normalized images but with 11 types of features which has \( = 2048 \) different
The classification success rate refers to the number of the objects classified correctly into either cancer or non-cancer classes divided by the total number of objects of the both classes. Also, the sensitivity metric is the accuracy of the disease (cancer) prediction and the specificity measure the accuracy of prediction of the non-cancer images.

According to the Table 5.5, the overall results obtained by color normalized images are higher than the experiment result using original images. Therefore, applying color
normalization on the input images is preferred for the next experiments. The best classification success rate which is corresponding to the normRGB images, 61.93%, is not a good rate for reliable bone cancer detection.

To select the most informative features in each of the experiments, we report the related feature set to each of them in the Table 5.6.

Table 5.6: Best feature sets and data normalization methods of the experiments using original and color normalized images- all limb regions

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Normalization Method</th>
<th>Number of Images per Class</th>
<th>Classification Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>_______ (original image)</td>
<td>Texture Energy Texture Correlation Histogram StdDev Histogram Skew</td>
<td>Standard Normal</td>
<td>cancer: 99 non-cancer: 98</td>
<td>59.39%</td>
</tr>
<tr>
<td>lum</td>
<td>Texture Correlation Histogram Mean Histogram Entropy</td>
<td>Soft-max, r = 1</td>
<td>cancer: 99 non-cancer: 98</td>
<td>61.42%</td>
</tr>
<tr>
<td>normGrey</td>
<td>Histogram Mean Histogram StdDev Histogram Skew</td>
<td>Standard Normal</td>
<td>cancer: 99 non-cancer: 98</td>
<td>57.36%</td>
</tr>
<tr>
<td>normRGB</td>
<td>Texture Energy Texture Entropy Histogram StdDev</td>
<td>Soft-max, r = 1</td>
<td>cancer: 99 non-cancer: 98</td>
<td>61.93%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>Texture Energy Histogram StdDev Histogram Skew</td>
<td>Soft-max, r = 1</td>
<td>cancer: 99 non-cancer: 98</td>
<td>60.91%</td>
</tr>
</tbody>
</table>

As it is shown in the Table 5.6, among all the texture and histogram features, texture energy, histogram standard deviation, and histogram skew mostly are extracted in the best classification success rates.

There is an assumption that the reason behind the low success rates may be due to the diverse variety of bone data, which are taken from dogs with different breeds, genders, and
limb regions. Moreover, the bone images include all different limb regions with different tissue types. One way to address the problem above is to divide the experiment based on gender, breed, and/or limb region. Performing gender and/or breed specific experiments is difficult as the bone cancer is a rare disease with limited amount of available data. However, the amount of data for each limb region makes it possible to implement region based experiments. In addition to all the main solutions, implementation of the other data classification methods using Partek may result in better success rates.

5.2.1.2 Partek Experiments

The Partek software is used to get the results for other object classification methods such as linear discriminant analysis with equal prior probability (LDAE), linear discriminant analysis with proportional prior probability (LDAP), nearest centroid with equal prior probability (NCE), and also K-nearest neighbor (KNN) with K=1, 3, and 5. The new six classification methods are implemented on the sets of features identified by CVIP-FEPC best results (Table 5.6). Therefore, the features sets of the five experiments (original and four color normalization) are imported to the Partek application. In the Table 5.7, the best results of the object classification methods applied by Partek are shown for each experiment, in which the testing method is leave-one-out. Although all the six classification methods are applied on each of the experiments, only the classification method which results the best success rate is reported in the Table 5.7.
Table 5.7: Best classification success rates by using six classification methods of Partek - all limb regions

<table>
<thead>
<tr>
<th>Color Norm* Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Norm* Method</th>
<th>Classification Method</th>
<th>Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>Texture Energy</td>
<td>Standard Normal</td>
<td>KNN K = 1</td>
<td>57.36%</td>
<td>55.00%</td>
<td>60.00%</td>
</tr>
<tr>
<td></td>
<td>Texture Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram StdDev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lum</td>
<td>Texture Correlation</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 1</td>
<td>61.42%</td>
<td>65.00%</td>
<td>58.00%</td>
</tr>
<tr>
<td></td>
<td>Histogram Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram StdDev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normGrey</td>
<td>Histogram Mean</td>
<td>Standard Normal</td>
<td>LDAE</td>
<td>59.39%</td>
<td>58.00%</td>
<td>61.00%</td>
</tr>
<tr>
<td></td>
<td>Histogram StdDev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normRGB</td>
<td>Texture Energy</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 1</td>
<td>61.42%</td>
<td>61.00%</td>
<td>62.00%</td>
</tr>
<tr>
<td></td>
<td>Texture Entropy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram StdDev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>Texture Energy</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 3</td>
<td>60.41%</td>
<td>63.00%</td>
<td>58.00%</td>
</tr>
<tr>
<td></td>
<td>Histogram StdDev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The best correct classification rate is 61.42%, which is corresponding to the normRGB experiment with K-nearest neighbor classification, K = 1.

5.2.2 Results and Discussions of Section 2

Based on the low success rate of the previous experiments, the experiments are divided into four different limb regions such as full-limb, hip/shoulder, knee/elbow, and wrist. The experiments are implemented using CVIP-FEPC and Partek with the same methods implemented in the Section 1 to follow a constant procedure to evaluate results only based on the region division.
5.2.2.1 CVIP-FEPC Experiments

The object classification method of K-nearest neighbor with K=4 is applied, where the Euclidean is the distance measure. The testing method is full leave-one-out. Also, as mentioned in Section 1, the applied data normalizations are standard normal density normalization and softmax scaling normalization with r=1. As well, the extracted features are the same, for the original images the total number of 10 types of features, and for the color normalized images, 11 types of features.

