Animal Classification Using Appearance-Based, Model-Based and Texture-Based Methods

Shaafan Jameel Othman

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Engineering

Eastern Mediterranean University
November 2018
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Assoc. Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Computer Engineering.

Prof. Dr. H. Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Computer Engineering.

Assoc. Prof. Dr. Önsen Toygar
Supervisor

Examining Committee

1. Assoc. Prof. Dr. Duygu Çelik Ertuğrul
2. Assoc. Prof. Dr. Önsen Toygar
3. Asst. Prof. Dr. Mehtap Köse Ulukök

ABSTRACT

Animal recognition and their classification have become increasingly popular area in pattern recognition and computer vision. Distinguishing between images of different animals and different species as means of classifying animals is an easy task for humans. However, it is difficult to distinguish animal species automatically even in simple cases such as distinguishing cats and dogs. Animal bodies easily get disfigured, they can appear in images in a way that they self-occlude and often the background in such images could be very complex and noisy. Furthermore, just as all objects in an image, illumination may vary from image to image, the dimension and viewpoints may also differ. There has been attempts to carry out animal recognition from images but this problem has gained not enough attention.

In the literature, there are some visual animal biometrics algorithms categorizing specific animal species such as zebra, elephant, chimpanzees, tiger, whales, pet animals like dogs, etc. Moreover, there are a few publicly available animal face databases including the facial images of different animals such as LHI Animal Faces, HiT and KTH Databases. In this thesis, we studied on different animal face images to classify different animal species. We implemented appearance-based, texture-based and model-based feature extraction methods to categorize animals from their faces and a comparative study is performed at the end of the thesis.

Keywords: Animal classification, visual animal biometrics, appearance-based methods, model-based methods, texture-based methods.

ÖZ

Hayvan türlerinin sınıflandırılması ve hayvan tanıma, bilgisayarla görü ve örüntü tanıma alanlarında önemli bir uygulamadır. Hayvan sınıflandırma, değişik hayvan türü görüntülerinin ayırt edilmesi problemini çözmek için kullanılır. Bu problem insanlar için basit bir iş olsa da, bilgisayarla otomatik olarak, kedi ve köpekleri bile ayırt etmek zordur. Hayvanların vücutları deforme olup kendi kendini kapatma veya saklama özelliğine sahiptir ve bu işlem karmaşık arkaplan görüntüleriyle daha da belirgin hale gelir. Bunlara ek olarak, bütün nesneler gibi hayvanlar da farklı ışıklandırılmış, farklı görüş açısıyla ve farklı büyüklükte görünebilir. Hayvan görüntüleri üzerinde tanıma yöntemlerini uygulamak için girişimler yapılmış, ancak hayvanların kategorize edilmesi problemi üzerinde çok yaygın çalışmalar yapılmamıştır.

Literatürde bazı görsel biyometri algoritmaları kullanılıp, zebra, fil, şempanze, kaplan, balina, köpek gibi bazı hayvan türlerinin sınıflandırılması yapılmıştır. Ayrıca, LHI Hayvan Yüzleri, HiT ve KTH veritabanları gibi sınırlı sayıda halka açık hayvan görüntüsü veritabanı bulunmaktadır. Bu tezde, değişik hayvan türlerinin sınıflandırılması için hayvan yüzü görüntüleri üzerinde çalışılmıştır. Görüntüye-dayalı, dokuya-dayalı ve modele-dayalı öznitelik çıkarma yöntemleriyle hayvan yüzlerinin sınıflandırılması yapılmıştır. Bu yöntemlerin hayvan sınıflandırma problemi üzerindeki başarımları tezin sonunda karşılaştırılmıştır.

Anahtar kelimeler: Hayvan sınıflandırma, görsel hayvan biyometrisi, görüntüye-dayalı yöntemler, modele-dayalı yöntemler, dokuya-dayalı yöntemler.

I dedicate this thesis to my lovely family, especially to my Mom and Dad. I would like to thank them from all of my heart for their support through these years to finish my Master degree.

ACKNOWLEDGMENT

First, thank God for everything. I sincerely appreciate the role of Assoc. Prof. Dr. Önsen Toygar who was my supervisor for this work. Her encouragement and patience guided me during the thesis study. This work would never have been successful without her help. It is my honor to work with her. Additionally, I want to also express my appreciation to all academic and administrative staff of Computer Engineering Department.

To my lovely family, especially my parents who gave me more confidence and power to persevere always and never quit, and for massively assisting me throughout the course of this study, I thank you all.

Finally, I would like to thank to all of my friends, thanks to every person who supported me during my master degree study. Moreover, I thank to the EMU University.

God bless you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Classification of images involves the identification of objects available in an image depending on their appearance, textures and properties associated with the image as well as the process of labeling them according to these attributes. Detection of objects in an image is made up of two categories of processes in general; object instance recognition, and object class recognition. The former is a supervised classification process that aims to identify originally classified groups or images, while the later one is a category level process which involves the process of introducing and categorizing new and unknown objects according to predefined categories. The object class recognition aspect is more laborious to implement because of the possibility of having different variants in the same group arising from the differences in image capturing condition, texture of image, color and other possible factors. The preceding step in the classification of an image or an object is feature/descriptor extraction [1].

1.1 Animal Classification System

Animal classification system uses an image classification algorithm that is image classifier, which receives the photograph or a part of the photograph as query and returns what the image contains. This implies that, following a classification process, an output is given showing the class label (tiger, lion, duck etc.). Figure 1.1 shows the animal classification system which is divided into the following steps:

- Preprocessing: the input images undergo a series of preprocessing techniques in order to reduce the effect of factors capable of influencing the animal classification adversely.
- Feature extraction: the features are extracted from an image using several methods such as Principal Component Analysis (PCA) as an appearance-based method, Local Binary Patterns (LBP) as a texture-based method and Elastic Bunch Graph Matching (EBGM) as a model-based method.
- Matching: Comparing the feature vector of each test image with the feature vectors of all training images by using distance measures or similarity measures.
- Classification: classify the test image so that the returned matched train image is the closest one to the input test image.

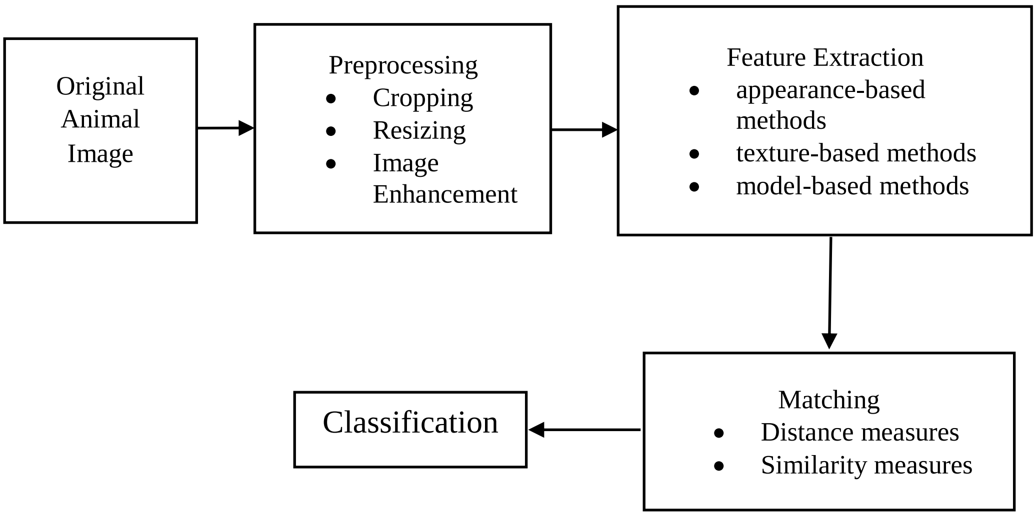

Figure 1.1: System for Animal Recognition and Classification.

It is noteworthy that many traditional classification algorithms in computer vision follow this pipeline described in Figure 1.1 [2]. In this thesis, the feature extraction methods namely PCA, LBP and EBGM methods and the distance measurement technique Manhattan Distance are used to classify the test images on four datasets.

1.2 The Work Done

Classification of animals face by the analysis using texture, appearance and model are applied on face images by using three different algorithm, namely Principal Component Analysis (PCA), Local Binary Patterns (LBP) and Elastic Bunch Graph Matching (EBGM). Each of the algorithms are applied to the whole face images as feature extractors and combined the techniques with preprocessing algorithms of Histogram Equalization(HE) as well as Mean-Variance Normalization(MVN) to frustrate the illumination influence.

Successful classification by these methods are mainly because of their statistics-based ability of extracting the features automatically and the subsequent classification into specified classes. These algorithms are commonly researched for the classification and recognition problems. These techniques are not without flaws as there are many challenges like variation in illumination, posture, facial colors, and shape. The methods are implemented using the animals' face images from four databases, namely LHI Animal Face Database(LHI)[3], Hybrid image Templates (HiT)[4] and KTH-animals (KTH)[5]. Experiments are carried out to test animal classification performance using PCA, LBP and EBGM methods. The results are shown in the further chapters. The accuracy of animal classification algorithms are also compared with each other at the end of the thesis.

1.3 Thesis Layout

The remaining of the thesis is organized as follows; in Chapter 2, existing methods and animal databases in the literature are presented and analyzed. Chapter 3 explains preprocessing used for facial image classification. In Chapter 4, feature extraction methods are discussed. Chapter 5 presents the experiments and results for image classification on animals. Finally, Chapter 6 presents the conclusions.

Chapter 2

LITERATURE REVIEW

Animal classification is an emerging research area in computer vision and pattern recognition. It is a promised research area that promote new innovation in the area of quantified algorithms and methodologies for representing individuals and recognition of animal characteristics and detection of visible features. At present, real world applications of animal classification systems have become more popular due to the existence of many kinds of applications [6].

Among the early system in the field of animal classification, Mizroch and Harkness from the National Marine Mammal Laboratory (NMML) enbacked on a process of compiling images of whale tail flukes that were captured in North Pacific waters [7]. NMML system is developed based on the concept of human brain that has the ability to spontaneously adjust, rotate, compensate, and recognize identical images. Testing is carried out on automated system for matching humpback whale tail fluke in 12,000 images. No significant difference was noticed when accuracy of matching with automated system was compared with matching a smaller catalog of about 200 to 400 pictures.

