Lappeenranta University of Technology
School of Engineering Science
Intelligent Computing Major

Master’s thesis

Tina Chehrsimin

ENHANCED METHODS FOR SAIMAA RINGED SEAL IDENTIFICATION

Examiners: Adjunct Professor Tuomas Eerola
           Professor Heikki Kälviäinen

Supervisors: Adjunct Professor Tuomas Eerola
             Professor Heikki Kälviäinen
ABSTRACT

Lappeenranta University of Technology
School of Engineering Science
Intelligent Computing Major

Tina Chehrsimin

Enhanced methods for Saimaa ringed seal identification

Master’s thesis
2015

59 pages, 35 figures, 2 tables.

Examiners: Adjunct Professor Tuomas Eerola
Professor Heikki Kälviäinen

Keywords: Saimaa ringed seals, segmentation, identification, animal biometrics, computer vision, image processing

Computer vision techniques have opened doors for practical, reliable and much safer methods for monitoring the animal populations in comparison with traditional methods. Old methods such as capturing the animals and installing sensors on their bodies are time consuming and may cause stress to the animals and affect their behavior. In this study, an automatic image based identification algorithm to identify endangered Saimaa ringed seals is proposed. The identification algorithm consists of three main steps including image segmentation, image enhancement, and identification. The image segmentation method contains two steps: unsupervised segmentation and classification of the superpixels. The algorithm provides promising results by using a combination of different feature descriptors and classifiers. In the enhancement step, morphological operations, contrast enhancement, and color normalization are applied for segmented images to increase the performance. In the identification step, two algorithms were tested: Wild ID and Hot spotter. The identification algorithms were implemented on original images, segmented images and enhanced segmented images. The best results were obtained using Hot spotter tested with the enhanced segmented images. With a challenging data set used, 44% of the seals were correctly identified within the first match. Meanwhile, 66% of the cases, the correct seals were one of the 20 best matches.
PREFACE

I take this opportunity to express gratitude to my supervisors at Lappeenranta University of Technology Adjunct Professor Tuomas Eerola and Professor Heikki Kälviäinen for their supervision, help, and support. A special thank I give to my love for his great support, remarkable encouragement, deep love and showing endless patience. I also would like to thank my parents for their faith in me, great support, and immense patience.

October 30th, 2016

Tina Chehrsimin
## CONTENTS

1 INTRODUCTION

1.1 Background ............................................................... 7
1.2 Objectives and delimitations ........................................... 9
1.3 Structure of the thesis .................................................. 9

2 IDENTIFICATION AND COMPUTER VISION

2.1 Biometric methods ..................................................... 11
  2.1.1 Fur and feather patterns ........................................ 11
  2.1.2 Fin patterns ....................................................... 12
  2.1.3 Nose prints ....................................................... 12
  2.1.4 Retinal patterns .................................................. 12
  2.1.5 Facial recognition ................................................ 13
  2.1.6 Ear vessel patterns .............................................. 13
  2.1.7 Weaknesses of the biometric methods ......................... 14
2.2 Computer vision methods .............................................. 14
2.3 Image-based identification of animals ............................. 16
2.4 Image segmentation .................................................. 17

3 METHODS

3.1 Saimaa ringed seal identification system .......................... 20
3.2 Image segmentation ................................................... 20
  3.2.1 Unsupervised segmentation .................................... 21
  3.2.2 Feature descriptors ............................................ 25
  3.2.3 Segment classification ......................................... 26
3.3 Segmented image processing and enhancement ..................... 28
  3.3.1 Morphological operations ..................................... 29
  3.3.2 Color normalization ............................................ 31
  3.3.3 Contrast enhancement ......................................... 32
  3.3.4 Automatic image cropping .................................... 33
3.4 Identification ......................................................... 34
  3.4.1 Wild ID ........................................................... 34
  3.4.2 Hot spotter identification algorithm ......................... 35

4 EXPERIMENTS

4.1 Data ................................................................. 36
4.2 Segmentation ......................................................... 38
  4.2.1 Unsupervised segmentation .................................... 38
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CMS</td>
<td>Cumulative Match Score</td>
</tr>
<tr>
<td>CV</td>
<td>Computer Vision</td>
</tr>
<tr>
<td>gPb</td>
<td>Globalized Probability of Boundary</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphic Processing Unit</td>
</tr>
<tr>
<td>GT</td>
<td>Ground Truth</td>
</tr>
<tr>
<td>k-d</td>
<td>k-dimensional</td>
</tr>
<tr>
<td>K-NN</td>
<td>K-Nearest Neighbour classifier</td>
</tr>
<tr>
<td>LBP</td>
<td>Local Binary Patterns</td>
</tr>
<tr>
<td>LBP-HF</td>
<td>Local Binary Pattern Histogram Fourier</td>
</tr>
<tr>
<td>LPQ</td>
<td>Local Phase Quantization</td>
</tr>
<tr>
<td>MCG</td>
<td>Multiscale Combinatorial Grouping</td>
</tr>
<tr>
<td>OWT</td>
<td>Oriented Watershed Transform</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>RGB</td>
<td>Red Green Blue</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>SFTA</td>
<td>Segmentation-based Fractal Texture Analysis</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machines</td>
</tr>
<tr>
<td>UCM</td>
<td>Ultrametric Contour Map</td>
</tr>
<tr>
<td>UEF</td>
<td>University of Eastern Finland</td>
</tr>
<tr>
<td>VHF</td>
<td>Very-High-Frequency</td>
</tr>
<tr>
<td>WPI</td>
<td>Wild Photo Identification</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 Background

Saimaa Lake is the largest amongst thousands of lakes in Finland and also habitat of many different animal species which are unique in the world. Saimaa ringed seals are one of these species which only exist in this area. They have been identified as being in danger of total extinction with conservation status of critically endangered. According to the latest studies the population of ringed seals in 2013 was around 310 individuals and the pop production is something between 44 and 66 individuals per year [1]. Annual number of death is approximately between 30 and 60 individuals [1]. Finnish Ministry of the Environment is supervising the conservation program of the seal. The program aims to increase the population of the Saimaa ringed seals up to 400 until 2020 [2].

Humans represent one of the main threats for the ringed seal, and although the species has been protected for the past half-century, it is still considered as being threatened of extinction. The major contributions to the extinction are: effects of fishing on their food, climate change and reduction of the ice and snow, and fragmentation of the population structure [2]. Therefore, effective monitoring of the populations of Saimaa ringed seals is necessary to provide a comprehensive conservation program. Monitoring of breeding conditions and birth rate, causes of death, and the natural state of the breeding grounds are among the actions which are made in the annual monitoring procedure [2].

University of Eastern Finland (UEF) provides current population tracking and conservation methods of the Saimaa ringed seals. Evaluation of the premature death and enhancement of the pup survival are among the main focus areas of the protection program. In addition, individual identification methods give possibilities to have more precise estimation of the population size and distribution in the future [2].