The first examined limb region is full-limb (see Figure 5.1), and Table 5.8 is to show the best classification success rates corresponding to this limb region. All together 26 full-limb images are used in the experiments for which the half of the images belongs to the cancer class.

Table 5.8: Classification results for original and color normalized images of full-limb region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>cancer: 13 non-cancer: 13</td>
<td>full-limb</td>
<td>65.38%</td>
<td>61.54%</td>
<td>69.23%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 13 non-cancer: 13</td>
<td>full-limb</td>
<td>69.23%</td>
<td>61.54%</td>
<td>76.92%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 13 non-cancer: 13</td>
<td>full-limb</td>
<td>69.23%</td>
<td>61.54%</td>
<td>76.925</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 13 non-cancer: 13</td>
<td>full-limb</td>
<td>73.08%</td>
<td>69.23%</td>
<td>76.92%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 13 non-cancer: 13</td>
<td>full-limb</td>
<td>76.92%</td>
<td>84.62%</td>
<td>69.23%</td>
</tr>
</tbody>
</table>
In the experiment of full-limb region, normRGB-lum color normalization method result has the best overall success rate, 76.92%, than other methods. The success rate of normRGB method is the second best at 73.08%.

The feature sets extracted to obtain the best success rates are illustrated in the Table 5.9. The feature which plays an important role in the cancer and non-cancer classification of the full-limb images is the histogram skew feature. This feature is extracted in four out of five experiments with the best classification results.

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Normalization Method</th>
<th>Number of Images per Class</th>
<th>Classification Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>________ (original image)</td>
<td>Histogram Skew Histogram Entropy</td>
<td>Standard Normal</td>
<td>cancer: 13 non-cancer: 13</td>
<td>65.38%</td>
</tr>
<tr>
<td>lum</td>
<td>Histogram Skew</td>
<td>Soft-max, r = 1</td>
<td>cancer: 13 non-cancer: 13</td>
<td>69.23%</td>
</tr>
<tr>
<td>normGrey</td>
<td>Histogram StdDev</td>
<td>Standard Normal</td>
<td>cancer: 13 non-cancer: 13</td>
<td>69.23%</td>
</tr>
<tr>
<td>normRGB</td>
<td>Spectral Texture Energy Texture Inertia Histogram Mean Histogram Skew Histogram Energy</td>
<td>Soft-max, r = 1</td>
<td>cancer: 13 non-cancer: 13</td>
<td>73.08%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>Texture InvDiff Histogram Skew Histogram Entropy</td>
<td>Standard Normal</td>
<td>cancer: 13 non-cancer: 13</td>
<td>76.92%</td>
</tr>
</tbody>
</table>

In the Table 5.10, the best classification success rates for the hip/shoulder images (see Figure 5.2) are stated in which lum color normalization makes the highest success rate of
74.28%. In the experiments, the feature extraction and classification techniques are implemented on 18 cancer and 17 non-cancer images. Also, the best feature sets producing the best success rates are displayed in the Table 5.11.

Table 5.10: Classification results for original and color normalized images of hip/shoulder region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>cancer: 18 non-cancer: 17</td>
<td>hip/shoulder</td>
<td>62.86%</td>
<td>61.11%</td>
<td>64.71%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 18 non-cancer: 17</td>
<td>hip/shoulder</td>
<td>74.28%</td>
<td>83.33%</td>
<td>64.71%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 18 non-cancer: 17</td>
<td>hip/shoulder</td>
<td>60.00%</td>
<td>66.67%</td>
<td>52.94%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 18 non-cancer: 17</td>
<td>hip/shoulder</td>
<td>57.14%</td>
<td>61.11%</td>
<td>52.94%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 18 non-cancer: 17</td>
<td>hip/shoulder</td>
<td>68.57%</td>
<td>77.78%</td>
<td>58.82%</td>
</tr>
</tbody>
</table>
Table 5.11: Best feature sets and data normalization methods of the experiments using original and color normalized images- hip/shoulder

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Normalization Method</th>
<th>Number of Images per Class</th>
<th>Classification Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>Texture Inertia</td>
<td>Standard Normal</td>
<td>cancer: 18</td>
<td>62.86%</td>
</tr>
<tr>
<td></td>
<td>Texture InvDiff</td>
<td></td>
<td>non-cancer: 17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Texture Entropy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lum</td>
<td>Histogram StdDev</td>
<td>Soft-max, r = 1</td>
<td>cancer: 18</td>
<td>74.28%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>non-cancer: 17</td>
<td></td>
</tr>
<tr>
<td>normGrey</td>
<td>Histogram Skew</td>
<td>Soft-max, r = 1</td>
<td>cancer: 18</td>
<td>60.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>non-cancer: 17</td>
<td></td>
</tr>
<tr>
<td>normRGB</td>
<td>Texture Correlation</td>
<td>Soft-max, r = 1</td>
<td>cancer: 18</td>
<td>57.14%</td>
</tr>
<tr>
<td></td>
<td>Histogram Mean</td>
<td></td>
<td>non-cancer: 17</td>
<td></td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>Texture Correlation</td>
<td>Soft-max, r = 1</td>
<td>cancer: 18</td>
<td>68.57%</td>
</tr>
<tr>
<td></td>
<td>Histogram Mean</td>
<td></td>
<td>non-cancer: 17</td>
<td></td>
</tr>
</tbody>
</table>

As it is shown in the Table 5.11, in the experiments on the hip/shoulder images, there is no specific feature which is used in the majority of the experiments.