Ramanan et.al [8] described and built a model of animals from tracking videos. It is a completely automatic system where articulated 2D models known as pictorial structures are built from moving pictures of the animals. This developed tracking system is a generalized one with the ability to track and model objects at the same time. The learned model is compared to a visual library, moreover the system can be seen as a video recognition method. The system can be viewed as a learning model method for object recognition as well. The system is able to improve the pictorial structures significantly by employing distinguishing texture model learned from a texture library to enhance them. The authors developed a novel texture descriptor which performed better than other state-of-the-art descriptors for animal textures. They demonstrated the system using real video series of three different animals (tiger, giraffe, and zebra). The learned models are used to identify animals from two datasets which include high quality images from the web from Google and images captured from the Corel collection by professional photographers. The performance was considerably good on both datasets. Comparing results with simple baselines, the authors showed that the Google set is able to detect, localize, and recover defected images from a stockpile significantly impossible for object recognition. A novel simple representation of texture is one contribution of the work, instead of using a histogram of textons that represents texture as a patch of pixels. They demonstrated that this representation performed better than the contemporary for the problem of animal detection.

Domeier and Nasby-Lucas [9] proposed a reliable automated method for white sharks identification using underwater images. The system was developed and adapted for examining site integrity at Guadalupe Island, Mexico. The most consistent features for identification in a number of years were the gill flaps, caudal fins and pelvic fins pigment patterns. Pigment patterns in each of the three regions were asymmetrical on the right and left sides making it important to capture both sides in order to record all individuals. Once recorded however, re-identification of an individual could done using a partial body image. Employing this system, 73 individuals were identified between 2001 and 2005. Integrity of the site was shown by repeating yearly sighting of individuals observing and identifying 78 percent of sharks in about 2 years. Identifying sharks individually through a systematic method gave an important baseline data for animal identification. They postulated that accurate estimation and identification of white shark population in Guadalupe Island is possible over a period of time using mark and recapture statistics, which enables us to track the population trends and new recruitment to an aggregation site. Moreover, documentation of migration of individuals from one aggregation site to another in a cheap way instead of tagging methods which are expensive. Trends in population could also be compared between sites.

An identification system for elephants was developed in 2007 [10] that was based on characterization of elephants' ears by shape comparison. The proposed system is popular among zoologists for pictorial identification of elephants in the wild because the method has the ability to work well in noisy and clustered images. Segmentation issues are dealt with by employing rough position information provided by the user. This input is employed in the system as a start point for segmentation and normalization hypotheses which aim at comparing an input image *A* with all images available in the database of the system and matching A with an image belonging to *A*. This implies that the system starts from positions given by the user, and formulates a set of normalization hypotheses with the aim of aligning the shape of the input with the curves saved in the system's database. The method is based on matching multiple curves which are not connected and it can be applied to different retrieval problems that are based on shape comparison. The method does not depend on unique guess on the elephant sample space and the images generally have low resolution and are highly cluttered. The accuracy rate is better in this system than identical single curve matching based animal recognition methods that are tested with low to medium quality images.

Mayank et.al proposed a computer-facilitated experimental approach in 2011 for animal identification especially in the field of ecology [11]. The system built a biometrics database of respective animals discriminated by markings on their coats using a database of noisy images captured in the wild. An unknown image of an unknown animal will then be introduced to the system to compare the markings on its coat to those in the database to check if the same animal is known to the system or not. The technique, known as StripeCodes, is able to identify and extract image features with high efficiency and compare images using a powerful programming algorithm. The accuracy of the system was tested against two different methods namely matching multiscale histograms of differential image features, which is a signal processing approach, and Eigenfaces [12], which is an appearance-based algorithm. StripeCodes system outperformed all methods compared for their dataset, and scales well with size of database.

A rapid system for species classification as well as for face detection of African Great Apes was proposed by Ernst and Küblbeck [13]. It was developed from human face detection algorithm with the aim of identifying and protecting endangered species populations. The authors demonstrated the performance of the system using chimpanzee and gorilla faces as examples. Moreover, they describe two different methods for species classification and compare the ability of the system to differentiate between both species. The system took advantage of the power of combining different feature types with Real-AdaBoost that is illumination invariant.

It uses fast pixel-based features for searching possible features and expressing more region-based features in the verification process. The authors composed two test sets which had gorillas and chimpanzees with varying facial dimension and qualities of image for evaluation. They also checked the ability to differentiate between chimpanzees and gorillas based on the scores detection and what is returned by separate classifier.

On the other hand, Zhang et.al[14] having been encouraged by the achievement of face filtering method used in both MSN live image search and Google image search, applied an animal head detection system. This system has numerous applications in real-world online image processing. The initial experiment was carried on a set of relatively large land animals that are commonly found on the internet, such as panda, cat, cheetah, tiger and fox. First of all, the authors proposed a new set of Haar of Oriented Gradients (HOOG) to which are able to detect shape and texture features on animal head effectively. They also introduced two methods for feature detection; Brute-force detection and Deformable detection. Their experiment was carried out on 14379 well labeled images of animals, with results that confirms the dominance of the proposed approach.

In 2012, Loos and Pfitzer proposed and evaluated a face recognition method for identifying great apes individually [15]. The authors presented an enhanced version of the Randomfaces approach, referred to as Hybridfaces, that augments the global recognition output with information retrieved from local facial regions. They evaluated the method on three open source databases of chimpanzees in captive, chimpanzees in the wild and wild western lowland gorillas. The proposed framework indicates promising results and the Hybridfaces approach clearly performs better than the previously used basic Randomfaces method on all three datasets. The proposed approach was evaluated in combination with two alignment methods on three publicly available datasets of chimpanzee and gorilla individuals. The presented Hybridfaces approach outperformed the standard Randomfaces approach on all three datasets if the facial images are sufficiently aligned. Even though the identification results, when congealing is used as alignment technique, are not as high as when a projective transform is applied. The experiments indicate that congealing is a powerful unsupervised alignment algorithm. Therefore, it can be used as a preprocessing step for automatic identification of great apes in the wild as well as in controlled environments.

Kumar and Singh proposed a biometric recognition system for pet animals [16]. They were inspired by rising cases of missing pets, swapping and reallocation of pet animals like dogs and non-genuine insurance claims. Dog's identification needs innovative research to protect the pet animal. Kumar and Singh tried to demonstrate that facial recognition system can be used to recognize the dog effectively to solve the problem for identification of pet animal (such as cat or dog). They proposed a fusion-based method for the recognition of pet animals with 94.86% accuracy. The contributions of this research are that it is the first research to employ existing biometrics algorithm in order to reduce the effects of covariates in dogs.

Finn et.al [17] studied the characterization of individual animal life history that is crucial for efforts to preserve wildlife. Sloop is an operational pattern retrieval system for identification of animals which was broadened by attaching crowdsourcing with image retrieval. The integrated system delivers scalable performance by using cumulative computational deduction to give precision effectively and also improve recall efficiently using feedback from humans. This is the first coupled human-machine animal biometrics system and its performance on various species indicates that the system is effective for large-scale usage. The MIT Sloop is an identification framework for identifying animals individually, where, using aggregated systems, the combined human skills with computer resources resulting in recognition systems with very high performance for conservation efforts on a large-scale.

Table 2.1 demonstrates a summary of the literature review on animal classification. For different animal species, the table presents various research studies with the details of the systems and the performance obtained in different metrics.

Table 2.1: Summary of Literature Review on Animal Classification

Species	Number of Images in database	Modalities	Algorithms	Features	Performance measure	Reference
Humpback Whale	24,428 images	marks or scars	pattern matching	black and white pigment	56%	[7,9]
Zebra	1722 images	texture patches	texton	SIFT	recognition rates (82% for all), 89.30% for zebras, 55.6% for tigers and 59.7% for giraffe	[8]
Elephant	332 images ×268 subjects	ear nicks	shape features	shape and global matching	92% (recognition rate)	[10]
Zebra	120 images	coat pattern	stripe-codes, CO-1+	joint strips of coat pattern	85%	[11]
Ape + Chimpanzee+ Gorilla	15,000 images	face	local features	LBP	92.3%	[12]
Tiger + Panda	14,379 images	head face	HOOG	texture features + dynamic programming	92%	[13]
Chimpanzees	6526 images	face	computer vision	pixel intensity	26.36% FRR (for Chimp Tai) and 21.29% FRR (Chimp Zoo)	[14]
pet animal (Dog)	3000 facial images	face	facial features	PCA+LDA+ICA + texture features	96%	[15]
Otago Skink	6000 images	visual	geometry feature	LBP+SIFT+PCA	92%	[16]

Chapter 3

PREPROCESSING

Images of animals captured by cameras are usually differ from one another due to differences in background, posture and orientation of the head, animals' color, the intensity levels and dimension of the images and other factors. However, majority of the available algorithms cannot automatically solve all these problems, rather, these deficiencies must be solved before introducing such images to the system as query. A good number of these deficiencies are dealt with during classification process. The steps of the whole animal classification system with the details of the preprocessing stages are shown in Figure 3.1.The next subsections give information about each step of preprocessing.

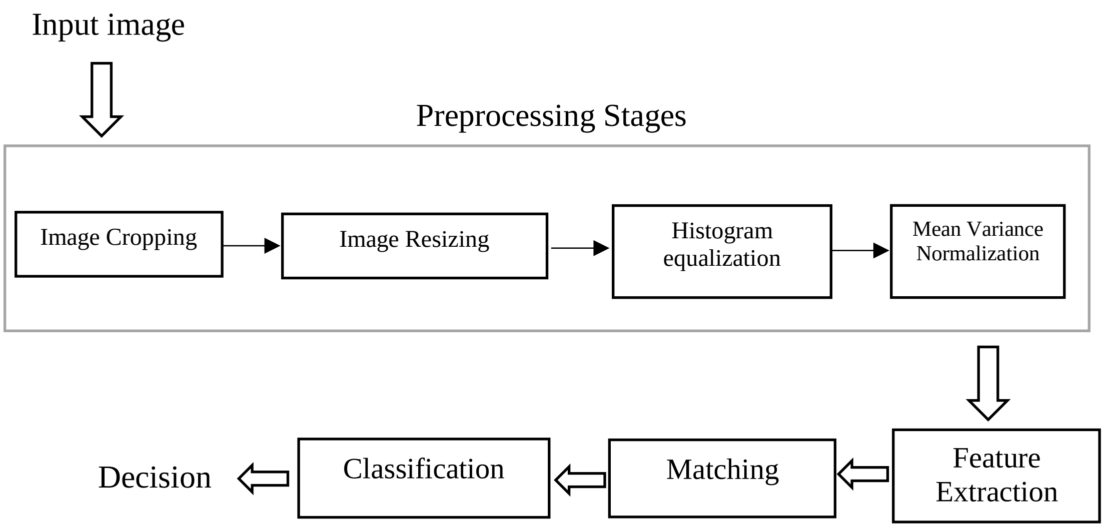

Figure 3.1: The Whole Animal Classification System with Preprocessing Stages.