At the moment monitoring of the seals is performed by sensors installed on the seal body. This method has some important drawbacks such as manual identification of the new pups and the need to check the sensors. Also, it can have side effects with producing regular stress by catching seals and the health of seals in general. Therefore, a need to find a new smart solution for monitoring of seal individuals is deeply felt. Other alternative solutions already exist and are tested for other species [2]. These methods can also be useful for monitoring Saimaa ringed seals.
One of these methods is Wildlife Photo Identification (WPI) which is a method for identification of the individuals and tracking the movement of the animals. The method works based on manual identification or automatic image processing utilizing the image of the animals. This method is useful for Saimaa ringed seals because of the unique patterns on their body (see Fig.1). WPI can be operated manually by experts with obvious drawbacks including slow speed and human errors. This encourages developing of automatic techniques for WPI using computer vision methods.

Figure 1. Example of a Saimaa ringed seal image.

One of the first efforts to provide computer vision techniques to identify the ringed seals was made in [3, 4]. In the research an automatic image-based algorithm for identification of Saimaa ringed seals was developed. Their algorithm contains two main steps of the image segmentation and identification. In the segmentation, the seal body is segmented from the background and prepared for extraction of the features describing the patterns on the seal body. In the algorithm individuals are identified based on the features in the identification step. The main idea was to introduce a computer vision method in order to improve the identification of the seal individuals by partial or complete substitution of the time-consuming and error-prone manual works.


1.2 Objectives and delimitations

This study continues the previous study by further developing the identification method for identifying the individual ringed seals. The segmentation method proposed in [3, 4] is used as a starting point for the research and the main focus of the project is the identification part. Moreover, the segmentation part is enhanced by applying appropriate methods to process the segmented seal images, such as morphological operations, color normalization, and contrast enhancement.

The main targets of this thesis are as follows:

1. To improve the time of the segmentation computation to use the original size of the images instead of downsampled images.
2. To apply pre-processing methods to enhance the segmentation algorithm.
3. To improve the identification part by testing different algorithms.
4. To evaluate the performance of the selected methods based on the larger data set established in this thesis.
5. To develop an automatic image-based algorithm for the identification of Saimaa ringed seals.

There are several limitations. The data set used in this study includes only Saimaa ringed seals images and the other species of the seal is not considered. However, the developed method might be useful for identification of other ringed seals. Other delimitation is that only the image processing aspects of the problem are considered and the issues such as data collection, management, and biological aspects are left out. Moreover, images are assumed to have only one seal per image.

1.3 Structure of the thesis

The rest of the thesis includes: Section 2 contains related work. Section 3 starts with a short explanation of the proposed method of the segmentation algorithm and continues with the pre-processing methods to apply on segmented images. The section also considers the identification methods, showing the main approaches that can be applied to the segmented images. Section 4 contains the description of the data used to produce the
results and the experiments. Section 5 discusses the results, potential weaknesses, and possible development in the future. Finally, Section 6 summarizes the main features of the study.
2 IDENTIFICATION AND COMPUTER VISION

2.1 Biometric methods

The word biometrics originally constructed by two Greek words, bios (life) and metron (measure). For decades biometrics methods have been utilized to measure the biological data such as human and animal populations. Today this word is mainly referred to non-invasive methods for identification of animal individuals [5].

Traditionally, marking animal body was one of the most common methods for identifying the animals. Due to the stress related to the capturing, handling, and restraining the animal caused by this method, biometrics methods have been developed. Biometric methods have been utilized to identify both humans and the animals according to physical and behavioral attributes [5].

An animal biometric trait is a measurable identifier which can be used to identify the animals based on distinctive physical trait. This trait should be simply delivered to a sensor and transformed to a quantitative format and it must be consistent over the time. Biometric methods do not cause pain and do not make any changes on the animal body [5]. In the following subsections common examples of biometric identification methods are presented.

2.1.1 Fur and feather patterns

Some species of animals have some unique external characteristic which makes the identification of individuals easier. These characteristics include ringed patterns on snakes, stripes patterns on body of zebras, spots on geese’s bellies and wings of butterflies. These marks can be photographed for identification purposes. Sometimes changing light settings or backgrounds may make it harder to identify the animals. However, modern digital cameras can decrease the error significantly. Moreover, the photos in these methods can be taken far away to minimize the stress and behavioral changes [6].
2.1.2 Fin patterns

Identification of animals such as dolphins and whales by using photos goes back to the 1970s. For example fins of dolphins can be used for photographic identification where curves, notches nicks and tears are used as a unique characteristic. Whales also can be identified by the unique callosity patterns on their heads [7]. Different patterns on dorsal fins are shown in Figure 2.

![Figure 2. Examples of unique patterns of dorsal fins used for identification. [8]](image)

2.1.3 Nose prints

Another biometric method for identification of animal is nose prints. In this method, ink is smeared on the nose of animal and the resulting pattern is reflected on the paper. Nose printing is the same as the fingerprint identification method for humans. This method is relativity cheap and avoids making pain related to old methods including ear tags, branding and, tattooing. Some factors affect the performance such as type of the paper and ink and the force applied on the animal’s nose for printing [9]. Figure 3 shows bovine nose prints.

2.1.4 Retinal patterns

Vascular pattern on the eyes of animal is a unique characteristic which can be used for identification of individuals. It is based on the unique patterns created by vessels on retina of eyes which are recognizable from birth and it is consistent for whole life of the animal. Computer software and high performance scanners provide fast, accurate, and reliable identification [10]. In Figure 4, retinal patterns are presented.
2.1.5 Facial recognition

Facial recognition has already been used as an identification method for sheep. This method was adapted from a study about the independent-components algorithm for human face recognition. However, it is difficult to design accurate face recognition instruments despite the fact that this technique has been usually used since long time ago by humans [11].

2.1.6 Ear vessel patterns

The patterns made by blood vessels of the ears can also be considered as a unique characteristic to identify the animals. This identification method has been inspired by fingerprint
identification of humans. In this technology, a detailed, high-contrast picture of the blood vessels is obtained by a special back lighting that allows capturing the unique pattern [6]. An example of ear vessels is presented in Figure 5.

![Example of vessel patterns on a rabbit ear. [12]](image)

**Figure 5.** Example of vessel patterns on a rabbit ear. [12]

### 2.1.7 Weaknesses of the biometric methods

Despite the fact that biometric methods can provide non-invasive approaches for identification of the animals, all of them are not practical for the wildlife. Some of the methods require a near contact which makes it difficult for wild animals in the nature. For this reason, alternative approaches such as computer vision techniques can make it possible to have easier identification [5]. The next section describes how modern computer vision techniques can introduce safer and easier identification solutions.

### 2.2 Computer vision methods

Computer vision techniques aim to understand and analyze the scenes in the real world by using images and videos. There are many tasks that are solved by using computer vision methods. Someone might want to know what is the distance between a camera and building whether a car is driven in the middle of its lane, how many people are in an image, or even someone wants to know who or what is in an image. All of these
questions can be answered based on images and recorded videos by implementation of computer vision methods [13].

Today computer vision is utilized in a vast variety of applications including optical character recognition, machine inspection, retail, construction of 3D model, automotive safety, motion tracking, surveillance, fingerprint and biometrics, and medical industry [14].

The structure of the computer vision system is very dependent on the application. Some of the computer vision systems solve only a single problem, while others are designed to perform more sophisticated tasks requiring sub systems. A computer vision system is also dependent on its functions which can be either pre-specified or partially learned or transformed during the activity. Many computer vision systems typically contains steps including [15]:

1. **Image acquisition**: A digital image is taken by different kind of cameras including light-sensitive cameras, range sensors, tomography devices, radar, and ultra-sonic cameras. Type of images can vary based on the sensor of cameras including an ordinary 2D image, a 3D volume, or an image sequence.