The knee/elbow limb region images (see Figure 5.3) are the largest region based data in this research with the total number of 84 images including 41 cancer and 43 non-cancer images. There is Table 5.12 containing the best success rates of the experiments, in which lum and norm-Grey color normalization methods classified the knee/elbow images better than other color normalization methods. Table 5.13 illustrates a wide variety of features with no common feature in the experiments.
Table 5.12: Classification results for original and color normalized images of knee/elbow region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>________ (original image)</td>
<td>cancer: 41 non-cancer: 43</td>
<td>knee/elbow</td>
<td>61.90%</td>
<td>60.98%</td>
<td>62.79%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 41 non-cancer: 43</td>
<td>knee/elbow</td>
<td>65.48%</td>
<td>68.29%</td>
<td>62.79%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 41 non-cancer: 43</td>
<td>knee/elbow</td>
<td>65.48%</td>
<td>68.29%</td>
<td>62.79%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 41 non-cancer: 43</td>
<td>knee/elbow</td>
<td>59.52%</td>
<td>53.66%</td>
<td>65.12%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 41 non-cancer: 43</td>
<td>knee/elbow</td>
<td>64.28%</td>
<td>65.85%</td>
<td>62.79%</td>
</tr>
</tbody>
</table>
Table 5.13: Best feature sets and data normalization methods of the experiments using original and color normalized images- knee/elbow

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Normalization Method</th>
<th>Number of Images per Class</th>
<th>Classification Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>Histogram StdDev, Histogram Entropy</td>
<td>Soft-max, $r = 1$</td>
<td>cancer: 41, non-cancer: 43</td>
<td>61.90%</td>
</tr>
<tr>
<td>lum</td>
<td>Histogram Mean, Histogram StdDev, Histogram Energy</td>
<td>Soft-max, $r = 1$</td>
<td>cancer: 41, non-cancer: 43</td>
<td>65.48%</td>
</tr>
<tr>
<td>normRGB</td>
<td>Histogram Skew</td>
<td>Soft-max, $r = 1$</td>
<td>cancer: 41, non-cancer: 43</td>
<td>59.52%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>Texture InvDiff, Histogram Skew</td>
<td>Soft-max, $r = 1$</td>
<td>cancer: 41, non-cancer: 43</td>
<td>64.28%</td>
</tr>
</tbody>
</table>

The last limb region analyzed in this study is the wrist of the dogs (see Figure 5.4). The results corresponding to each color normalization method is shown in the Table 5.14. As seen in this table, the classification success rate is higher for the wrist experiments compared to other limb regions. This may be due to the high similarity that exists between the forelimb wrist and hind limb wrist, which is not the case for other regions such as knee/elbow and hip/shoulder.
Table 5.15 illustrates the feature sets for the best resulting classification experiments. It can be observed in the table that texture correlation and histogram energy features are common in most of the best resulting experiments indicating that these features are very useful and informative for the task in hand.

Table 5.14: Classification results for original and color normalized images of wrist region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>cancer: 27 non-cancer: 25</td>
<td>wrist</td>
<td>71.15%</td>
<td>77.78%</td>
<td>64.00%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 27 non-cancer: 25</td>
<td>wrist</td>
<td>73.08%</td>
<td>85.19%</td>
<td>60.00%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 27 non-cancer: 25</td>
<td>wrist</td>
<td>69.23%</td>
<td>81.48%</td>
<td>56.00%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 27 non-cancer: 25</td>
<td>wrist</td>
<td>73.08%</td>
<td>81.48%</td>
<td>64.00%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 27 non-cancer: 25</td>
<td>wrist</td>
<td>67.31%</td>
<td>74.07%</td>
<td>60.00%</td>
</tr>
</tbody>
</table>
Table 5.15: Best feature sets and data normalization methods of the experiments using original and color normalized images - wrist

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Normalization Method</th>
<th>Number of Images per Class</th>
<th>Classification Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>Texture Correlation</td>
<td>Soft-max, r = 1</td>
<td>cancer: 27</td>
<td>71.15%</td>
</tr>
<tr>
<td></td>
<td>Texture Entropy</td>
<td></td>
<td>non-cancer: 25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Energy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lum</td>
<td>Texture Correlation</td>
<td>Standard Normal</td>
<td>cancer: 27</td>
<td>73.08%</td>
</tr>
<tr>
<td></td>
<td>Histogram Mean</td>
<td></td>
<td>non-cancer: 25</td>
<td></td>
</tr>
<tr>
<td>normGrey</td>
<td>Texture Inertia</td>
<td>Standard Normal</td>
<td>cancer: 27</td>
<td>69.23%</td>
</tr>
<tr>
<td></td>
<td>Texture Correlation</td>
<td></td>
<td>non-cancer: 25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Texture InvDiff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Energy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normRGB</td>
<td>Texture InvDiff</td>
<td>Soft-max, r = 1</td>
<td>cancer: 27</td>
<td>73.08%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>non-cancer: 25</td>
<td></td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>Texture Entropy</td>
<td>Soft-max, r = 1</td>
<td>cancer: 27</td>
<td>67.31%</td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td>non-cancer: 25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Energy</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2.2.2 Partek Experiments

As mentioned in the Partek experiments of the section 1, the features sets corresponding to the best results of CVIP-FEPC are considered for the classification in the Partek application. Also, the same six classification methods of partek mentioned before are used in the section too; *linear discriminant analysis with equal prior probability* (LDE), *linear discriminant analysis with proportional prior probability* (LDP), *nearest centroid with equal prior probability* (NCE), and also *K-nearest neighbor* (KNN) with K=1, 3, and 5. There are four tables of the best classification results of Partek experiments done on four different limb regions. In the all the tables, only the classification method of the best success rate is mentioned. In Table 5.16, Table 5.17, Table 5.18, and Table 5.19, the best results of the
Partek experiments are illustrated for full-limb, hip/shoulder, knee/elbow, and wrist limb regions, respectively.