3.1 Image Cropping

Image background and some additional parts of the face such as hairs and neck are factors that cause to failure in the classification of animals, therefore, effort should be made to remove them from the image. Figure 3.2 illustrates the difference between the original image and a cropped image. Cropping the images also helps to reduce memory consumption and raises the speed of detecting face. The background information that is removed from images are usually not useful.

(a) (b)

Figure 3.2: (a) Original Image (b) Cropped Image from LHI Database.

3.2 Image Resizing

Images that are introduced to the algorithm as an input are mostly in varying dimensions. Therefore, in this step, it is compulsory to resample images so that all input images are of the same size. Matlab Image Processing Toolbox is used in all experimental implementations. Interpolation techniques are used in this work in order to resize the images. Interpolation refers to the process of resizing images through the use of known data to estimate the values at an unknown location. There are mainly three types of interpolation techniques namely Nearest Neighbor Interpolation, Bilinear Interpolation and Bicubic Interpolation. Nearest Neighbor

Interpolation is used in this thesis because it is the simplest one that requires the smallest processing time among all the interpolation methods. It allocates to each new location the intensity of its nearest neighbor in the original image and it performs this process in a very short time.

3.3 Histogram Equalization (HE)

Histogram Equalization **(**HE) in image processing uses the relative frequency of instances of different gray levels in the photographs. Images may have a different amount of intensity levels. Crowding of intensities at different levels might be different and these variations among images reduces the performance of the animal grouping system. The following mathematical formula is used to obtain Histogram equalization:

$$S_k = (L - 1) \sum_{j=0}^{k} P_{r(r_j)} \tag{3.1}$$

where L is the number of all possible intensity levels where K= 0, 1, 2… L-1. P_r is the projection of the rate of occurrence of instances of intensity levels in an image. HE is a function of rounded S_k and n_k where n_k is the count of pixels that possess intensity value, and r_j is the level of intensity of query images. Figure 3.3 shows the difference between levels of intensity of dark and bright images.

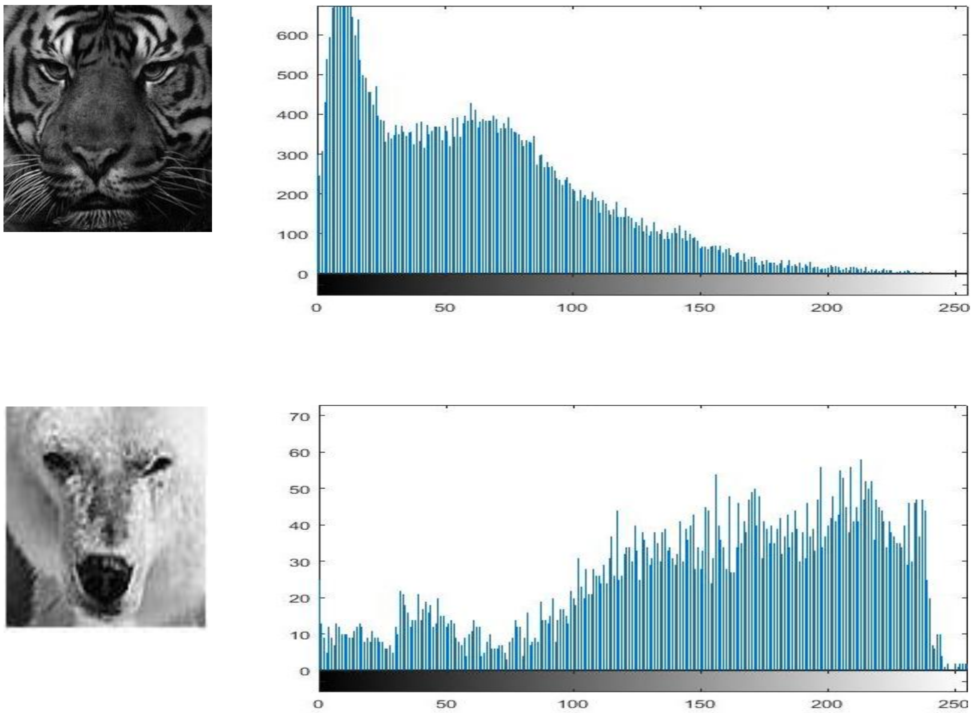

Figure 3.3: Histogram of Dark and Bright Images.

The intensity levels of two images in one database might vary, therefore, to circulate their degrees of intensities, HE algorithm can be applied. This raises the range of intensity and distributes the intensity in a way that are better than having depressed peaks and valleys for images in terms of a histogram [18]. This operation raises the contrast of the low areas in the image while not adversely influencing the overall image contrast.

HE algorithm raises the rate of classification of animals by equalizing the degrees of intensities of all images as much as possible. HE implementation method is close to equalizing the spread of intensity rates in different images. Figure 3.4 shows the variations between images prior to implementing HE algorithm and after its implementation. The contrasts of both images are close to each other after performing HE on them. The original dark image in Figure 3.4 (a) becomes lighter and the original bright images in Figure 3.4 (b) become darker after applying

Histogram Equalization. These results are demonstrated in Figure 3.4 (c) and (d), respectively.

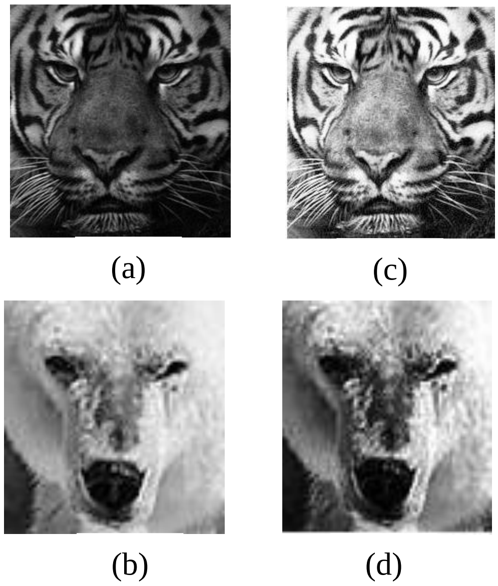

Figure 3.4: (a) and (b) Original Images; (c) and (d) Images after Histogram Equalization Process.

On the other hand, Figure 3.5 shows the distribution of intensity levels of both dark and light images after applying Histogram Equalization.

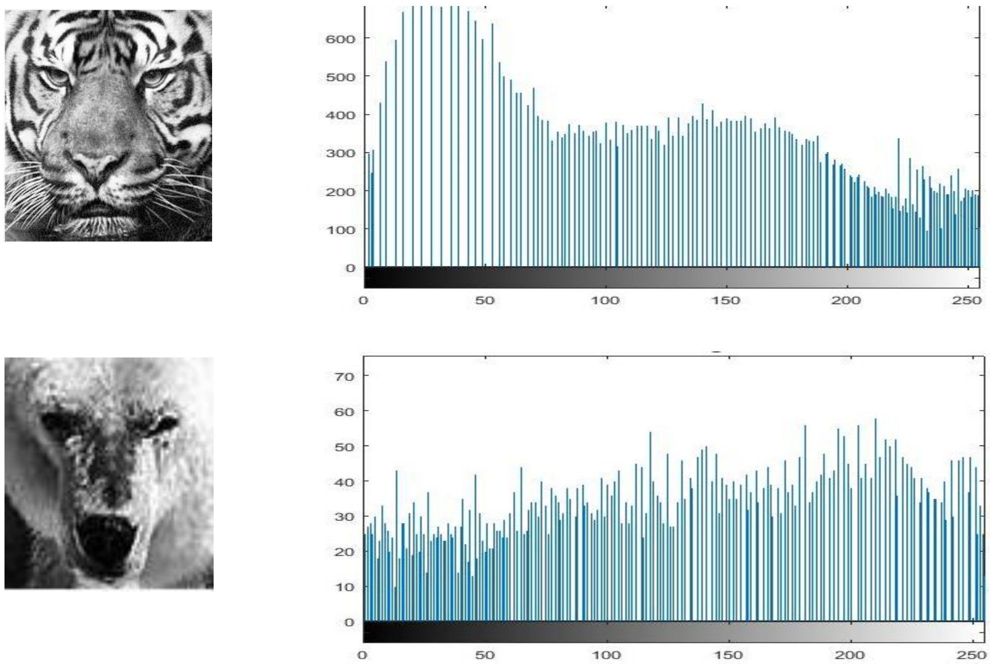

Figure 3.5: Histograms of Dark and Bright Images after Applying Histogram Equalization.

3.4 Mean-Variance Normalization (MVN)

Mean And Variance Normalization (MVN) technique is commonly used to improve the robustness of recognition features [19]. Feature normalization is able to largely decrease the actual mismatch between training and testing conditions. Therefore, HE and MVN [20] have been used to process feature vector in order to significantly increase the classification performance. Mean-variance normalization is applied through the following formula:

$$R = X - M_x \, , MVN = \frac{R}{std(R)} \tag{3.2}$$

where X is a matrix containing the intensity values of images in grayscale,

M_x is mean value of X and $std(R)$ is standard deviation of R.

After the completion of these steps as well as the preprocessing on the animals face, the images are ready to enter the next step called training and testing phases which include the feature extraction techniques that will be explained in the next chapter.

Chapter 4

EXTRACTION OF FEATURES

Animal faces possess many shapes and color of skin referred to as facial attributes that are the fundamental components of a good performing biometrics system. Feature extraction is the most important step for the recognition of facial images. There are many types of feature extraction techniques such as appearance-based, texture-based and model-based approaches. In this thesis, facial images are used for the classification of animal subjects, in which the feature extraction step is done using the appearance-based method Principal Component Analysis (PCA), texture-based Local Binary Patterns (LBP) method and model-based Elastic Bunch Graph Matching (EBGM) method.

PCA projects images on to a plane in a way that the first orthogonal dimension of this sub-plane represents the largest amount of variation among the images and the last dimension of this subspace represents the smallest amount of variation among the images [21] [22]. In this regard, the eigenvectors of the covariance matrix are found which correspond to the direction of the principal component of the initial data and their corresponding eigenvalues represents their statistical quality.

A way of achieving the texture classification of images in gray-scale is employing the LBP texture descriptors to create many local descriptions of the face image and join them together to form a global description. The LBP operator is among the most efficient algorithms to implement texture descriptors as it can be seen in various applications [23]. Smaller regions are formed by dividing the animal facial image, then a vector of weights for each of the smaller regions are computed. LBP builds to a binary number system for each locality and then the histogram of labels can be used as a texture descriptor for all regions. This local feature based method is very effective in the presence of posture variations [24] [25].