2. **Pre-processing**: Usually in order to obtain some certain information it is necessary to process the image data before applying a computer vision method. These processes are including noise reduction, contrast enhancement, and re-sampling.

3. **Feature extraction**: There are different image features with various levels of complexity. Lines, edges, and ridges are examples of simple features. Corners and points are in another category called localized interest points. Motion, shape, and textures are examples of more complex features which are used to describe the images.

4. **Detection or segmentation**: Sometimes a decision has to be made in order to extract the regions or points in the image which have a special meaning for further processing. Segmentation of the image regions or extraction of the interest points containing particular object of interest are examples of the decision.

5. **High-level processing**: High-level processing is usually applied to segmented images where the input is a small set of data including a set of points or image regions containing a particular object.

6. **Decision making**: In some applications final decisions are required to be made by computers, for example, pass or fail on automatic inspection applications and match or no-match in identification tasks.
2.3 Image-based identification of animals

By recent development in technology of digital cameras offering inexpensive cameras with high-quality images, alternative approach has become possible. Digital cameras make it possible to produce large image data sets of animals. Moreover, recent cameras provide recording both time and location of the images by using built in clock and a global positioning system (GPS) unit. Digital images can be collected in different ways: by tourists, scientists and professionals, local residents, and photographers [16].

Basically, extracting the useful information of image data for analyzing of animal population is dependent on location and identification of the animals in each image. In the identification process, a comparison between an image of animal and a database of images that the animal individuals are known is made. Manual comparison of individuals even for small population is very time consuming and very likely error prone and it is not practical for large populations. In these cases, computer-based methods can provide automatic solutions which is practical for both small and large populations [16].

Based on this requirement, different studies on the identification of animals ranging from zebras to whale sharks to penguins have been conducted [17, 18]. Some methods are species specific such as [19, 20, 21], while the others are species independent algorithms such as [22, 23]. Especially, the Wild-ID algorithm [23] utilizes the key points matching algorithm from [24] to discover if animals are the same in two images.

The Stripe Spotter method [21] is based on recognition of features called stripe-codes. In this method zebras are analyzed and identified according to stripe patterns on their body by using two dimensional strings of binary values. The similarity between stripe-patterns is calculated by an algorithm called edit-distance dynamic-programming. In this method queries are run by calculation of similarity to each image of a database separately and providing the top matched images. This algorithm was designed specifically for zebras and the region of the interest on the image or the body of zebra is cropped manually by a user.

The Wild-ID [23] algorithm utilizes the SIFT features and descriptors [24] to extract the patterns on the animals body. The algorithm is independent on the species and has potential to be useful for any animal with unique pattern on its body for identification purposes including Saimaa ringed seals. The details of the algorithm are described later in Section 3.4.
The hot spotter algorithm [16] is a species independent algorithm tested on variety of animals to identify the individuals of species based on the key points or hot spots on their bodies. There is software, called IBEIS which utilize the latest version of the hot spotter algorithm to identify the animals in addition to a lot of features for recording and managing the information from analyzing the population of the different species and animals.

One of the improvements in the IBEIS software is accepting the cropped images without requiring selecting the region of the interest manually by clicking two points on the image and the orientation of the region of interest (ROI) by user. This gives ability to use cropped images and keeping the whole identification process automatically. Details of the theory of the identification algorithms tested on this study are described in Section 3.4 and the results of utilizing these algorithms for ringed seal identification is presented in Section 4.

An algorithm was proposed for identification of Saimaa ringed seals in [3, 4]. The algorithm was applied on images of Saimaa ringed seals. However, the algorithm can be used for any animals with unique pattern on its body. The proposed identification algorithm consists of different steps. First, a supervised segmentation method is utilized to extract features describing the texture of seal fur. Feature descriptors used in this study are including Segmentation-based Fractal Texture Analysis (SFTA), Local Binary Patterns (LBP), and Local Phase Quantization (LPQ). The features are classified with well-known classifiers to determine that the image (seals) belongs to which specific class (seal individuals). Principal Component Analysis (PCA) was applied to the feature descriptors with long feature vectors to reduce the size of the features. The reason is that not all classifiers work efficiently with a large set of features. In this method, the best matches are given to experts to make the final decision with a limited number of seals. The best result was obtained by utilizing the SFTA features transformed by PCA and the Naive Bayes classifier.

2.4 Image segmentation

Before the identification, it is often useful to perform image segmentation for easier extraction of the useful information from the image. Image segmentation is the problem of dividing a given image into smaller regions. Image segmentation is known as one of the most crucial topics in computer vision and its main goal is to change the representation of the image to something more meaningful which makes further analysis on the image easier [25]. Segmentation can be described as grouping of photographic information, where
the data are grouped into objects and objects into classes of objects [26].

Although the goal of the segmentation is clear, there is not a single method which works perfectly for all images since each image has its own certain characteristic. Different approaches advanced to enhance the segmentation task. Three algorithms have shown to be popular practically according to their performance and running time. These methods are including the mean shift method [27], the normalized cuts method [28] and the graph based method [29].

A segmentation algorithm proposed by [30, 31] has recently gained impressive results on commonly available image databases. The method is called Globalized Probability Boundary (gPb) and begins with optimized local edge extraction using learning techniques. Then, edge extraction results are globalized using normalized cuts in a spectral partitioning procedure. The purpose of this globalization procedure is to focus on the most salient edges in the image. The globalization is an effective method, but it requires solving of a large, sparse eigensystem which takes a lot of time and computational effort.

There have been efforts to speed up the gPb method [32]. In this method, the gPb method is effectively accelerated by paralleling on a Graphic Processing Unit (GPU). This is a fast method but needs to use a powerful GPU. Another research was conducted by Taylor [33] to make gPb segmentation method more efficient by utilizing structure of the underlying problem which can be implemented by a various computational platforms. In this research, the overall approach was inspired by the work on gPb [30], begins with a set of edges derived from local pixel differences and then a globalization method is implemented according to the evaluation of the eigenvectors provided by a normalized cuts method. The process includes three main steps: an edge extraction phase, a normalized cuts phase, and a region merging phase. The method utilizes a simple watershed over segmentation technique to reduce the size of the eigenvector problem to accelerate the procedure. However, the accuracy of the segmentation is reduced. In another study [34], called Multiscale Combinatorial Grouping (MCG) method, similar approach was taken to solve the eigenvector problem in a reduced space by using a simple image-pyramid operation on the affinity matrix. In this method a major speed improvement is obtained without scarifying the performance.

Another segmentation method which can be useful to segment the seals from the background is called semantic segmentation [35]. Semantic segmentation is considered as a task of partitioning an image in to regions or segments which determines objects with specific meanings. In this method the regions are labeled with an object category. The
objects can be anything like animals, humans or plants. The method constructs intensity edge between one object and the other one. The target of semantic segmentation is recovering the regions to the objects directly and also labeling them with relevant object category. The task is usually implemented as supervised or semi-supervised setting, by utilizing a manually labeled and segmented set of training images. Then, the learning algorithm identifies related image features which help to segment the regions corresponding to different classes in the unknown test images.
3 METHODS

3.1 Saimaa ringed seal identification system

In order to obtain useful information about the Saimaa ringed seal population for the conservation purposes, different techniques have been used so far. The techniques which have been used to identify the Saimaa ringed seals and to track their movements are including audiovisual tracking, GPS telemetry, very-high-frequency (VHF) transmitters in [36]. However, it is required to find efficient, non-invasive, and long-term methods which can be used to monitor the whole population.