Table 5.16: Best classification success rates by using six classification methods of Partek- full-limb

<table>
<thead>
<tr>
<th>Color Norm* Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Norm* Method</th>
<th>Classification Method</th>
<th>Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
</table>
| ______ (original image) | Histogram Skew
Histogram Entropy | Standard Normal | KNN K = 1 | 65.38% | 62.00% | 69.00% |
| lum | Histogram Skew | Soft-max, r = 1 | KNN K = 5 | 65.38% | 54.00% | 77.00% |
| normGrey | Histogram StdDev | Standard Normal | KNN K = 1 | 65.38% | 62.00% | 69.00% |
| normRGB | Spectral Texture Energy
Texture Inertia
Histogram Mean
Histogram Skew
Histogram Energy | Soft-max, r = 1 | LDE/LDP | 57.69% | 54.00% | 62.00% |
| normRGB -lum | Texture InvDiff
Histogram Skew
Histogram Entropy | Standard Normal | KNN K = 1 | 80.77% | 100.00% | 62.00% |

As it is shown in Table 5.16, the best result of the full-limb region, 80.77%, is related to the normRGB-lum experiment which is classified by K-nearest neighbor with K = 1. In addition, in the most of experiments of full-limb region, K-nearest neighbor with K = 1 results the highest classification success rate.
Table 5.17: Best classification success rates by using six classification methods of Partek- hip/shoulder

<table>
<thead>
<tr>
<th>Color Norm* Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Norm* Method</th>
<th>Classification Method</th>
<th>Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>_______ (original image)</td>
<td>Texture Inertia Texture InvDiff Texture Entropy Histogram Skew</td>
<td>Standard Normal</td>
<td>KNN K = 1</td>
<td>62.86%</td>
<td>61.00%</td>
<td>65.00%</td>
</tr>
<tr>
<td>lum</td>
<td>Histogram StdDev</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 3</td>
<td>71.43%</td>
<td>72.00%</td>
<td>71.00%</td>
</tr>
<tr>
<td>normGrey</td>
<td>Histogram Skew</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 5</td>
<td>65.38%</td>
<td>54.00%</td>
<td>77.00%</td>
</tr>
<tr>
<td>normRGB</td>
<td>Texture Correlation Histogram Mean</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 1</td>
<td>60.00%</td>
<td>72.00%</td>
<td>47.00%</td>
</tr>
<tr>
<td>normRGB -lum</td>
<td>Texture Correlation Histogram Mean</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 1</td>
<td>60.00%</td>
<td>72.00%</td>
<td>47.00%</td>
</tr>
</tbody>
</table>

The best result of the hip/shoulder region experiments is obtained from the lum color normalized images classified by K-nearest neighbor with K = 3. Moreover, K-nearest neighbor with K = 1 is the common classification method produces the best success rates in the experiments shown in Table 5.17.
Table 5.18: Best classification success rates by using six classification methods of Partek - knee/elbow

<table>
<thead>
<tr>
<th>Color Norm* Method</th>
<th>Feature Sets (texture pixel dist=6)</th>
<th>Data Norm* Method</th>
<th>Classification Method</th>
<th>Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>Histogram StdDev</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 1</td>
<td>58.33%</td>
<td>41.00%</td>
<td>74.00%</td>
</tr>
<tr>
<td></td>
<td>Histogram Entropy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lum</td>
<td>Histogram Mean</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 3</td>
<td>63.09%</td>
<td>68.00%</td>
<td>58.00%</td>
</tr>
<tr>
<td></td>
<td>Histogram StdDev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normGrey</td>
<td>Spectral Texture Correlation</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 3</td>
<td>60.71%</td>
<td>68.00%</td>
<td>53.00%</td>
</tr>
<tr>
<td></td>
<td>Texture Entropy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Texture Entropy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histogram Energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normRGB</td>
<td>Histogram Skew</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 5</td>
<td>58.33%</td>
<td>54.00%</td>
<td>63.00%</td>
</tr>
<tr>
<td>normRGB -lum</td>
<td>Texture InvDiff</td>
<td>Soft-max, r = 1</td>
<td>KNN K = 5</td>
<td>65.48%</td>
<td>66.00%</td>
<td>65.00%</td>
</tr>
<tr>
<td></td>
<td>Histogram Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The best Partek results of the classification of the knee/elbow region images are displayed in the Table 5.18, in which normRGB-lum experiment produces, 65.48%, the best classification success rate among all of the region experiments done by Partek.
As shown in the Table 5.19, the highest result of wrist region experiments belongs to the normGrey color normalized images which are classified by the K-nearest neighbor with K value of ‘1’.