On the other hand, EBGM algorithm extracts a stack of jets from the points identified in the face image, then these are employed in training the system. By putting this bunch of jets together, a graph is created that is referred to as a graph node. Later a stack graph, which is a stockpile of face images, is created and it is later compared with the inputted test image for recognition. During the matching process, the jets obtained from the facial features of the query image are matched to every jet in the related stack appended to the stack graph and the one with the most matching score is selected [26].

In our study, we mainly used PCA, LBP and EBGM methods for feature extraction on the animal's face images. Classification strategy is then applied to generate the classification accuracy of the techniques and the results of different systems using PCA, LBP and EBGM methods are compared.

4.1 Appearance-Based Methods

Appearance-based methods work with the entire image, not some parts or specific feature. In the process of analyzing the entire animal face, it endeavors to compare the templates and to the global representation, it aims to identify animal faces [27]. Taking the entire image as one and performing all required operations on it implies the following: there exist just one matrix with one dimension of all pixels in the images and all the information in the images are saved and processed together. This can be both a pro and a con. It is beneficial to have any information about the animal, as much as possible, to discriminate it from the others. However, there may be unwanted data in the image such as the items like tree and grass parts instead of animal face in the background of an image. The performance rate of the face classification system is affected if the second case is present. Figure 4.1 (a) shows an image without any background items, it is an example of animal image after cropping so that the entire animal face, that is every pixel, is useful to distinguishing the animal from others. However, Figure 4.1 (b) has some items such as tree branches, which may be also analyzed during the classification process and may mislead the method that is used.

(a) (b)

Figure 4.1: Example of Cropped Animal Faces.

A one dimensional matrix containing every pixel of the image can be very large in size, and therefore difficult to be analyzed by a machine. Therefore, statistical dimensionality reduction method PCA can be used as an appearance-based method where the features are obtained from the entire animal face image. Afterwards, dimensionality reduction and classification is applied.

Statistical dimensionality reduction algorithms such as Principal Component Analysis have been proven to be largely successful in many academic researches as well as profit-making applications. The success and popularity are mainly due to their statistics-based capability to automatically derive the feature instead of depending on human for their definition. In the next section, PCA method is explained in detail.

4.1.1 Principal Component Analysis

Principal Component Analysis (PCA) is among the most common algorithms utilized in pattern recognition and for compression [28]. It is a dimension decreasing method and feature extraction method [29]. It is a popular statistical method which employs comprehensive method to locate patterns in data with high dimensions. The aim of using PCA come from the information theory approach, where face images are broken down into small sets of distinctive feature images called Eigenfaces that are utilized to describe both the existing as well as introduced faces [30]. In PCA technique, the 2-dimensional matrices of face image are converted into a 1-dimensional vector. This 1-Dimensional array could be either row or column vector, this leads to high dimensional space for representation of images [31].

4.1.1.1 Principal Component Analysis Algorithm

The PCA technique can be done as follows:

1: Add every animal image (x_i) into a row vector of dimension V: $x_i = [x_{i1} \ldots x_{iv}]$, V is the total count of pixels in an image.

2: Calculate the average m and subtract the average from each x_i

$$\mathrm{m} = \left(\sum_{i=0}^{n} x_i\right) * 1/N \tag{4.1}$$

$$y_i = x_i - m \tag{4.2}$$

3: Combine the vectors above to create the data matrix as; $Z = [\, y_1^i \dots \, y_v^i]$

4: Covariance matrix is calculated: $C = ZZ^T$ (4.3)

5: Eigenvalues of covariance matrix C is determined.

6: Sort the eigenvectors and eigenvalues in descending order.

7: Assume $N \leq V$**,** choose the first $d \leq R$ eigenvectors and develop the dataset in the new representation.

8: Match the query images projection matrix with the projection matrices of all training images through the use of a similarity measure. This returns the training image with the closest projection to the test image.

In this thesis, Manhattan Distance Measure is used which is known as City-Block Distance Classifier, L1 Distance, L1 Norm and Rectilinear Distance. It is used with PCA, LBP and EBGM methods as a similarity measure. It shows the distance between points in a city road grid. It considers the absolute value obtained by subtracting the coordinates of two objects from each other [32], as follows:

$$D(x, y) = \sum_{i=1}^{n} |x_i - y_i| \qquad (4.4)$$

In the classification stage, we use Nearest-Neighbor classifier which is also known as NN-based classifiers. Nearest-Neighbor classifier makes classification decision directly based on the data which does not require training or learning of parameters, Nearest-Neighbor classifier classifies an image by the class of its nearest (most similar) image in the database [33].

4.2 Texture-Based Methods

Texture-based methods divide the images into equal length regions that do not overlap and then obtain the local features from all created regions using a texture-based method such as Local Binary Patterns. The extracted local features represent the local projection, which are finally concatenated into a global feature of the initial image for texture classification. The global feature analogous to the training image projections are matched with the query image projections using a distance measure such as Manhattan distance measure to detect the animal face. LBP method is implemented in this study with 5x5 regions.

4.2.1 Local Binary Patterns (LBP)

Local Binary Patterns is an efficient but simple approach employed to get features that are utilized for classification process in computer vision. The most advantageous properties of LBP features are the insensitivity against differences in illumination and computation simplicity [34]. These useful features allow LBP approach to be used in real time [35]. This implies that, given a pixel position (x_c, y_c), LBP is described as a sorted set of binary matching of pixel intensities between the pixel at the center and its neighboring pixels as follows:

$$LBP(x_{c,} y_c) = \sum_{n=0}^{7} s\,(I_n - I_c)\,2^n, \qquad (4.5)$$

where I_c corresponds to the gray value of the center pixel $(x_{c,} y_c)$, I_n to the gray values of the 8 surrounding pixels, and thresholding function $s(k)$ is defined as:

$$s(k) = f(x) = \begin{cases} 1, & if\ k \geq 0 \\ 0, & if\ k < 0 \end{cases} \qquad (4.6)$$

LBP approach gives an efficient way of texture description [36]. LBP features are obtained in gray scale and texture operators are robust against rotation. LBP approach is a fast feature extraction method and it provides a good performance that makes it the most commonly researched feature extractor approach. Figure 4.2 illustrates an example of LBP feature extraction process.

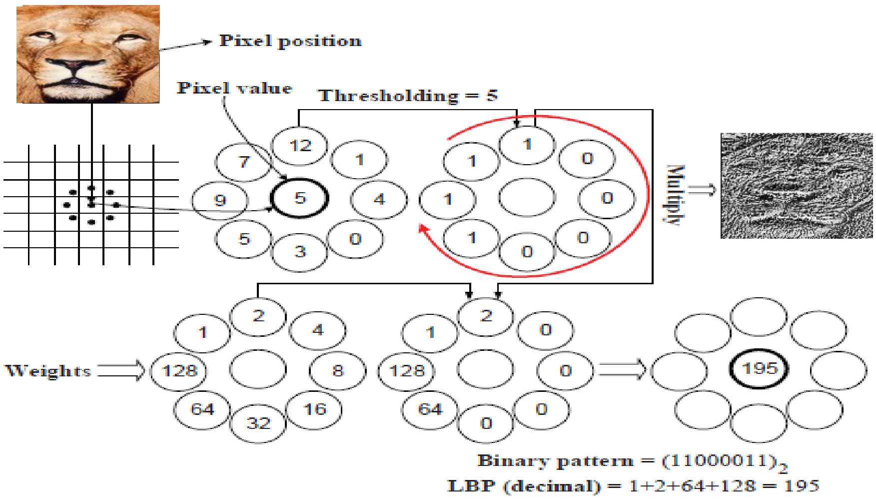

Figure 4.2: An Example of an LBP Feature Extraction Process [34].

LBP approach is applied on images that are divided into 5x5 partitions obtaining a total of 25 partitions (blocks). For each block, LBP histogram is obtained separately. Histograms of these blocks are then concatenated and used as a texture descriptor. An example of LBP approach carried out on an original animal image that is divided into 5x5 blocks is shown in Figure 4.3.

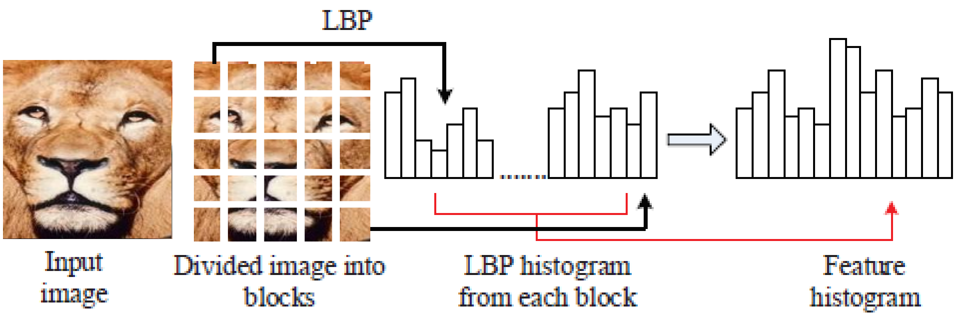

Figure 4.3: An Example of LBP Approach [34].

We can obtain LBP features using the following technique:

- Divide images into blocks of equal size (e.g. 5x5 blocks for each query image) with the aim of extracting the LBP histogram from every segment.
- For every pixel in a segment, take its 8 neighbors and compare these to the pixel (that is, on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels in a clockwise manner.
- If the value of pixel at the center is greater than the value of its neighbor, give it a "1" else, give it a "0". This generates an 8-digit binary number which is usually changed to decimal value.

LBP feature extraction is carried out on a query image (for example, lion face image) and the output features can be seen in Figure 4.4 along with the histogram of the feature. LBP approach tries to raise the strength of the feature and to generate more details, therefore the query images are grouped into a number of blocks (such as 5x5 blocks in our case).