As it was mentioned before, computer vision techniques can provide promising identification systems of wild animals avoiding significant effects on their daily life and overall health. The first efforts to construct the computer vision identification system for Saimaa ringed seals were made in [3, 4]. The proposed identification system contains two main steps: image segmentation and identification. In image segmentation the seal body is extracted from background to make the identification step easier. After segmentation, identification of the seal individuals is implemented using the algorithm utilizing feature descriptors and well-known classifiers.

This master thesis aims to develop an identification system for Saimaa ringed seals inspired by the previous work [3, 4]. There are three major steps in the identification system implemented in the thesis: image segmentation, preprocessing, and identification. The image segmentation method is based on [3, 4]. However, a new method in unsupervised part is used to obtain a more efficient result with higher performance. In the second step, different preprocessing methods are applied for the first time in identification of Saimaa ringed seals to enhance the quality of images for obtaining high performance result in the identification step. In the final step, different publicly available identification algorithms are applied to study the potentials to improve the identification methods proposed in [3, 4]. Figure 6 shows the whole proposed Saimaa ringed seal identification system.

3.2 Image segmentation

In this study, segmentation is the first part of the system. Image segmentation is used to extract the seal body in the image from the background for further analysis later in the identification. The method consists of two major steps: 1) unsupervised segmentation
and 2) classification of the superpixels. In unsupervised segmentation, the image of seal is partitioned to several segments called superpixels. In the second step, all superpixels are classified to two classes of the seal and the background based on the information extracted by feature descriptors. The superpixels classified as seal construct the whole seal body as a single object in the image separated from the background. The segmented image constructed in this step is used for further analysis for identification later.

Figure 7 shows the main steps of the segmentation method, based on Saimaa ringed seal segmentation proposed in [3, 4]. However, unsupervised segmentation has been modified to enhance the computational efficiency and segmentation performance.

### 3.2.1 Unsupervised segmentation

The method used in this step is called hierarchical segmentation. Hierarchical organization is one of the main characteristics of human segmentation. The human’s eyes segment an image by marking the boundaries of the objects until a certain level of detail by identifying them. If different subjects segmenting a single image perceive the same objects then
Figure 7. Segmentation method: 1) Unsupervised segmentation, 2) Training, 3) Classification. [3]

up to changes in the exact localization of boundaries the intersection of their segments define the finest level of the detail [26]. An example of human segmentation is shown in Figure 8.

The unsupervised segmentation method used for Saimaa ringed seal images in [3, 4] is called Globalized Probability of Boundary, oriented Watershed Transform, Ultrametric Contour Map (gPb-owt-ucm) [30] which has been shown to provide the state-of-the-art accuracy with the Berkeley Segmentation Benchmark database [38]. The method contains three main parts. First, gradients of brightness, color, and texture are computed. By
spectral graph partitioning the local contour cues are globalized which constructs the gPb contour detector. According to the average gPb strength adjacent regions are merged to implement the hierarchical segmentation. A tree of regions are produced at several levels of homogeneity in texture, brightness, color and the boundary strength of its ultrametric contour map (UCM) can be interpreted as a measure of contrast. The map determines a duality between closed, nonself-intersecting contours and a hierarchy of regions utilized for showing uncertainty about segmentation.

The gPb edge detector which is the first step in the proposed unsupervised hierarchical image segmentation has one main drawback. The gPb method requires solving of a large, sparse eigensystem which takes a lot of time and computational effort [34]. This problem limits the size of the image which can be used for segmentation resulting in a lower quality of segmented image. There have been studies with the motivation of speeding up the gPb method. For this step, an approach to multiscale hierarchical segmentation called Multiscale Combinatorial Grouping (MCG) [34] was selected which requires less computational effort in comparison with gPb edge detector.

The problem of contour detection and hierarchical image segmentation can be unified by introducing UCM [34]. UCM is an image constructed by weighting of the boundary of each two adjacent segments in the hierarchy with the index at which they are merged. For example, considering an image segmentation which divides the image into different segments. A segmentation hierarchy is a group of family of the partitions such as \( \{ S^*, S^1, ..., S^L \} \) where \( S^* \) and \( S^L \) represent the smallest set of superpixels and the whole domain respectively. In this case large level regions are unions of regions from smaller levels. A hierarchy where each level \( S_i \) is assigned a real-valued index \( \lambda_i \) can be shown by a tree of regions. In this method segmentation is constructed by giving a threshold at levels \( \lambda_i \) [34].

In order to reduce the computational cost of the gPb detectors which limits the scalability, in the MCG method an efficient normalized cuts algorithm was introduced. Practically this method keeps full performance of the contour detection with 20 times faster speed and lower memory requirements. In the MCG method UCMs are computed in three main parts including single-scale segmentation, hierarchy alignment, and multiscale hierarchy [34]. These three methods are described as follows:

In single scale segmentation local contour cues are constructed based on 1) brightness, color, and texture differences in half-disks of three sizes [39], (2) sparse coding on patches [40], and 3) structured forest contours [41]. Then, mean contour based strength UCM
is computed by globalizing the contour cues by implementing fast eigenvector gradients which combines global and local contours. Final hierarchies are evaluated instead of open contours by trying learning weights on the training set as explained in [30].

In the second step the single scale segmentation described before is applied on a multiresolution pyramid with N scales. This multiresolution pyramid is created by subsampling and supersampling the original image. In order to keep details, the smallest superpixels in the highest-resolution are declared as a set of possible boundary locations. Then the smallest superpixels of each hierarchy are extracted and re-scaled to the original image size. Finally, their sequential projections are pre-computed and transferred recursively the strength of all the larger UCMs [34].

In the final step, UCM strengths are combined together. After alignment which provides set of boundary locations and strength for each of them coming from different scale this is formulated as a binary boundary classification problem and the classifiers are trained to combine N features into a single probability of boundary estimation and uniform weights are transform into probabilities [34]. The main steps of multiscale hierarchy segmentation are shown in Figure 9.

![Figure 9. Multiscale segmentation hierarchy. Starting from a multiresolution image pyramid, hierarchical segmentation is performed at each scale independently. The multiple hierarchies are aligned and combined into a single multiscale segmentation hierarchy. [34](image)](image)

In the final step of unsupervised segmentation, superpixels are generated by giving a threshold to UCM computed in the previous step. In this step the main challenge is to find a value of a threshold giving more uniform superpixels without making superpixels
too small. The selection of the optimum threshold value for UCM is explained in detail in Section 4.

3.2.2 Feature descriptors

In next step of the segmentation, features describing the image segments are utilized. These features are used to train the superpixel classifier (Step 2 in Figure 7) and to classify the segments in the new images (Step 3 in Figure 7). The segment description and labeling is utilized to construct an automatic supervised segmentation system which makes it possible to identify the regions of the image containing parts of the seal body and combine them together to construct the segmented seal body.