### 5.2.3 Summary of the Results

Totally, 175 different sets of experiments are implemented by the CVIP-FEPC and Partek applications, from which 25 sets of experiments are executed in the CVIP-FEPC application. The experiments are created by different combinations of the limb regions (five limb regions) and color normalization statues (four color normalizations and one original). Among all the CVIP-FEPC experiments, the full-limb region with normRGB-lum color
normalized images produces the best classification success rate of 76.92% in the bone cancer diagnosis using thermographic images.

A total number of 150 sets of experiments are executed in the Partek application, which is including all combinations of five limb regions, five color normalization statuses, and six classification methods. However, in the results chapter, only the best results of the six classification methods were discussed which are only 25 sets of experiments. In the experiments, the best correct classification rate, 80.77%, is obtained by the full-limb region images which are color normalized by normRGB-lum and classified by K-nearest neighbor with K = 1.

The overall results of the CVIP-FEPC and Partek applications are depicted and compared in the following graphs. Figure 5.6 is to show the best results obtained by CVIP-FEPC, in which K-nearest neighbor with K = 4 is used as classification method.
Figure 5.6: CVIP-FEPC best classification success rates for the bone cancer detection with different limb regions images. Note: the best results are obtained by either original or color normalized images.

The best correct classification rates obtained by Partek are illustrated in Figure 5.7, in which different classification methods are applied including K-nearest neighbor (KNN) with K = 1, 3, and 5. However, in the most of cases, the classification method of KNN with K=1 results the best classification success rates.
Figure 5.7: Partek best classification success rates for the bone cancer detection with different limb regions images. Note: the best results are obtained by either original or color normalized images.

In order to evaluate the performance of each color normalization method, the average of the best results obtained by the CVIP_FEPC and the Partek software is calculated based on each color normalization method. The evaluation of each color normalization method is illustrated in the Figure 5.8.
According to the information reported in the Figure 5.8, normRGB-lum and lum color normalization methods perform better than other color normalization methods. The average of the best results obtained by lum color normalized images is 67.02%, and normRGB-lum color normalized images produce the average success rate of 67.20%.

The frequency of each feature occurrence in the best results is shown in the Figure 5.9. The frequency is measured from the feature sets of the best classification sets obtained by K-nearest neighbor with K = 4 (see CVIP-FEPC experiments Sections 5.2.1 and 5.2.2). The features frequency measure is calculated by the number of experiments each is appeared divided by the total number of experiments with the best results (25 experiments of CVIP-FEPC).
As shown in the Figure 5.9, the features including histogram skew, texture correlation, and histogram standard deviation are appeared frequently in the feature sets with best classification success rates. The frequency measures of the features are displayed by red color in the Figure 5.9.

5.3 Canine Hair Type Effect in Bone Cancer Diagnosis

This section of the experiments is to investigate whether the hair type of the dogs affects the thermographic images for bone cancer detection. Based on the breeds and hair types of the dogs participated in this research, the thermal images can be divided in two categories of long-hair and short-hair. In this section, the bone cancer detection experiments are implemented on long-hair and short-hair image categories separately. In the first part, the experimental results with the long-hair are shown and discussed, and the short-hair results are investigated in the second part. In each part, the experiments are applied on both original and color normalized images based on different limb regions.
5.3.1 Experiment Results of the Long-Hair Category

The total number of 25 different sets of experiments of long-hair category is implemented in CVIP-FEPC to evaluate the bone cancer detection success rates according to this hair type. The 25 experiments encompass the all combinations of the five limb region based conditions and five color normalization statuses. The five limb region based conditions include all limb regions, full-limb, hip/shoulder, knee/elbow, and wrist regions, and the five color normalization methods consist of none-color normalization (original images), lum, normGrey, normRGb, and normRGB-lum color normalization methods. In each of the experiments using color normalized images, 11 types of features are extracted; five types of histogram features, five types of texture features, and one type of spectral feature (see Section 4.2). However, in the experiments with original images all the mentioned features are extracted except mean histogram feature (10 types of features in total). The testing method is full leave-one-out and the object classification method is $K$-nearest neighbor with $K = 1$, which is called nearest neighbor as well. The reason behind using the classification method is the best correct classification rates obtained in the bone cancer detection experiments without hair type consideration (Partek results), in the previous sections of the chapter.

The results are shown in five tables corresponding to five limb region based conditions. In each table the results of the all five color normalization statuses are illustrated and compared as well. The Table 5.20 shows the experiments results including all limb regions, original and color normalized images. In the experiments, 71 thermal images are used for the classification, from which 35 images belong to the cancer class and 36 images belong to the non-cancer class. Based on the results shown in Table 5.20, the original and normGrey images produce the highest classification success rate, 64.79%, for the long-hair category with all limb regions images.
Table 5.20: Classification results of long-hair category for original and color normalized images of all limb regions

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>_______ (original image)</td>
<td>cancer: 35 non-cancer: 36</td>
<td>all limb regions</td>
<td>64.79%</td>
<td>65.71%</td>
<td>63.89%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 35 non-cancer: 36</td>
<td>all limb regions</td>
<td>61.97%</td>
<td>62.86%</td>
<td>61.11%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 35 non-cancer: 36</td>
<td>all limb regions</td>
<td>64.79%</td>
<td>62.86%</td>
<td>66.67%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 35 non-cancer: 36</td>
<td>all limb regions</td>
<td>61.97%</td>
<td>60.00%</td>
<td>63.89%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 35 non-cancer: 36</td>
<td>all limb regions</td>
<td>61.97%</td>
<td>65.71%</td>
<td>58.33%</td>
</tr>
</tbody>
</table>