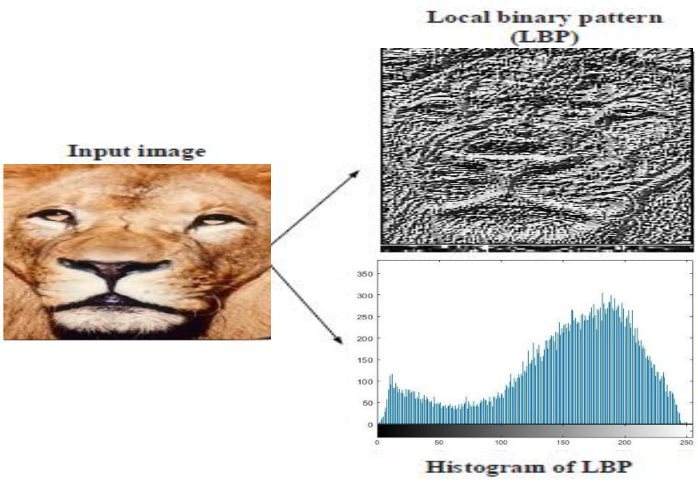

Figure 4.4: An Example of Local Binary Patterns and its Histogram [34].

It should be noted that not all groups in the LBP histogram must have important information for image description. Therefore it is advantageous to recognize the discerning LBP-Histogram (LBPH) bins for improved image description [37].

4.3 Model-Based Methods

Model based approaches depend on information about well-defined facial attributes and the area around the points to describe a face in the problem space and carry out the recognition process. These facial attributes include the mouth, eyes, nose, etc. Ideally, the actual coordinates of the eyes are given, however, in practice the technique is able to estimate based on the values obtained from face detection and eye zone locator module used before the recognition process [38]. An example of a model-based technique is the Elastic Bunch Graph Matching (EBGM) algorithm [39, 40], which stores spectral information about the neighborhoods of facial features by convolving these areas with Gabor wavelets (masks). EBGM method is explained in the following subsections in detail.

4.3.1 Elastic Bunch Graph Matching (EBGM) Method

Elastic Bunch Graph Matching approach trains the system by extracting a stack of jets from the key-points available in the face image. By appending stack of jets, a graph is generated which is referred as a graph node. Later a bundle graph, which is a stockpile of facial images, is extracted which are later matched with the query image given for identification. While matching an image to the bundle graph, the jets obtained from the facial features of the query image are matched with all jets in the corresponding bundle corresponding to the bundle graph and the most similar one is chosen [41].

A bunch graph consists of different faces having varying features or properties to obtain a model graph. The nodes attaching the model graph are placed over the query image and jets are obtained from the fiducial points. These are employed for comparing the model graph obtained at the training stage with the test image. The final stage is using a similarity function to obtain a match between the query image and the model graph [42].

4.3.1.1 Elastic Bunch Graph Matching (EBGM) Algorithm

In the EBGM algorithm, faces are described as labeled graphs, with nodes placed at fiducial points (mouth, eyes, nose, etc.) dependent on Gabor Wavelet Transform (GWT) [39, 40]. EBGM is believed to be the best in respect to rate of identification and performance reliability but, bad illumination adversely affects recognition. EBGM algorithm is applied in this thesis using the same method as in [43]. A set of Gabor wavelet coefficients for every point is produced. Several feature points describing the local features are obtained from the training faces. Each feature point is described by a feature vector combining the GWT coefficients and its coordinates. All feature vectors are added together to constitute a face that is used in matching, recognition and classification processes. The EBGM technique is represented as follows:

Input: A training image $I = [\, I(x_{1)}, I(x_2), \ \ldots\ldots\ I(x_n)]$; where $I(x_i)$ is the image intensity value at $\boldsymbol{x_i}$, and x_i is the coordinate vector of point I , n is the total number of points.

Output: A collection of feature arrays represented as $v_k = \{x_k, y_k, R_j(x_k, y_k)\ , j = 1, 2\ , \ldots\ \ , 40\ \}$, which represents the k^{th} feature array of the face image. $R_j(x_k, y_k)$ represents the GWT coefficients at (x_k, y_k) point.

EBGM Algorithm:

1) Apply GWT to the image $I, R_i = \int I(x')\psi_i(x-x')d^2\,x'$, with Gabor filter

$$\psi(x) = \frac{\|k_i\|^2}{\sigma^2} e^{\frac{\|k_i\|^2\|x_i\|^2}{2\sigma^2}} \left[e^{jk_i \cdot x - e^{-\frac{\sigma^2}{2}}} \right], \quad (4.7)$$

where

$$k_i = \begin{pmatrix} k_{ix} \\ k_{iy} \end{pmatrix} = \begin{pmatrix} k_\upsilon & \cos\theta_\mu \\ k_\upsilon & \sin\theta_\mu \end{pmatrix};\ k_\upsilon = 2^{\frac{\upsilon+2}{2}}\pi;$$

$$\theta_\mu = \mu\frac{\pi}{2}\ ;\ \upsilon = 0, \ldots, 4; u = 0, \ldots, 7; i = \mu + 8\upsilon. \quad (4.8)$$

2) Locate feature points by searching the location in a window W_0 of size $w \times w$. A feature point is found at (x_0, y_0) if

$$R_j(x_0, y_0) = \max(R_j(x, y)) \quad (4.9)$$

$$R_j(x_0, y_0) > \frac{1}{N_1 N_2} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} R_j(x, y), \quad (4.10)$$

where $(x, y) \in W_0, j = 0, \ldots, 40$ and N_1, N_2 is the dimension of the animal face image. Then the window size should be chosen to be small enough to represent the relevant features and also big enough to deal with cases of redundancy.

3) Obtain feature vectors at feature points as a composition of GWT coefficients:

$$\upsilon_k = \{x_k, y_k, R_j(x_k, y_k), j = 1,2, \ldots, 40\}. \quad (4.11)$$

EBGM technique has two benefits over other techniques [44, 45] including its easy implementation and unnecessariness to manually localize the training graphs. However, in order to decrease the computation cost significantly, the face images must be well partitioned in order to get rid of meaningless/repeated information concerning background from the initial images.

Chapter 5

EXPERIMENTS AND RESULTS

Several preprocessing techniques and feature extraction methods are used in this thesis to classify the animals' faces. The experiments are performed using 4 datasets namely LHI animal faces, KTH (profile and front) and HiT to evaluate the performance of appearance-based (PCA), texture-based (LBP) and model-based (EBGM) techniques.

In experiments, all the training images as well as test images used are cropped and resized to 150x150, 100x100 and 50x50 pixels from the original sizes in different experiments. All images only include the head of animals as explained in the third chapter.

Manhattan distance measure [32] was used for the comparison of the training and test images. This is the classification of a test face image according to the corresponding faces learned by the system. Manhattan distance measure is calculated by subtracting the test image projections from the training image projection and summing up this difference for each animal. The minimum difference obtained by the algorithm is taken as the most similar face to the query image, therefore it is returned as the matched image. Classification accuracy for a given dataset is computed by the number of images correctly classified to the total number of test images in the dataset.

5.1 Animal Databases

There exist many animal databases in the literature. However, a few of these databases are publicly available. These databases include images of many animal species. The following table (Table 5.1) demonstrates some of these databases with the detailed information including animal species, number of images available and the modalities used in these images for performing animal classification. In this thesis, we used four different animal datasets using three databases namely LHI, HiT and KTH databases. The details related to each database are explained in the following subsections.

Table 5.1: List of Animal Databases.

#	Database	Species	Number of images	Modalities
1	NMML	humpback whale	24,428	Marks Or Scars
2	The Corel collection	Zebras, tigers, giraffes and random animals	304	Texture Patches
3	Elephant photo database by Zoological Dep. Of the University of Rome	Elephants	332 and 268 subjects	Ear Nicks
4	MPI2	Ape + chimpanzee+ gorilla	15,000	Face
5	PASCAL 2007	cats	679	Face
6	ChimpZoo, ChimpTai and Gorilla	chimpanzee	6526	Face
7	Dogster, Catster and Mydogspace	pet animal (dog and cat)	3000	Face
8	LHI Animal Faces	Random animals, 19 subjects	2312	Face
9	Hybrid Image Templates (HiT)	18 different subjects of animal	662	Face
10	KTH-Animals(Front and profile)	14 subjects of random animals	193 frontal and 159 profile images	Frontal and Profile Face

5.1.1 LHI Animal Faces Database

The LHI-Animal-Faces database is a good database for classification of animals, because of the interesting intra-class variation and inter-class confusion exhibited by the animal face categories. Intra-class variation may include rotation and flip transforms, such as rotated lion faces and faces oriented left or right, panda heads; posture differences, such as rabbits with upright straight ears against those with relaxed ears. Additionally, the within-class variation include sub-types, such as male and female lions. The inter-class confusion is majorly arises from shared parts such as wolves and cats both having pointed ears roughly at the same position; cow and sheep faces sharing a similar contour. The images of LHI significantly display variability in resolution, quality and illumination. For our experimental studies, we used just faces of animals. The total number of the images is 2312. Among these, 950 images were used for train and 1362 for test. The LHI database includes 19 subjects namely bear, cat, dog, chicken, cow, elephant, lion, monkey, deer, duck, eagle, mouse, panda, pigeon, pig, rabbit, sheep, tiger, and wolf. Figure 5.1 shows some of the samples of animals faces of LHI database.