To describe the superpixels, various features were tested including 1) mean RGB colors, 2) mean variance of local range, local standard deviation, and local entropy, 3) Local Binary Pattern Histogram Fourier (LBP-HF) descriptor [42], Local Phase Quantization (LPQ) descriptor [43]. In Table 1 all the feature descriptors are shown with brief description and length of their vectors.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean colors</td>
<td>Mean intensity of R, G, B color channels.</td>
<td>3</td>
</tr>
<tr>
<td>Mean Variance</td>
<td>Mean of local range of an image, local standard deviation of an image and local entropy of a gray scale image.</td>
<td>3</td>
</tr>
<tr>
<td>LBP HF</td>
<td>Discrete Fourier transforms of local binary pattern (LBP) histograms.</td>
<td>38</td>
</tr>
<tr>
<td>LPQ</td>
<td>Local Phase Quantization.</td>
<td>256</td>
</tr>
</tbody>
</table>

Local Binary Pattern Histogram Fourier

Local Binary Pattern Histogram Fourier (LBP-HF) is an effective and powerful descriptor for texture analysis. LBP-HF developed by [44]. It is a modified version of the ordinary Local Binary Pattern. In this descriptor, pixels of the image are labeled by thresholding their neighborhood with the center value as a binary number. The histogram of the labels is used as a texture descriptor [45]. LBP-HF is computed from discrete Fourier transforms of the local binary pattern histogram. One of the main advantages of this descriptor is
being invariant to rotation. In addition to being rotation invariant, the LBP-HF features preserve the highly discriminative nature of LBP histograms [42].

**Local Phase Quantization**

Another texture descriptor utilized to describe the superpixels for training the classifiers is called Local Phase Quantization (LPQ) [43]. LPQ is a blur insensitive texture classification method based on quantized phase of the discrete Fourier transform (DFT) computed in local image windows. LPQ operator is insensitive to centrally symmetric blur, including motion, out of focus, and atmospheric turbulence blur. The LPQ operator is implemented for identification of the texture by local computation at location of pixels.

**Principal Component Analysis**

One of the factors affecting the performance of the classifiers is the length of the feature descriptors. Principal Component Analysis (PCA) as a non-parametric method is utilized to make the complex data set simpler and computational task easier by extraction of relevant information from data sets. PCA by reducing dimension of a complex data set minimizes the effort to solve the task [46]. Dimensionality of the large data sets is reduced by a vector space transform. The original data set may have many variables which can be represented in just a few variables called principal components by a mathematical projection. Reducing dimension of dataset is useful allowing the user to spot trends, patterns and outliers in the data much easier than without performing PCA [47]. In this study, feature descriptors and their combinations make quite long vectors which might have negative effect on the performance of the classification. For this reason the PCA which is a popular method in dimensionality reduction of the feature descriptors was tested to study how reduction in dimensions of the features affects the performance of the classifiers.

**3.2.3 Segment classification**

In the final step of the proposed segmentation method, superpixels are classified in two classes: the background and the seal body to construct the segmented image. For the classification of the superpixels, three widely known classifiers were selected: Naive Bayes (NB) [48], k-nearest neighbor (K-NN) [48], and support vector machine (SVM) [48].
three classifiers have their own advantage and weaknesses. In the classification procedure classifiers are first trained by the features descriptor explained in Section 3.2.2. Depending on the nature of the descriptors each classifiers have a different performance. Some classifiers show good results with some features while failing in the other ones. In this regard, all features and combinations of them are tested with the classifiers independently. A brief description of each classifier is given below.

**Naive Bayes**

One of the simple classifiers which has been studied extensively since the 1950s is called Naive Bayes. The classifier comes from the probabilistic classifiers family. Naive Bayes is competitive with more complex classifiers such as support vector machines since Bayes theorem with strong independence assumptions between the features with appropriate preprocessing is performed. Naive Bayes in construction of the classifier makes an assumption that the value of the feature is not dependent on the other features. This means that it assumes that each features independently contributes to the probability of belonging a variable to a particular class, regardless of any possibility of relations between different features. In supervised learning, Naive Bayes classifier is trained very efficiently. Naive Bayes works quite well in many complex classification tasks despite the simplified assumption and naive design [48].

**k-nearest neighbor**

The k-Nearest Neighbors (k-NN) is a non-parametric classifier used in pattern recognition. In the algorithm, there are k nearest training examples in the feature space. Class membership is the output for the classification. The classification does not make any evaluation on the underlying joint probability density function. But K-NN utilizes the data sample to compute the density directly. The object is classified according to a major vote of its neighbors, which the object is assigned to the most common class among certain number of its closest neighbors defined by k which is a positive integer. If k = 1, it means that the object is assigned to its single closest neighbor class [48].

**Support vector machine**

Another well-known classifier in machine learning is called support vector machines (SVMs). SVM is a supervised learning model meaning that labeling the data is required
to train the classifier. When the data is not labeled supervised approach is not working and an unsupervised method such as support vector clustering is required to perform the classification. In classification process, SVM creates hyperplanes to separate the data into different classes. There might be different hyperplanes which can be used to separate the data but the best hyperplane is the one that has the maximum distance between two categories of the data. Such a hyperplane is called maximum marginal hyperplane. In fact, finding the optimal hyperplane is done during the training the SVM by using pre-labeled data [48]. In Figure 10, the optimal hyperplane separating data in two classes is shown.

![Figure 10](image)

**Figure 10.** Discrimination of two classes of data with an optimum hyperplane with maximum margin. [48]

All new data with unknown classes to be classified are discriminated by the hyperplane made during the training process. Sometimes the data stated in the finite space to be classified is not separable linearly and a non-linear method is required to perform the classification. SVM can perform both linear and non-linear classification efficiently. In this case, SVM maps the original finite-dimensional space into a much higher-dimensional space making the separation easier in the new space by using kernel functions [48].

### 3.3 Segmented image processing and enhancement

Since all original images do not have the same quality, contrast, color, and the performance of the segmentation algorithm is not perfect, the identification of the ringed seals
is a difficult task by using raw image data. For these reasons, it is useful to enhance the quality of the segmented images to improve the identification performance. Four major enhancement operations applied on segmented images including morphological operations, color normalization, contrast enhancement, and automatic cropping. Next a brief explanation of each enhancement steps is given.

3.3.1 Morphological operations

Mathematical morphology plays a crucial role in nonlinear signal and image processing. They can be used for image description as well as image manipulation [49]. Binary images might have many imperfections, particularly when the binary regions are produced by thresholding and distorted by noise and texture. The image imperfections can be eliminated significantly by performing morphological image processing accounting for the form and structure of the image. Morphological operations are based on the relative ordering of the pixel values regardless of their numerical values, and therefore are useful for the processing of binary images. There are two fundamental morphological operators, called erosion and dilation. There are two more operators called opening and closing which are both constructed by a combination of erosion and dilation [50].

The simple effect of the erosion operation on the image is to wear away the boundaries of the foreground regions. This is done by defining a structuring element. As an example, the structuring element can be a $3 \times 3$ square shown in Figure 11. The central pixel or element is fully surrounded by 8 pixels. In the erosion operation, all pixels of foreground which are not surrounded fully by 8 pixels are removed. Therefore, foreground regions become smaller and the holes within those areas become wider [50]. Figure 12 illustrates the erosion operation using $3 \times 3$ square structuring elements.