The classification success rates of full-limb region images are illustrated in Table 5.21. In the experiments, five cancer and five non-cancer images are used for the investigation of long-hair category. Two color normalization methods of normGrey and normRGB-lum as well as original images result in the best classification rates of 80.00% for the specific limb region.
Table 5.21: Classification results of long-hair category for original and color normalized images of full-limb region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>_______ (original image)</td>
<td>cancer: 5 non-cancer: 5</td>
<td>full-limb</td>
<td>80.00%</td>
<td>80.00%</td>
<td>80.00%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 5 non-cancer: 5</td>
<td>full-limb</td>
<td>70.00%</td>
<td>80.00%</td>
<td>60.00%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 5 non-cancer: 5</td>
<td>full-limb</td>
<td>80.00%</td>
<td>60.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 5 non-cancer: 5</td>
<td>full-limb</td>
<td>70.00%</td>
<td>40.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 5 non-cancer: 5</td>
<td>full-limb</td>
<td>80.00%</td>
<td>80.00%</td>
<td>80.00%</td>
</tr>
</tbody>
</table>

As shown in Table 5.22, the experiments corresponding to the hip/shoulder region are implemented on a total number of 14 thermographic images which contain seven cancer and seven non-cancer images. The best result is 78.57% which is related to the normGrey color normalized cancer and non-cancer images.
Table 5.22: Classification results of long-hair category for original and color normalized images of hip/shoulder region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>________ (original image)</td>
<td>cancer: 7, non-cancer: 7</td>
<td>hip/shoulder</td>
<td>71.43%</td>
<td>71.43%</td>
<td>71.43%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 7, non-cancer: 7</td>
<td>hip/shoulder</td>
<td>71.43%</td>
<td>71.43%</td>
<td>71.43%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 7, non-cancer: 7</td>
<td>hip/shoulder</td>
<td>78.57%</td>
<td>57.14%</td>
<td>100.00%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 7, non-cancer: 7</td>
<td>hip/shoulder</td>
<td>71.43%</td>
<td>71.43%</td>
<td>71.43%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 7, non-cancer: 7</td>
<td>hip/shoulder</td>
<td>71.43%</td>
<td>71.43%</td>
<td>71.43%</td>
</tr>
</tbody>
</table>

All the results of the knee/elbow region experiments which are done on long-hair category are displayed in Table 5.23. The highest correct classification rate, 78.79%, of the limb region is generated by normGrey color normalization method.
Table 5.23: Classification results of long-hair category for original and color normalized images of knee/elbow region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>________ (original image)</td>
<td>cancer: 16 non-cancer: 17</td>
<td>knee/elbow</td>
<td>72.72%</td>
<td>62.50%</td>
<td>82.35%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 16 non-cancer: 17</td>
<td>knee/elbow</td>
<td>69.70%</td>
<td>68.75%</td>
<td>70.59%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 16 non-cancer: 17</td>
<td>knee/elbow</td>
<td>78.79%</td>
<td>75.00%</td>
<td>82.35%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 16 non-cancer: 17</td>
<td>knee/elbow</td>
<td>66.66%</td>
<td>68.75%</td>
<td>64.71%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 16 non-cancer: 17</td>
<td>knee/elbow</td>
<td>66.66%</td>
<td>68.75%</td>
<td>64.71%</td>
</tr>
</tbody>
</table>

The last limb region studied for the long-hair category is the wrist region, and Table 5.24 is to display its classification success rates. The best success rate of the experiments is 92.86% which is even higher than the best results of the other limb regions. The classification rate is generated by color normalized images of normRGB-lum method.
Table 5.24: Classification results of long-hair category for original and color normalized images of wrist region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>________ (original image)</td>
<td>cancer: 7 non-cancer: 7</td>
<td>wrist</td>
<td>78.57%</td>
<td>85.71%</td>
<td>71.43%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 7 non-cancer: 7</td>
<td>wrist</td>
<td>71.43%</td>
<td>71.43%</td>
<td>71.43%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 7 non-cancer: 7</td>
<td>wrist</td>
<td>71.43%</td>
<td>85.71%</td>
<td>57.14%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 7 non-cancer: 7</td>
<td>wrist</td>
<td>85.71%</td>
<td>100.00%</td>
<td>71.43%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 7 non-cancer: 7</td>
<td>wrist</td>
<td>92.86%</td>
<td>85.71%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

5.3.2 Experiment Results of the Short-Hair Category

The same methodology of the long-hair category experiments is applied in the short-hair category too. The same histogram, texture, spectral features and the same classification method of the nearest neighbor are used for the experiments. The only difference between these experiments and the experiments related to the long-hair category is the hair type of the dogs.

Similar to the results of the long-hair category, there are five tables of data for different limb region conditions. For each of the limb regions, four color normalization methods are applied that with no-color normalization consideration, five experiments are implemented.

The Table 5.25 displays the classification success rates with total images of the short-hair category. There are 126 thermal images of all limb regions including 64 cancer and 62
non-cancer thermograms. In the experiments, original and normRGB-lum images generate the best classification success rates of 65.87%.