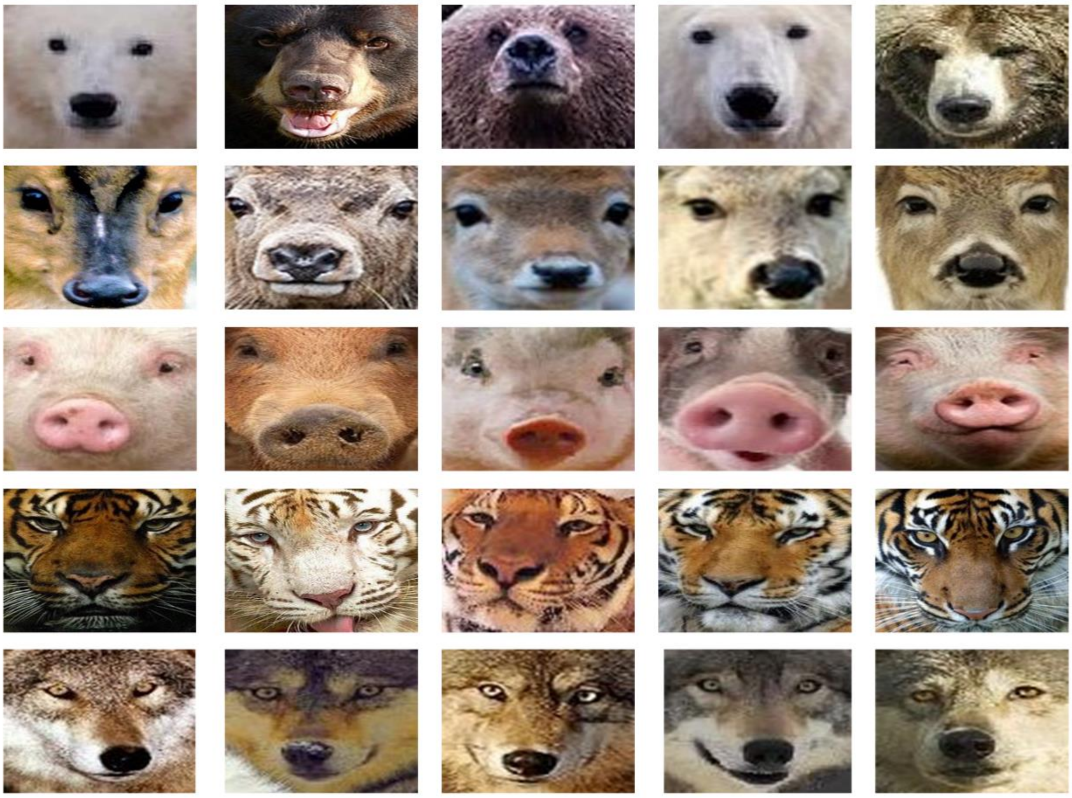

Figure 5.1: Five Subjects from LHI Animal Faces Database

5.1.2 KTH (Front and Profile) Database

KTH database is an animal database including both frontal and profile face images. The images are low quality JPEG images like most of the images found on the internet. Both datasets, namely Front and Profile, have 14 subjects and the names of these subjects are bear, cougar, coyote, deer, giraffe, gorilla, kangaroo, leopard, lion, panda, penguin, sheep, tiger, and zebra. First of all, KTH-Front database contains 193 images in which 70 of them are used for train and 123 for test. There are approximately 5 images per subjects for train. KTH-Profile dataset includes 159 images, in which 70 images are used for train and 89 for test. Only one subject (panda) has just 9 images and 5 of them are used for train and 4 images for test. In Figures 5.2 and 5.3, some samples of images for both KTH Frontal and Profile datasets are demonstrated.

Figure 5.2: Examples of Some Faces of Animals from KTH-Front Dataset.

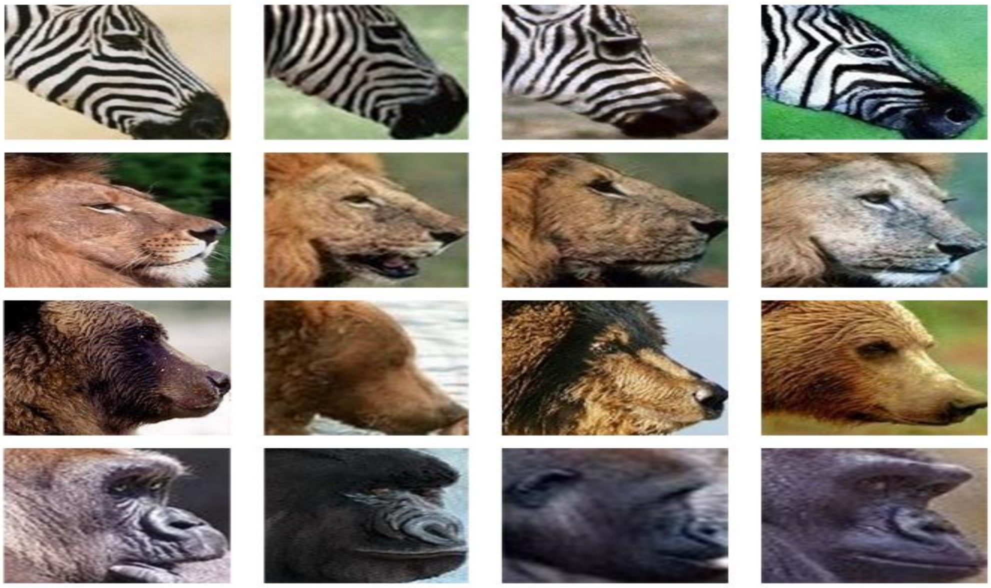

Figure 5.3: Examples of Some Faces of Animals from KTH-Profile Dataset.

5.1.3 HiT Database

HiT database is the third database we used in our experiments. The images are good quality JPEG images like most of the images found on the LHI Animal Faces Database. The database has 18 subjects and HiT database subjects are bear, cat, cow, dog, eagle, elephant, fish, lion, monkey, panda, pigeon, pig, rabbit, rabbit2, sheep, tiger, wolf, and zebra. Figure 5.4 shows some face images of animals in HiT database. This database contains 662 images. In our experiments, we used 180 images for training and 482 images for testing. Additional details about the amount of training and probe images used in each database are shown in Table 5.2 and Table 5.3 with the name of subjects in all databases.

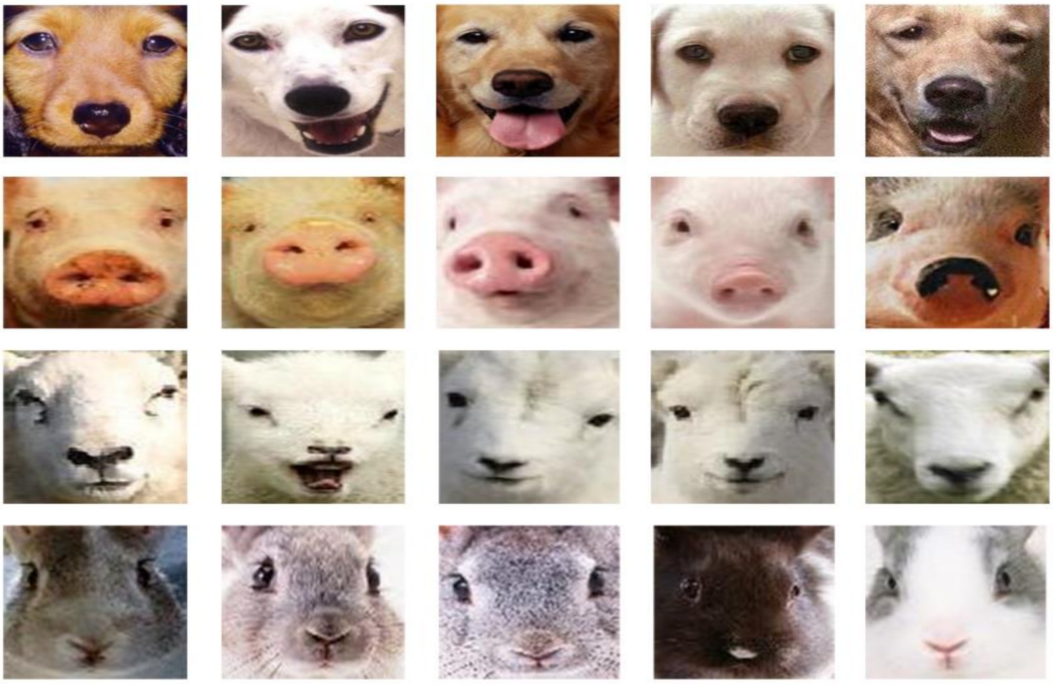

Figure 5.4: Examples of Some Faces of Animals from HiT Database

Table 5.2: Number of Subjects and Number of Samples for Each Subject in Animal Databases.

Database Name	Number of subjects	Number of samples for each subjects		
		Min	Max	Total number of images
LHI Animal Faces	19	100	389	2312
KTH-Animals(Front)	14	10	24	193
KTH-Animals(Profile)	14	9	24	159
Hybrid Image Templates (HIT)	18	13	88	662

Table 5.2 shows the number of subjects in each animal database, minimum and maximum number of images for different animal species in the databases and the total number of images for each database. Table 5.3 shows the name of subjects (the name of animal species) in four databases.

Table 5.3: Subject Names in Animal Databases.

No	LHI	KTH-Animals(Front)	KTH-Animals(Profile)	Hybrid Image Templates(HiT)
1	BearHead	bear	bear	bearHead
2	CatHead	cougar	cougar	catHead
3	ChickenHead	coyote	coyote	cowHead
4	CowHead	deer	deer	dogHead
5	DeerHead	giraffe	giraffe	eagleHead
6	DogHead	gorilla	gorilla	elephantHead
7	DuckHead	kangaroo	kangaroo	fish
8	EagleHead	leopard	leopard	lionHead
9	ElephantHead	lion	lion	monkeyHead
10	LionHead	panda	panda	pandaHead
11	MonkeyHead	penquin	penquin	pigeonHead
12	MouseHead	sheep	sheep	pigHead
13	PandaHead	tiger	tiger	rabbitHead
14	PigeonHead	zebra	zebra	rabbitHead2
15	PigHead	-	-	sheepHead
16	RabbitHead	-	-	tigerHead
17	SheepHead	-	-	wolfHead
18	TigerHead	-	-	zebraHead
19	WolfHead	-	-	-

5.2 Animal Classification Using Several Train/Test Ratios

The experimental results using different train and test ratios are presented in this section. The experimental results shown in Table 5.4 are the results we obtained in this thesis and for comparison, results from Reference [2] are included in Table 5.5. In our experiments, we used PCA, LBP and EBGM approaches as shown in Table 5.4. We used the same experimental setup as in Reference [2]. In this experimental setup, we just selected five subjects from LHI database (bear, deer, pig, tiger, and wolf) and each subject has 100 images. The size for images is 150x150 pixels and the experiments are grouped into six major categories; A, B, C, D, E, and F. The first part of our experiments utilized 90 % of the images for training and 10 % for testing. The second part used 80 % of the images for training and 20 % for testing. The third part of our experiments utilized 70 % of the images for training and 30 % for testing. The next part used 60 % of the images for training and 40 % for testing. The following part utilized 50 % of the images for training and 50 % for testing. The final part of our experiments used 40 % of the images for training and 60 % for testing.

Table 5.4: Six Categories of Experiments on LHI Database.

Categories / Accuracy (%) using	A	B	C	D	E	F
PCA	88	79	73	72	69	64
LBP	90	85	77	75	71	69
EBGM	94	89	87	88	86	85

Table 5.5: Six Categories of Experiments from [2].

Categories / Accuracy (%) using	A	B	C	D	E	F
PCA	85	77	72	64	62	61
LBPH	88	84	76	73	71	67

The best classification rate (accuracy of 94 %) is achieved using model-based EBGM approach for the first part of the performed experiments. For category A, 90 % of the images were used as training images while 10 % were used as test images. The worst classification rate of 64 % accuracy was obtained for the last part of our experiments, category F where 40 % of the images were used for training and 60 % for testing. When we compare our results to the results in [2], PCA results are close to each other for all categories. Our PCA results are slightly better than the results obtained in [2]. On the other hand, LBP and LBPH results are compatible in all categories. Moreover, the best results are obtained using EBGM method and the obtained accuracies using EBGM method in all categories outperform the other methods as shown in Table 5.4.