Dilation affects oppositely in comparison with erosion by adding a layer of pixels to both the inner and outer boundaries of the image regions [50]. Figure 13 shows how a dilation operators use $3 \times 3$ square structuring elements to make the region and gaps between different segments smaller, and fill small intrusions into the boundaries of a region.

The opening operation by a structuring element is equivalent to erosion followed by dilation. It widens the gaps between regions connected by a thin layer of pixels. By the dilation any segments that have remained from the erosion are returned back to their previous size [50]. Figure 14 presents effect of opening using a $3 \times 3$ square structuring element.
The closing operation by a structuring element is equivalent to a dilation followed by an erosion. It fills holes in the segments while keeping the primary size of the regions [50]. Figure 14 presents effect of closing using a $3 \times 3$ square structuring element.

Due to errors in the segmentation process of Saimaa ringed seals such as creating non-uniform superpixels or classification of the superpixels, some pixels are remained in the segmented image from the background or some pixels are missing from the seal body making holes in the seal body region. These errors may have some negative effects on the identification step. Opening and closing morphological operations are useful to improve the quality of the segmented images by removing unwanted segments and filling the holes. Since the seal images are not binary image, in this case the morphological operation is implemented on the mask provided by classification of the superpixels in the previous step.
3.3.2 Color normalization

Similar objects in different illumination condition appear in different colors. In order to remove the effect of the variation in illumination, causing negative effect in identification, color normalization is applied. Color normalization is concerning with artificial color vision and object recognition. Generally, color values distribution in the image is dependent on the illumination, which is affected by various factors including cameras and lighting conditions. Color normalization makes it possible to eliminate these variations for object recognition process [51].

In order to implement the color normalization on segmented images a method called histogram matching was utilized. In this method one image is taken as a reference and other images are normalized according to the reference image by matching its histogram to the reference image histogram. Histogram matching is done by defining a transformation function into input image on the input intensity values such that the transformed image
Figure 15. Effect of closing operation. [50]

output has wanted histogram. Often, the target histogram will be taken from a similar image. The histogram matching task is constructed as a derivation of a mapping \( f(x) \) between the input intensities \( x \) and the output intensities \( z \) in a way that output intensity values contains the required probability density function (PDF). The mapping \( f(x) \) is derived by the formula [50]:

\[
f(x) = C_z^{-1}[C_x]
\]  

(1)

where \( C_x \) and \( C_z \) are the cumulative distributions (CDF) of the input and reference images respectively. In Figure 16 an example of histogram matching is shown.

3.3.3 Contrast enhancement

The seal images used in this study have taken in various situations with different illumination conditions. Sometimes the contrast of the image is not good to have a clear image for extracting the pattern of the seal body for the identification. In this case contrast enhancement algorithms are used to enhance the contrast. The function which showed a good result to make the patterns more visible is called adaptive histogram equalization.

Enhancing the contrast of an image may cause over-saturation caused by the basic function of the histogram equalization. In order to prevent this happening it would be preferable to keep the brightness of some areas of the image constant while enhancing the contrast in other areas. In order to make this happen, different transformations need to
Figure 16. An example of histogram matching

be applied in different portion of the image to get a suitable result. The Contrast-Limited Adaptive Histogram Equalization technique can be used for these purposes. The described algorithm analyzes different portions of the image and computes the appropriate transformations [50].

3.3.4 Automatic image cropping

Another image pre-processing operation performed in this study is called automatic cropping of segmented image. In some available identification algorithms it is required to define the ROI before starting the identification process. Automatic cropping keeps the whole process of the identification completely automatic. Otherwise, in the identification
step, the user must select the ROI for each image manually. The cropping task is made by recognizing all isolated objects in the image and generating a bounding box around the biggest one. Since usually in the segmented images the biggest segment includes the seal body, the resulting bounding box represents exactly ROI for the identification. Figure 17 shows two segmented seal images automatically cropped by a bounding box around the seal body.

![Segmented images and bounding boxes](image)

**Figure 17.** Labeling; from left to right: Segmented image, bounding box around the seal body, and cropped image.

### 3.4 Identification

Identification is the third step of the seal identification system. This section describes two identification methods to be applied to ringed seal images.

#### 3.4.1 Wild ID

The first identification method is called WILD ID [23]. There are three main steps in this method. The first step is to extract the SIFT features [24] for each image in a database. In
this step potential interest points are identified by utilizing the scale and orientation invariant difference of a Gaussian function. Then, the local image gradients are measured at the selected scale in the regions near to the each interest point. In the second step, candidate matched pairs of SIFT features are identified from two images by locating features of one image in to the other image and minimizing the Euclidean distance between descriptors of the features. In the third step, to choose reasonable matches among all potential matches RANSAC algorithm [52] is employed to obtain a geometrically consistent subset of the candidate matches.

3.4.2 Hot spotter identification algorithm

In the hot spotter method, there are two algorithms available to solve the animal identification problem. The first one is similar to Wild ID and is called one-vs-one and the second one is referred to as the one-vs-many [16].

In the one-vs-one algorithm, for every image the key points are recognized and associated descriptor is extracted. Then by making comparison between descriptors, image matches are determined. In the one-vs-one the query image matches against each database image separately and it sorts out the data base images by the similarity score to find the final rank of the result. This algorithm includes various minor modifications in comparison with Wild ID resulting in a higher performance algorithm [16].

In the second algorithm called one-vs many [16] each descriptor is matched from the query image against all descriptors from the image database. In this method, a fast, approximate nearest neighbor search data structure is utilized. The scores are generated for each image according to these matches, and finally the scores are aggregated to generate the final similarity scores for each image. The most important contribution in this algorithm in addition to introducing fast matching is a new method for scoring matches based on the Local Naive Bayes Nearest Neighbor methods [53].
4 EXPERIMENTS

In this section, the image data and the results of the segmentation and identification algorithms on ringed seal images are presented.

4.1 Data

An image database of Saimaa ringed seals including 892 images of 147 individual seals was used. Figure 18 shows sample of images from the database.

![Figure 18. Examples from the Saimaa ringed seals database. [3]](image)

However, all images were not suitable for this study. Since the images are taken in different situations and lighting conditions with different cameras from various directions the quality of images varies significantly. Images with objects covering some part of the seal body, images with very low and high contrast are examples of such a low quality images. Preliminary experiments with such images showed poor performance because of the difficult extraction of required features. So, for further work 591 images of 131 individuals
were selected. Examples of high and low quality images are shown in Figure 19.

![Image of high and low quality images]

**Figure 19.** Examples of image quality: (a) Covered; (b) Low contrast; (c) High brightness; (d) Good quality.

**Manual segmentation**

In order to evaluate the segmentation method and to find the best threshold value for the UCM maps, the segmentation ground truth (GT) of the seal images was generated manually by the author. In these images the seals were segmented from the background and the contour around the seal body was created. In total 280 images of 40 individual seals were segmented manually by the author. Figure 20 shows an example of the seal image GT.

![Image of ground truth creation]

**Figure 20.** Ground truth creation: (a) Seal image; (b) Contour; (c) Mask.
Manual labeling of superpixels

The classification of the superpixels requires training data to prepare a classifier for the superpixel classification. To provide the training data superpixels created by thresholding the UCM were labeled manually by the author. Superpixels were labeled into two classes: the seal segment and the background segment. Manual labeling was performed on all 591 images in the database.