Table 5.25: Classification results of short-hair category for original and color normalized images of all limb regions

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>cancer: 64 non-cancer: 62</td>
<td>all limb regions</td>
<td>65.87%</td>
<td>68.75%</td>
<td>62.90%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 64 non-cancer: 62</td>
<td>all limb regions</td>
<td>64.28%</td>
<td>64.06%</td>
<td>64.52%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 64 non-cancer: 62</td>
<td>all limb regions</td>
<td>60.32%</td>
<td>57.81%</td>
<td>62.90%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 64 non-cancer: 62</td>
<td>all limb regions</td>
<td>60.32%</td>
<td>54.69%</td>
<td>66.13%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 64 non-cancer: 62</td>
<td>all limb regions</td>
<td>65.87%</td>
<td>68.75%</td>
<td>62.90%</td>
</tr>
</tbody>
</table>

The classification results of the 16 full-limb region images for the short-hair category are shown in the Table 5.26. As the results show, the best classification rates are produced by three color normalization methods including lum, normGrey, and normRGB.
Table 5.26: Classification results of short-hair category for original and color normalized images of full-limb region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>cancer: 8 non-cancer: 8</td>
<td>full-limb</td>
<td>50.00%</td>
<td>62.50%</td>
<td>37.50%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 8 non-cancer: 8</td>
<td>full-limb</td>
<td>81.25%</td>
<td>87.50%</td>
<td>75.00%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 8 non-cancer: 8</td>
<td>full-limb</td>
<td>81.25%</td>
<td>75.00%</td>
<td>87.50%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 8 non-cancer: 8</td>
<td>full-limb</td>
<td>81.25%</td>
<td>87.50%</td>
<td>75.00%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 8 non-cancer: 8</td>
<td>full-limb</td>
<td>68.75%</td>
<td>75.00%</td>
<td>62.50%</td>
</tr>
</tbody>
</table>

As shown in Table 5.27, in the experiments of short-hair category using 21 hip/shoulder images, the best classification result is 85.71% produced by lum color normalization method. The experiments of the limb region, hip/shoulder, produces the highest success rates than all other limb regions in the short-hair category.
Table 5.27: Classification results of short-hair category for original and color normalized images of hip/shoulder region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(original image)</td>
<td>cancer: 11 non-cancer: 10</td>
<td>hip/shoulder</td>
<td>76.19%</td>
<td>72.73%</td>
<td>80.00%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 11 non-cancer: 10</td>
<td>hip/shoulder</td>
<td>85.71%</td>
<td>81.82%</td>
<td>90.00%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 11 non-cancer: 10</td>
<td>hip/shoulder</td>
<td>71.43%</td>
<td>81.82%</td>
<td>60.00%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 11 non-cancer: 10</td>
<td>hip/shoulder</td>
<td>61.90%</td>
<td>54.55%</td>
<td>70.00%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 11 non-cancer: 10</td>
<td>hip/shoulder</td>
<td>80.95%</td>
<td>90.91%</td>
<td>70.00%</td>
</tr>
</tbody>
</table>

In the experiments of short-hair category illustrated in the Table 5.28, the total number of 51 images of the knee/elbow region is applied. According to the results, the highest correct classification rate is 72.55% which is the corresponding result to the lum color normalized images.
The table below shows the classification results for the short-hair category for original and color normalized images of the knee/elbow region.

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>original image</td>
<td>cancer: 25</td>
<td>knee/elbow</td>
<td>66.66%</td>
<td>56.00%</td>
<td>76.92%</td>
</tr>
<tr>
<td></td>
<td>non-cancer: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 25</td>
<td>knee/elbow</td>
<td>72.55%</td>
<td>72.00%</td>
<td>73.08%</td>
</tr>
<tr>
<td></td>
<td>non-cancer: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 25</td>
<td>knee/elbow</td>
<td>68.63%</td>
<td>72.00%</td>
<td>65.38%</td>
</tr>
<tr>
<td></td>
<td>non-cancer: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 25</td>
<td>knee/elbow</td>
<td>60.78%</td>
<td>60.00%</td>
<td>61.54%</td>
</tr>
<tr>
<td></td>
<td>non-cancer: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 25</td>
<td>knee/elbow</td>
<td>64.70%</td>
<td>60.00%</td>
<td>69.23%</td>
</tr>
<tr>
<td></td>
<td>non-cancer: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The wrist region is the last limb region analyzed in the short-hair category. In the hair type category, there are 38 thermograms including 20 cancer and 18 non-cancer images. In the experiments, the color normalized images by normGrey method make the best correct classification rate of 81.58%.
Table 5.29: Classification results of short-hair category for original and color normalized images of wrist region

<table>
<thead>
<tr>
<th>Color Normalization Method</th>
<th>Number of Images per Class</th>
<th>Experiment Group</th>
<th>Classification Success Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>__________</td>
<td>cancer: 20 non-cancer: 18</td>
<td>wrist</td>
<td>71.05%</td>
<td>80.00%</td>
<td>61.11%</td>
</tr>
<tr>
<td>lum</td>
<td>cancer: 20 non-cancer: 18</td>
<td>wrist</td>
<td>76.31%</td>
<td>75.00%</td>
<td>77.78%</td>
</tr>
<tr>
<td>normGrey</td>
<td>cancer: 20 non-cancer: 18</td>
<td>wrist</td>
<td>81.58%</td>
<td>80.00%</td>
<td>83.33%</td>
</tr>
<tr>
<td>normRGB</td>
<td>cancer: 20 non-cancer: 18</td>
<td>wrist</td>
<td>73.68%</td>
<td>65.00%</td>
<td>83.33%</td>
</tr>
<tr>
<td>normRGB-lum</td>
<td>cancer: 20 non-cancer: 18</td>
<td>wrist</td>
<td>73.68%</td>
<td>80.00%</td>
<td>66.67%</td>
</tr>
</tbody>
</table>

5.3.3 Summary of the Results

To determine the effect of the dog’s hair type in the thermal images for bone cancer detection, a total number of 50 sets of experiments are implemented by the CVIP-FEPC. Among the experiments, half of them are applied on the long-hair category images and the other half on the short-hair images. In each of the long-hair and short hair category experiments, all the combinations of the five different limb regions and five color normalization methods are examined by using nearest neighbor classification method.