5.3 Animal Classification Based on Five Subjects

The second set of experiments is conducted using five animal subjects. The experimental results are shown in Table 5.6. This set of experiments is done with different sizes (150x150, 100x100 and 50x50 pixels). We selected 5 subjects from each database. The selected LHI database subjects are BearHead, DeerHead, PigHead, TigerHead, and WolfHead. KTH front and profile datasets include the same subjects such as BearHead, DeerHead, LionHead, TigerHead, and ZebraHead and HiT database includes BearHead, LionHead, PigHead, TigerHead and

WolfHead. Using PCA, LBP and EBGM algorithms, we performed feature extraction and then the Nearest Neighbor classification results are shown below in Table 5.6 and the results are in (%) for classification rates.

Table 5.6: Animal Classification Rates (%) on Four Datasets.

		LHI	HiT	KTH(front)	KTH(profile)
1		BearHead	BearHead	BearHead	BearHead
2		DeerHead	LionHead	DeerHead	DeerHead
3		PigHead	PigHead	LionHead	LionHead
4		TigerHead	TigerHead	TigerHead	TigerHead
5		WolfHead	WolfHead	ZebraHead	ZebraHead
Number of train images		250	50	25	25
Number of test (min-max) images		250(50 – 50)	174(9 – 78)	50(5 – 19)	42(5 – 19)
PCA	Size: 150 x 150	69	53	50	61
LBP		71	70	42	59
EBGM		**86**	**78**	**75**	**87**
PCA	Size: 100 x 100	69	53	50	61
LBP		71	70	46	58
EBGM		**90**	**85**	**84**	**83**
PCA	Size: 50 x 50	70	53	45	60
LBP		63	53	43	50
EBGM		**87**	**83**	**77**	**78**

When we use the size of 150x150, the best result we obtained was 87% accuracy using EBGM on KTH (profile) dataset. For the second size of 100x100, classification accuracy we get was 90 % using EBGM algorithm on LHI database.

For the last size of 50x50, we obtained 87% accuracy on LHI database. In general, the best result we obtained for classification accuracy was achieved using the size of 100x100. Because of that we decided to use the size of 100x100 in the rest of the experiments.

5.4 LHI Database Animal Classification Results

These set of experiments is done with all subjects in LHI database and the total number of subjects are 19 subjects in which the images are resized to 100x100 pixels. Total number of training images used are 950 images, and for testing images we used 1362 images. The classification performance using 3 methods (PCA, LBP and EBGM) are presented in Table 5.7. Individual accuracy using each feature extraction method for each subject and the total accuracy for the database are shown in Table 5.7. For PCA, the best accuracy result we obtained is 72% (TigerHead). For LBP, the best accuracy is obtained as 86% (ElephantHead) and for EBGM it is 94% (DuckHead). In general, for all subjects, the best accuracy is achieved using EBGM method which is 69 %.

5.5 HiT Database Animal Classification Results

HiT database includes images of animal faces of 18 subjects with resized to 100x100 pixels. Performance obtained using this database is better than the performance obtained by LHI database. The overall animal face classification performances using PCA, LBP and EBGM methods are presented in Table 5.8.

Table 5.7: Animal Classification on LHI Database using PCA, LBP and EBGM with Size 100x100.

#	Subjects	Number of train images	Number of test images	Individual Accuracy (%)		
				PCA	LBP	EBGM
1	BearHead	50	51	35	67	76
2	CatHead	50	110	26	61	77
3	ChickenHead	50	50	36	40	58
4	CowHead	50	54	30	43	63
5	DeerHead	50	53	30	72	62
6	DogHead	50	339	15	25	43
7	DuckHead	50	53	42	40	94
8	EagleHead	50	51	22	24	59
9	ElephantHead	50	50	44	86	66
10	LionHead	50	52	31	67	85
11	MonkeyHead	50	50	14	30	70
12	MouseHead	50	50	22	30	52
13	PandaHead	50	69	30	59	86
14	PigeonHead	50	65	28	43	82
15	PigHead	50	51	20	37	53
16	RabbitHead	50	50	30	40	56
17	SheepHead	50	50	26	38	56
18	TigerHead	50	64	72	83	88
19	WolfHead	50	50	58	76	84
	Total	950	1362	32	51	69

Table 5.8: Animal Classification on HiT Database using PCA, LBP and EBGM with Size 100 x 100.

#	Subjects	Number of train images	Number of test images	Individual Accuracy (%)		
				PCA	LBP	EBGM
1	bearHead	10	39	33	54	51
2	catHead	10	10	10	40	90
3	cowHead	10	27	30	56	56
4	dogHead	10	19	21	11	37
5	eagleHead	10	10	60	40	80
6	elephantHead	10	22	45	95	73
7	fishHead	10	25	60	60	96
8	lionHead	10	38	50	71	95
9	monkeyHead	10	40	8	10	38
10	pandaHead	10	6	83	67	100
11	pigeonHead	10	48	60	48	96
12	PigHead	10	78	19	46	29
13	rabbitHead	10	10	50	50	90
14	rabbitHead2	10	37	27	27	51
15	sheepHead	10	51	29	39	55
16	tigerHead	10	10	50	80	100
17	WolfHead	10	9	56	100	100
18	zebraHead	10	3	67	33	100
	Total	180	482	42	51	74

5.6 KTH Frontal Dataset Animal Classification Results

Classification using KTH Frontal dataset with 14 subjects of size 100x100 is performed using PCA, LBP and EBGM methods. Performance using this database is not good compared to the accuracies obtained with HiT database. The overall animal face classification performance using three feature extraction methods are presented in Table 5.9.

Table 5.9: Animal Classification on KTH Front Dataset using PCA, LBP and EBGM Size 100 x 100.

#	Subjects	Number of train images	Number of test images	Individual Accuracy (%)		
				PCA	LBP	EBGM
1	Bear	5	15	20	53	47
2	Cougar	5	5	20	100	80
3	Coyote	5	12	50	17	83
4	Deer	5	5	20	20	60
5	Giraffe	5	9	11	78	22
6	Gorilla	5	5	40	60	80
7	Kangaroo	5	10	20	10	10
8	Leopard	5	8	13	13	63
9	Lion	5	5	20	40	60
10	Panda	5	9	44	56	100
11	Penguin	5	5	40	20	40
12	Sheep	5	10	40	10	40
13	Tiger	5	19	21	53	84
14	Zebra	5	6	67	67	83
	Total	70	123	30	43	61

5.7 KTH Profile Dataset Animal Classification Results

Classification rates using KTH Profile dataset with the same subjects as in KTH Front dataset are obtained with images resized to size 100x100 pixels. Performance obtained using this dataset is near to the accuracy result obtained in KTH Front dataset. Animals face classification performance using three methods on KTH profile dataset are shown in Table 5.10.

Table 5.10: Animal Classification on KTH Profile Dataset using PCA, LBP and EBGM Size 100 x 100.

#	Subjects	Number of train images	Number of test images	Individual Accuracy (%)		
				PCA	LBP	EBGM
1	Bear	5	8	63	38	75
2	Cougar	5	5	40	40	60
3	Coyote	5	7	29	14	71
4	Deer	5	5	40	60	80
5	Giraffe	5	7	43	43	29
6	Gorilla	5	3	33	33	67
7	Kangaroo	5	5	20	60	20
8	Leopard	5	5	20	20	40
9	Lion	5	5	20	20	40
10	Panda	5	4	75	75	100
11	Penguin	5	6	17	33	17
12	Sheep	5	5	20	20	20
13	Tiger	5	5	40	40	100
14	Zebra	5	19	47	26	74
	Total	70	89	36	37	57

5.8 Overall Animal Classification Results on Four Datasets

The animals face classification performance review using PCA, LBP and EBGM methods on all databases are presented in Table 5.11.

Table 5.11: Animal Classification Accuracy of all Databases using Three Methods.

#	Databases	Accuracy (%)		
		PCA	LBP	EBGM
1	LHI	32	51	**69**
2	HiT	42	51	**74**
3	KTH-front	30	43	**61**
4	KTH-profile	36	37	**57**

The best classification accuracy result is obtained using EBGM method as 74% using HiT database. EBGM achieves the best accuracies for all the databases compared to PCA and LBP results.

5.9 Human Classification Experiments

Humans can be considered as one of the species of animals. Classification of human face in order to identify different people is another interesting research area. At the end of this thesis, the similarities and differences between animal classification and human classification are studied by conducting some experiments on human identification. Animal images are taken in the wild, therefore these images represent large amount of intra-class and inter-class variations which decrease the performance of animal classification systems. However, controlled human face images help biometrics system to identify human with high performance.

There is a large number of human face databases available such as FERET, LFW, ORL, FRGC, AR and etc. Experiment on human faces is important to see the efficiency of the methods used in this study. Therefore, we also performed classification on human face to analyze the results of classification accuracy using PCA, LBP and EBGM methods. For this purpose, we use two human face databases namely FERET database and ORL database. We summarized information about these two available databases that have been used in the next subsections. We got good accuracy results on human face databases FERET and ORL that are presented in the next subsection.

5.9.1 FERET Database

The Face Recognition Technology (FERET) program database is a major database of facial images. Its evaluation methodology is made up of a consolidated data acquisition process and testing program. These two important stages are put together through the FERET database of face images. The database is subdivided into two parts: the training or learning portion of the database which is available to researchers, and a second part is for testing. This division of the database facilitates the training and testing of the algorithms using different sets of images for both processes, though, the images should be related. The FERET evaluation system administer tests for face-recognition independently. The test was fashioned to direct comparison between different algorithms possible, correctly choose the most efficient approaches, evaluate the state-of-the-art in face recognition, detect future areas of research, and encourage modern hybrid researches in area of face recognition [46]. Figure 5.5 shows 5 subjects with 2 sample face images in different lighting condition and expressions with cropped and rescaled face images.

Figure 5.5: Examples of Images of FERET Database.

5.9.2 ORL Database

This is a database of face images which contains a set of face images captured in a laboratory during a period of 2 years (between April 1992 and April 1994). ORL have 40 different subjects and each subject has 10 different images with the total number of 400 images. Each image has a size of 92x112 pixels, and 256 gray levels per pixel. The images were captured at different times for some subjects, lighting was varied for some, different facial posture and expressions such as open versus closed eyes, different expressions like smiling and not smiling were introduced as well as other facial details like putting glasses on and removing glasses. All the capturing of images were done by positioning the subjects in an upright, frontal position, with tolerance for some slight head posture change, and against a dark consistent background [47]. Figure 5.6 shows 3 sample subjects having 10 different images for each.