4.2 Segmentation

The segmentation method contains three main steps: unsupervised segmentation, training the classifier, and classification of the superpixels.

4.2.1 Unsupervised segmentation

Deciding which threshold is optimal for all images is a challenging task since larger superpixels are more probably nonuniform which means that they are not completely inside or outside the seal body. These nonuniform superpixels cause error in segmentation. So it is crucial to find a threshold which gives superpixels as large as possible while at the same time most of them being uniform. In this case, selection of the optimal value for the threshold is required.

Figure 21 shows the results of percentage of uniform superpixels, mean and median size of the superpixels, and the number of superpixels greater than 50, 100 and 200 pixels for threshold values between 0.1 to 9. The UCM computation was performed for all 591 images in the database and the UCM computation took around 300 to 500 seconds per image depending on the size of the image. The percentage of uniform superpixels was calculated by comparing the superpixels and the GT.

Figure 21a shows that by increasing the threshold the number of the uniform superpixels is decreased. Figure 21b-21f show that by increasing the threshold the size of the superpixels is increased. Based on the results threshold value of 0.7 was selected.
Figure 21. Threshold value selection for the UCM: (a) Percentage of uniform superpixels; (b) Mean size of superpixels; (c) Median size of superpixels; (d) Number of superpixels greater than 50 pixels; (e) Number of superpixels greater than 100 pixels; (f) Number of superpixels greater than 200 pixels.
4.2.2 Feature extraction and classification

The next step was to find the features and the classifier that shows highest classification performance of the superpixels. In this regard, different types of classifiers, and different feature descriptors were tested on the ringed seal images. In Table 2, brief description of all the feature descriptors and the length of their vectors are presented.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean colors</td>
<td>Mean intensity of R, G, B color channels</td>
<td>3</td>
</tr>
<tr>
<td>Mean Variance</td>
<td>Mean of local range of an image, local standard deviation of an image and local entropy of a gray scale image.</td>
<td>3</td>
</tr>
<tr>
<td>LBP</td>
<td>Discrete Fourier transforms of local binary pattern (LBP) histograms</td>
<td>38</td>
</tr>
<tr>
<td>LPQ</td>
<td>Local Binary Pattern (LBP) operator</td>
<td>256</td>
</tr>
<tr>
<td>LPQ, LBP</td>
<td>Combination of LBP and LPQ</td>
<td>294</td>
</tr>
<tr>
<td>LPQ, LBP, Mean color</td>
<td>Combination of LPQ,LBP and Mean color</td>
<td>297</td>
</tr>
<tr>
<td>LPQ, LBP, Mean color, Mean Variance</td>
<td>Combination of LPQ,LBP and Mean color and Mean Variance</td>
<td>300</td>
</tr>
</tbody>
</table>

For the experiments 11 sets of 20 images from 591 images of the whole database were selected randomly, and 571 remaining images in each experiment were used for the training data. The accuracy of the segmentation with different features and classifiers were computed. The graphs obtained from this calculation are shown in Figures 22-24. The performances of the segmentation were calculated according to a number of pixels which were correctly classified based on the GT. The results of the segmentation performance are shown based on the percentage of the images in which specific amount of pixels were classified correctly. This specific amount varies between 0% to 100% of pixels in the image which is called threshold in the figures.

As it can be seen from the results, classifiers behave differently against different features. For the K-NN classifier the best result obtained using a combination of the LPQ and LBP features. The mean color feature descriptor shows significantly better performance for the Naive Bayes classifier. However, the best performance with the SVM classifier was obtained using combination of the LPQ, LBP, mean color, and mean variance features. This combination outperformed the other classifiers and combinations and was selected for the identification step. The performance of the segmentation using the different classifiers with their best feature descriptors are shown in Figure 25.
Figure 22. Segmentation performance of the Naive Bayes classifier.

Figure 23. Segmentation performance of the K-NN classifier.
4.2.3 Principal components analysis

According to the obtained results, it is obvious that K-NN and Naive Bayes suffer more than SVM from large dimensionality of the features. To solve this issue the PCA was performed on features with large dimensionality in order to make their dimensions smaller and see how it affects the performance of the classifiers.

With a smaller numbers of dimensions there is a risk of losing information and with a larger numbers of dimensions the benefits of using PCA are lost. The output dimensions for one set of features for different classifiers reduced by PCA. Figures 26-28 show the segmentation performance with the transformed features. For each classifier, the best features with three number of output dimensions: 50, 100 and, 200 were tested and the performances were calculated. The results show that PCA does not affect the K-NN classifier at all and the performance remains the same while the number of output dimension is changed. With SVM, a higher number of dimensions gives better result and with the Naive Bayes a smaller number of dimensions gives better performance.
4.3 Segmented image processing and enhancement

4.3.1 Morphological operations

Two morphological operations were implemented on a binary mask obtained in the segmentation step to improve the segmentation results. The first operation is called closing operation. Examples of imperfect segmentation results before and after the operations are shown in Figure 29.

The other morphological operation which was implemented on the binary masks called opening which removes small isolated segmented parts in the image. This helps to remove some small incorrect region in the segmented images causing errors in the feature extraction in the identification step. Examples are shown in Figure 30.

4.3.2 Color normalization

Another enhancement operation performed on segmented images was the color normalization. An example of color normalization is shown in Figure 31.
Figure 26. Segmentation performance of the LPQ-LBP-Color features with the SVM classifier with different number of PCA dimensions.

Figure 27. Segmentation performance of the LPQ-LBP-Color features with the Naive Bayes classifier with different number of PCA dimensions.
4.3.3 Contrast enhancement

In the final step, contrast enhancement was applied on segmented images. Figure 32 shows an example of applying the contrast-limited adaptive histogram equalization.

4.4 Identification

In this section, the results of the identification methods are presented.

4.4.1 Hot spotter

The software provided by the authors of the Hot Spotter method was utilized for the identification of Saimaa ringed seals. The first step was to import the images into the software. Then the ROI were determined manually by the operator by clicking the mouse on two points of each image. It is important in which direction these clicks are made which determines the orientation of the object. The orientation of the region of the interests must
be equal on all images. For example, if the orientation is from the head of the seal to the
tail of the seal in one image in the other images it must be the same. Figure 33 shows
selecting the region of the interest and their correct orientation.

After selecting the ROIs for all images, one of the images is selected to be identified. This
image is called the query image. After running the query the results of comparing that
image with remaining image are shown according to the score of the similarity. One of
the examples of correct matching and wrong matching are presented in Figure 34. As it
can be seen correct matching, significantly more hot spots have been identified resulting
in higher score, but in incorrect matching example just a few hot spots have been detected.

The IBEIS software uses the same algorithm as the Hot spotter, but it is more advanced
in database management. One benefit of this software a compare to the hot spotter is that
it is not necessary to select ROI manually by the user. The whole image can be selected
as ROI which enables to use the automatically cropped images. Thus, all the steps of the

Figure 29. Examples of filling holes on the seal body by closing operation.
identification process are automatic.