The best results of the long-hair category (Table 5.20 through Table 5.24) are illustrated in Figure 5.10, in which the classification success rate, 92.86%, of the wrist region is higher than the other regions results.
In the Figure 5.11, the highest classification rates of the short-hair category experiments (Table 5.25 through Table 5.29) are shown with their corresponding limb regions. The best results of the short-hair category do not change dramatically when the limb region changes. However among them, hip/shoulder images produce the best classification success rate of 85.71%.
Figure 5.11: Best classification rates for the bone cancer detection with different limb regions images of short-hair category. Note: the best results are obtained by either original or color normalized images.

In order to evaluate the overall results obtained by long-hair and short hair categories, the averages of the best results (Figure 5.10 and Figure 5.11) and the average of the total results of the 25 experiments for each hair type (Table 5.20 through Table 5.29) are calculated and compared in the Figure 5.12. According to the evaluation, although there is an expectation to obtain better results for short-hair category by veterinarians, the overall results of the long-hair category are higher.
Figure 5.12: The average of the best classification rates and also the average of the all classification rates for both long-hair and short-hair categories.
CHAPTER 6

CONCLUSION

The research study is to determine if it is possible to diagnose canine bone cancer by thermographic image analysis. Based on the conditions under which the study was performed, the answer is by the average success rate of 70.83%, it is possible to detect canine bone cancer with the thermograms.

Among 175 experiments with different limb regions, color normalization and pattern classification methods, the color normalized images of normRGB-lum of the full-limb region produced the best success rate of 80.77%. The number of thermal images in the specific experiment is 26 including 13 cancer and 13 non-cancer thermograms. Three features are extracted in the experiment which are texture inverse difference, histogram skew, and histogram entropy with using standard normal density data normalization method. The best success rate is obtained by K-nearest neighbor with K = 1.

Although the best classification success rate is 80.77%, the average of the best classification rates of all the different limb region categories is 70.83% calculated by data information of Figure 5.6 and Figure 5.7. This success rate is predicted to be improved by using more homogeneous images taken from a specific dog breed and gender.

In this study, also we tried to determine whether the hair type of the dog can influence on the bone cancer thermographic images and also in the disease diagnosis. According to the 50 experiments including long-hair and short-hair categories (Figure 5.12), the long-hair category performance in bone cancer detection is better than the short-hair category performance by a maximum of 2.00%, which is not determined to be significant.
All the canine bone cancer detection experiments are performed by using masks created manually, however the automatic mask creation algorithm is developed to facilitate the mask creation stage for finding the regions of interest in the future studies. Based on the XOR error metric with threshold value of 0.8, 47 out of 99 automatic masks are very similar to their related manual masks. Also, according to the subtraction-energy error metric with threshold value of 0.85, 36 of 99 created automatic masks extract similar regions of interest (ROIs) to their corresponding ROIs extracted by the manual masks. However, among the satisfying automatic masks, there are masks which include more than one object. The kind of masks cannot be applied in the CVIP-FEPC application for the region of interest extraction, so a solution is needed to change the multi-objects masks to the one-object ones. This solution is defined in the future scope chapter.
CHAPTER 7

FUTURE SCOPE

In this research study, the thermal images are taken from 22 cancer dogs whose diseases were confirmed by the biopsy results. If the number of thermal images used in the experiments is increased, the results and conclusion based on them should be more reliable. Recently, we have received eight more dogs and their corresponding cancer and non-cancer images and data sheets which can be added to the experiments and perform the investigation with total number of 30 dogs.

The feature sets used for the Partek experiments are selected based on the best classification results of K-nearest neighbor with $K = 4$ by CVIP-FEPC. As the results of the nearest neighbor classification method applied by Partek were higher than the same experiments (same feature sets) with K-nearest neighbor, $K = 4$, there is the possibility to improve the classification rates by implementing nearest neighbor by CVIP-FEPC at first, and using their best feature sets for the Partek experiments. The new feature sets selected by the best results of the nearest neighbor may increase the classification rates of the Partek experiments using other classification methods.

In addition, according to the problem faced in the multi-objects results of the automatic mask creation algorithm, it is possible to detect the biggest object among the all objects in the multi-objects masks and keep only the main one. It can be performed by programming and using labeling function. The area of each label (object) can be found and only the biggest object can be kept.
REFERENCES


http://www.vetinfo.com/dog-bone-cancer.html#b

National Institutes of Health (2008), “Bone Cancer Fact Sheet.”

http://www.cancer.gov/cancertopics/factsheet/Sites-Types/bone


http://www.meditherm.com/mms_default.htm

CVIPtools (2010), “Computer Vision and Image processing Tools”.

http://cviptools.ece.siue.edu

CVIP-ATAT (2010), “CVIP- Algorithm Test and Analysis Tool”.

http://cviptools.ece.siue.edu

CVIP-FEPC (2010), “CVIP- Feature Extraction and Pattern Classification”.

http://cviptools.ece.siue.edu

Partek (2012), “Partek Discovery Suite”.

http://www.partek.com/partekds