Figure 5.6: Three Sample Subjects with 10 Different Images from ORL Database.

5.9.3 Performance of Different Methods on FERET and ORL databases

Classification of humans in order to identify or recognize each person is studied with PCA, LBP and EBGM methods in this set of experiments. FERET and ORL databases are used in the experiments which include controlled images of human faces. Performances of each method on two databases are given in Table 5.12. EBGM method gives the best performance as %98.2 in FERET dataset and as %96.0 in ORL database.

Table 5.12: Human Classification Accuracy using PCA, LBP and EBGM Methods.

#	Methods	Accuracy (%)	
		FERET	ORL
1	PCA	92.8	84.5
2	LBP	98.2	93.5
3	EBGM	**98.2**	**96.0**

The best results for both animal and human classification experiments are obtained with EBGM method. However, the results on human classification are better than the results of animal classification since human face images are taken at controlled conditions and do not include large amount of variations.

5.10 Discussion on Experimental Results

In general, animal classification experiments are performed using three feature extraction methods namely PCA, LBP and EBGM. Each method is evaluated on different animal face databases. Different size of images as well as different amount of training and test images are utilized to obtain the performances. The results of the experiments show that EBGM method provides better results than PCA and LBP for different size and different amount of training images. The best experimental results of animal face classification are obtained using model-based EBGM method. The outcome of the experiments indicates that EBGM method provides the best classification rate for both animal classification on uncontrolled animal databases and human classification on controlled human face databases.

Chapter 6

CONCLUSION

A comparative analysis of appearance-based PCA, texture-based LBP and model-based EBGM methods for the solution of animal face classification task was presented in this thesis. Experiments were carried out using 4 different animal datasets namely LHI, HiT, KTH Profile and KTH Front datasets. The results demonstrated the efficiency of EBGM algorithm over the other techniques on experiments using several train/test ratios with the images. In the case of different size of animal faces, EBGM achieved better accuracies compared to PCA and LBP performances on all datasets. It can be stated that EBGM generally achieves better results for animals face classification on all subjects for different datasets of LHI, HiT and KTH (profile and front) databases.

Animal face databases including animal face images taken at uncontrolled conditions reduce the classification accuracy. However, it is not easy to take animal pictures in controlled conditions. Therefore, the experiments on animal classification present lower accuracies compared to human classification experimental results. The similarities and differences between animal and human classification are also presented in this thesis. The experimental results show that the accuracies are different for animal and human classification. Human classification accuracies are higher than animal classification accuracies whenever the same feature extraction methods are used. However, the same feature extraction method, namely model-based EBGM method, achieved the best results compared to appearance-based PCA and texture-based LBP methods on both animal and human classification experiments.

As a future work, different animal databases with larger amount of animal images can be used to perform animal classification experiments. Additionally, other appearance-based, texture-based and model-based approaches can be implemented to present the animal classification results on the available animal databases. Moreover, various deep learning techniques can be studied for the solution of animal classification problem.

REFERENCES

[1] D. Masri, Z. Aung, and W. L. Woon, “Image classification using appearance based features,” *2015 11th International Conference on Innovations in Information Technology (IIT)*, 2015.

[2] T. Trnovszky, P. Kamencay, R. Orjesek, M. Benco and P. Sykora, "Animal Recognition System Based on Convolutional Neural Network", *Advances in Electrical and Electronic Engineering*, vol. 15, no. 3, 2017.

[3] The LHI Database. Retrieved from http://www.stat.ucla.edu/~zzsi/HiT/exp5.html

[4] The HiT Database. Retrieved from http://www.stat.ucla.edu/~zzsi/hit/

[5] The KTH Database. Retrieved from https://www.csc.kth.se/~heydarma/Datasets.html

[6] S. Taheri and Ö. Toygar, "Animal classification using facial images with score-level fusion", *IET Computer Vision*, vol. 12, no. 5, pp. 679-685, 2018.

[7] S. A. Mizroch, & S. A. D. Harkness, “A test of computer-assisted matching using the North Pacific humpback whale, Megaptera novaeangliae, tail flukes photograph collection”. *Marine Fisheries Review*, 65, 25-37, (2003).

[8] D. Ramanan, D. Forsyth and K. Barnard, "Building models of animals from video", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1319-1334, 2006.

[9] M. Domeier and N. Nasby-Lucas, "Annual re-sightings of photographically identified white sharks (Carcharodon carcharias) at an eastern Pacific aggregation site (Guadalupe Island, Mexico)", *Marine Biology*, vol. 150, no. 5, pp. 977-984, 2007.

[10] A. Ardovini, L. Cinque and E. Sangineto, "Identifying elephant photos by multi-curve matching", *Pattern Recognition*, vol. 41, no. 6, pp. 1867-1877, 2008.

[11] M. Lahiri, C.Tantipathananandh, R. Warungu,D. I. Rubenstein and T.Y. Berger-Wolf, “Biometric animal databases from field photographs: identification of Individual zebra in the wild”, *In Proceedings of the 1st ACM International Conference on Multimedia Retrieval*, pp.6. (2011).

[12] M. Çarıkçı and F. Özen, "A Face Recognition System Based on Eigenfaces Method", *Procedia Technology*, vol. 1, pp. 118-123, 2012.

[13] A. Ernst and C. Küblbeck, "Fast face detection and species classification of African great apes," *2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS),* Klagenfurt, pp. 279-284, 2011.

[14] Weiwei Zhang, Jian Sun and Xiaoou Tang, "From Tiger to Panda: Animal Head Detection", *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1696-1708, 2011.

[15] A. Loos and M. Pfitzer, "Towards automated visual identification of primates using face recognition," *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, Vienna, pp. 425-428, 2012.

[16] S. Kumar and S. Singh, "Biometric Recognition for Pet Animal", *Journal of Software Engineering and Applications*, vol. 07, no. 05, pp. 470-482, 2014.

[17] C. Finn, J. Duyck, A. Hutcheon, P. Vera, J. Salas and S. Ravel, “Relevance feedback in biometric retrieval of animal photographs”, *Pattern recognition*, pp. 281–290, 2014.

[18] B. Weyrauch, B. Heisele, J. Huang and V. Blanz, "Component-Based Face Recognition with 3D Morphable Models," *2004 Conference on Computer Vision and Pattern Recognition Workshop*, Washington, DC, USA, pp. 85-85, 2004.

[19] P. Pujol, D. Macho and C. Nadeu, "On Real-Time Mean-and-Variance Normalization of Speech Recognition Features," *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, Toulouse, pp. I-I, 2006.

[20] G. Zhou, Y. Zhan and J. Zhang, "Facial Expression Recognition Based on Selective Feature Extraction," *Sixth International Conference on Intelligent Systems Design and Applications*, Jinan, pp. 412-417, 2006.

[21] M. Turk and A. Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[22] P. Belhumeur, J. Hespanha and D. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

[23] T. Ojala, M. Pietikäinen and D. Harwood, "A comparative study of texture measures with classification based on featured distributions", *Pattern Recognition*, vol. 29, no. 1, pp. 51-59, 1996.

[24] T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.

[25] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[26] A. Z. Pervaiz, "Real Time Face Recognition System Based on EBGM Framework," *2010 12th International Conference on Computer Modelling and Simulation*, Cambridge, pp. 262-266, 2010.

[27] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063-1074, 2003.

[28] N. Hazim, S. Sameer, W. Esam and M. Abdul, "Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance", *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 5, 2016.

[29] S. Kaushik, R. Dubey and A. Madan, “Study of Face Recognition Techniques”, *International Journal of Advanced Computer Research*, 4(17), pp.939-949, 2014.

[30] A. Singh, "Comparison of face Recognition Algorithms on Dummy Faces", *The International journal of Multimedia & Its Applications*, vol. 4, no. 4, pp. 121-135, 2012.

[31] N. Barnouti, “Face Recognition Using Eigen-Face Implemented on DSP Processor”, *International Journal of Engineering Research and General Science*, vol. 4(2), pp.107-113, 2016.

[32] H. Rady, “Face Recognition using Principle Component Analysis with Different Distance Classifiers”, *International Journal of Computer Science and Network Security*, vol.11, no.10, pp.134-143, October 2011.

[33] E. Alpaydın, Introduction to Machine Learning, Second Edition, the MIT Press, Cambridge, Massachusetts, London, England, 2010.

[34] P. Kamencay, T. Trnovszky, M. Benco, R. Hudec, P. Sykora and A. Satnik, "Accurate wild animal recognition using PCA, LDA and LBPH", *2016 ELEKTRO*, Strbske Pleso, pp. 62-67, 2016.

[35] K. Meena and A. Suruliandi, “Local binary patterns and its variants for face recognition”, I*nternational Conference on Recent Trends in Information Technology*, pp. 782-786, 2011.

[36] C. Shan, S. Gong and P. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study", *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009.

[37] M. Heikkilä, M. Pietikäinen and C. Schmid, "Description of interest regions with local binary patterns", *Pattern Recognition*, vol. 42, no. 3, pp. 425-436, 2009.

[38] M. Tiwari “Gabor Based Face Recognition Using EBGM and PCA”, *International Journal of Computer Science and Information Technologies*, vol. 3 (3), pp. 4322 – 4326, 2012,

[39] D. Bolme, "Thesis on Elastic Bunch Graph Matching", Master Thesis, Colorado State University, 2003.

[40] A. Stergiou, "Elastic Bunch Graph Matching Face Recognition: Performance and Comparison with Subspace Project Methods", Master Thesis, Athens Information Technology, 2004.

[41] A. Z. Pervaiz, "Real Time Face Recognition System Based on EBGM Framework," *2010 12th International Conference on Computer Modelling and Simulation*, Cambridge, pp. 262-266, 2010.

[42] L. Wiskott, J. Fellous, N. Kuiger and C. von der Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, 1997.

[43] B. Kepenekci, "Face recognition using gabor wavelet transform", Master Thesis, The Middle East Technical University, Sep., 2001.

[44] P. Phillips, H. Wechsler, J. Huang and P. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms", *Image and Vision Computing*, vol. 16, no. 5, pp. 295-306, 1998.

[45] L. Wiskott, J. Fellous, N. Kuiger and C. von der Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, 1997.

[46] P. Phillips, H. Wechsler, J. Huang and P. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms", *Image and Vision Computing*, vol. 16, no. 5, pp. 295-306, 1998.

[47] ORL Database. Retrieved from http://www.cl.cam.ac.uk/