4.4.2 Identification performance

Hot spotter

Performance of the identification algorithms were tested using different data sets. Original images, segmented images and enhanced cropped images were used and cumulative match score (CMS) against the rank of the correct seal id was calculated. In this step, 194 images were selected from all 220 images which were segmented in the segmentation step. The selection made manually based on the quality of the segmented images to make sure that most of the bodies of the seals have been remained after segmentation and at least there are two images representing one individual. The produced CMS histogram shows accumulative accuracy of the identification system. Nth bin in the histogram tells a percentage of images for which the correct match within the N best match proposed by the identification method. Figure 35 presents the CMS histogram of the Hot spotter identification performance.

This figure shows that enhanced cropped images gives better results as compare to the original image and segmented images as it was expected. 44% of the seals were correctly identified. Meanwhile, with the cropped enhanced images 66% of the cases, the correct seal was one of the 20 best matches and 57% and 53% for the segmented images and original images respectively.

Wild ID
The benefit of Wild ID method is that any images such as original, cropped or segmented images can be used. This helps the automation of the whole identification process. This software was tested for both 194 of cropped enhanced images and original images. 13% of the cropped enhanced images and 8% of the original images were correctly identified. Wild ID does not provide scores or ranks and therefore, it was not possible to produce the CMS histograms.
Figure 32. Before and after implementation of the adaptive histogram equalization.

Figure 33. Selecting the region of the interest for two images with different orientation.
Figure 34. Examples of hot spots matching in Hot spotter: (a) Correct matching; (b) Incorrect matching.

Figure 35. Identification performances of the Hot spotter algorithm.
5 DISCUSSION

5.1 Methods and results

An effort has been made in this study to improve the algorithm introduced in [3, 4] for the automatic identification of the individual ringed seals of Saimaa Lake. The ringed patterns on the seal body are unique like human fingerprints. This enables to use image processing methods to identify an individual seal in an image. The image processing algorithm utilized in this study includes three main steps: 1) image segmentation, 2) image processing and enhancement, and 3) identification.

Image segmentation is the first step which helps to isolate the seal body from the background of the image. This step is useful for identification of the seal where the main focus is on the patterns on the seal body. Segmentation of the seal avoids any disruption in extracting the features for unique patterns on the seal body. There are many different image segmentation methods available. However, none of them are perfect for all different images and situations. Segmentation methods are very dependent on the image and the objects which are going to be segmented. The method which was utilized in this study includes computation of UCM and creating superpixels by giving threshold to the UCM. One of the main challenges of the segmentation is how to compute the UCM faster with minimum computational efforts. As the computation of the UCM is an intensive computational task, it is not possible to segment a very large high quality image within a reasonable time by ordinary computation powers. In this study, the MCG algorithm was found a suitable method to compute UCM in a fast way without scarifying the quality of the image by making it smaller. In the MCG method, it is possible to segment three times larger images in comparison to the method used in [3, 4]. However, it is still impossible to compute UCM for the image with several megapixels with ordinary computers.

Another challenge of the segmentation part was creating uniform superpixels by giving a certain threshold for all images. It is impossible to find a value for the threshold to create completely uniform superpixels for all images without making superpixels very small since it is very dependent on the image. Therefore a trade-off has to be made in Section 4, 0.7 was selected for the threshold value of UCM according to comparison of different values and controlling the size and uniformity of the superpixels.

After creating superpixels, selecting the feature extraction and superpixel classification is the next step. Different feature descriptors and combinations of them were tested, and
superpixels were classified using the SVM, K-NN and Naive Bayes classifiers. The best result was with the combination of the features the LBP, LPQ, mean color and mean variance and the best classifier was SVM. 55% of the images were segmented with this combination which more than 90% of their superpixels were classified correctly. The main factors affecting the performance of the segmentation can be listed as: quality of the generated superpixels, performance of the classifiers, and the right selection of the feature descriptors and the combination of them. The optimization of all these factors together provides better segmentation result. In [3, 4], the best segmentation result was obtained with performance of 80% by using SVM. However, only 20 pre-selected images were tested which was too small data set to make reliable conclusion.

After the segmentation part enhancements were made for the segmented images to make the identification part more accurate and easier. These enhancements were morphological operations, color normalization, contrast enhancement, and automatic cropping of the image. Morphological operations include opening and closing. Both these operations improve the segmented image quality by filling holes and removing unwanted objects. Color normalization and contrast enhancement make the ringed patterns on the seal body more visible and make the feature extraction on identification part easier.

The Identification part was accomplished by testing different software for photo identification. These software included Wild ID and IBEIS. Both of these software follow almost the same procedure in identification of the individual animals. Wild ID uses the SIFT features and is completely automatic. However, it does not provide the rank of the correct seal ID. This software was tested for 194 images which 13% of cropped enhanced images and 8% of original images were correctly identified. The second software was IBEIS which is the advanced version of the Hot spotter. In this software is not necessary to select the region of the interest manually. So any images can be used as an input. IBEIS shows better results by using the segmented and cropped images. This software gives the best result by using enhanced cropped images. 44% of the enhanced cropped images were correctly identified while this value was 40% and 34% for the segmented images and the original images respectively.

5.2 Future research

The identification system proposed in this study can be further developed in two main areas. In segmentation, many different segmentation methods are available and none of them can be perfect for all kind of images. Other methods may give a better result for
segmentation of the ringed seals if they can offer faster methods to alternate UCM computation since it requires considerable computation effort. Superpixel classification and selection of the best threshold value are still challenging. It can be improved by finding better classifiers and feature descriptors as well as a method to find a more comprehensive method, instead of giving a threshold value to UCM. Identification part should be improved. The proposed method used SIFT descriptor or hot spots to identify the key points. By combining different features, better results might be obtained like in segmentation.
6 CONCLUSION

In this study the image-based Saimaa ringed seal identification algorithm introduced in [3, 4] has been improved. This identification algorithm is divided in three main parts including image segmentation, image processing and enhancement, and identification.

An image database of the Saimaa ringed seal was used from raw database collected by University of Eastern Finland. First, by implementing the proposed segmentation algorithm on the original image database the seal was separated from the background. Segmentation improves the identification process by reducing problems in extracting the feature descriptor of the unique patterns of Saimaa ringed seal body. Segmentation was done in two main steps. First by computing the UCM by using the MCG method to create the superpixels. The superpixels were created by giving a threshold to the UCM. Superpixels were classified in two classes of the seal and the background to create the segmented images. Three classifiers, SVM, K-NN and Naive Bayes were tested. The classifiers were trained by data obtained by manual labeling of the superpixels by the author and calculation of different feature descriptors such as LBP, LPQ, mean color, and mean variance. The best performance was obtained using combination of all these features using the SVM classifier.

In order to improve the quality of the segmented images, enhancement operations were applied on segmented images. First, unwanted segments remained in the segmented image were removed and missing segments restored by using two morphological operations, following by contrast enhancement and color normalization of the images.

In the next step two identification algorithm, Hot spotter and Wild ID were tested on the image database of Saimaa ringed seals. The identification was tested on the original images, segmented images, and cropped enhanced images. Since the both of the software for identification require the selection of the region of the interest, an algorithm was developed to crop the image automatically to keep the identification process automatic.

Hot spotter showed the better performance specially when the enhanced cropped images were used. The second best performance was for the segmented images and the last one obtained by using the original images. This study has shown that the performance of the identification was improved by applying segmentation and enhancement of the original images. However, these are not all the factors affecting the performance of the identifying the ringed seals. The quality and size of the image database have shown a great impact on the performance of the whole process.
REFERENCES


