National University of Science and Technology POLITEHNICA Bucharest

Doctoral School of Electronics, Telecommunications and Information Technology

Decision no. 303 from 17-01-2025

Ph.D. THESIS SUMMARY

Ing. Dumitru ABRUDAN

APPLICATION OF MATHEMATICAL MORPHOLOGY IN SHAPES RECOGNITION

THESIS COMMITTEE

Prof. dr. Ing. Florin DRĂGHICI UNSTPB	President
Prof. dr. Ing. Mihai CIUC UNSTPB	PhD Supervisor
Prof. dr. Ing. Romulus TEREBEȘ Technical University of Cluj	Referee
Prof. dr. Ing. Cristian GRAVA University from Oradea	Referee
Prof. dr. Ing. Ioana MARCU UNSTPB	Referee

BUCHAREST 2025

Content

Chapter 1

Introduction

Image processing is a constantly relevant research field, with particular importance in various sectors of activity, such as health, construction and agriculture. The goals of image processing include identifying and classifying shapes (eg, people, objects, structures), highlighting or hiding features of shapes in images, and minimizing noise introduced by acquisition methods or defects in image acquisition equipment.

Among the algorithms designed for recognizing shapes in images are face detection algorithms (both in static and dynamic images) and object detection algorithms (by performing contouring and smoothing operations, zooming in and out of certain areas in the processed images). The accuracy of these algorithms is crucial in the medical field (e.g., in medical imaging the correct identification of shapes allows the correct diagnosis of patients), in the field of security and protection of civilians (e.g., in certain congested perimeters, the precise detection of faces allows easier to identify people who commit crimes), etc. The accuracy of image recognition techniques can be compromised by noise introduced into images, with the undesired effect of distorting information in digital images.

In this context, the personal contribution to this research activity consists in demonstrating the usefulness and efficiency of the method based on the use of mathematical morphology operators in recognizing shapes and removing noise from images. The basic mathematical morphology operators (MMOs) are erosion and dilation and are used for shape recognition, and the combined mathematical morphology operators (opening and closing) are applied to reduce/remove noise from images.

Presentation of the field of the doctoral thesis

Digital image processing is the process by which relevant information is extracted from images using different types of methods. Thus, an algorithm designed for digital processing takes as input a digital image, and the output may be a digital image, certain features of the image/from the image, or a refined version of the input image. From the

point of view of the final objective, digital image processing procedures lead to obtaining images with significantly improved qualities. Among these techniques can be listed the *smoothing of the edges or the addition of colors*.

Image processing (regardless of the usage scenario) is dependent on the method of image acquisition, which can be done with digital cameras, scanners, or these images can be obtained by transferring them between various devices at specific target locations. The accuracy of these methods differs and, as a result, digital images can be affected by different types of noise. A general classification of these defines them as: *impulse, photoelectronic and structured noise*.

Scope of the doctoral thesis

The application of shape detection techniques (objects, people, animals, etc.) from images is of particular importance in fields such as medical imaging, air and ground traffic control, fraud detection, security (e.g. facial recognition).

Since the quality of an image to be processed depends on the noise superimposed on it during its realization, different methods of filtering this noise (linear, non-linear filtering techniques, etc.) can lead to acceptable results. Although some state-of-the-art methods primarily involve pixel-by-pixel filtering operations, upon review of existing approaches, a stand-alone solution that eliminates the requirement for visual inspection by a human operator is a must. Thus, the use of combinations of shape detection methods together with DL algorithms can represent a solution for the efficiency of the image processing process from the point of view of human interaction (by minimizing its presence), the time interval required to obtain the results (by its minimization) and, implicitly, the time interval for making decisions based on the results obtained.

Content of the doctoral thesis

The doctoral activity had as its main theme the study and practical analysis of some "Applications of Mathematical Morphology in Shape Recognition". The stages for the successful fulfillment of the objectives were focused both on the study of theoretical aspects and on the practical implementation of mathematical morphology operators in different use cases for shape recognition, feature extraction and impulse noise removal from the analyzed images.

Chapter 1 provides a brief but comprehensive introduction to image processing, focusing on shape detection. The traditional methods used (the Viola-Jones algorithm, the HOG detector and the DPM model) are mentioned, but also the modern methods (CNN networks, the YOLO algorithm and the morphology operators). The fields of practical applicability for these methods are stated (airport perimeters, conflict areas, border areas,

etc.), thus defining the rationales that were the basis of the choice of this doctoral research activity.

Chapter 2 illustrates in detail the traditional and current shape detection methods for each of the previously described methods (Viola-Jones, Aggregate Channel Features (ACF) and MMO algorithms), indicating implementation modalities, use cases and performance evaluation. The last sub-chapter highlights the experimental results obtained using the three methods for: identifying faces, eyes and masks worn by people during the COVID-19 pandemic, as well as estimating the weight of passengers on board an aircraft based on body contours to reduce fuel consumption in commercial aviation. Some conclusions related to the performances of these algorithms are found at the end of the chapter.

Since image acquisition can be affected by noise superimposed on them, techniques for removing noise from images are described in **Chapter 3**: non-linear filtering and linear filtering. Mathematical morphological operators are presented in the category of nonlinear filtering, and a comparative analysis of all image denoising techniques is highlighted in this chapter. The validation of the efficiency of using MMO for the purpose of noise filtering is obtained in a specific scenario involving the detection of defects in pine wood, and MMOs are used both for noise removal and for the identification of existing knots. The impulse noise considered in the experiments is ”salt *and pepper noise*”. Based on the experimental results, a summative analysis was carried out regarding the efficiency of morphological operators and the identification of their validation needs in an aggregated approach with modern image processing techniques (convolutional neural networks).

Chapter 4 contains the ways to validate MMO performances by merging with CNN. The main aspects related to VGG16, VGG19 and CNN-type networks algorithms are detailed: description, advantages and disadvantages depending on the fields of applicability. The last part of the chapter illustrates use cases where the effectiveness of this approach has been proven, these case studies including detection of handwritten mathematical symbols affected by noise and subsequently detection of handwritten letters, obtaining synthesized images using a model-based augmentation, correct identification of a ship model using a database, and addressing earthquake-induced structural defects in civil works. A new proprietary TRIB Crack database of images containing road infrastructure cracks is also introduced to the research. These images are processed to demonstrate the usefulness of this study in road infrastructure management.

In the last chapter, **Chapter 5**, the main results obtained together with own contributions are synthesized and a list of original works published during the doctoral research activity is presented. Finally, based on the results obtained in the doctoral research activity, the prospects for further development in the field of image processing (in a postdoctoral study program) are stated.

Chapter 2

Algorithms for detecting shapes in images

Image processing techniques can be used to identify and classify shapes detected in images to provide information that can later be used for different purposes depending on the field of applicability.

Following the studies carried out, certain conclusions related to the efficiency of the use of the three methods in the identification of different shapes in the images were stated. In addition, comparisons are made with similar research in the field and the strengths and limitations of the proposed approach are highlighted within the presented scenario.

2.1 Viola–Jones detection algorithm

The Viola-Jones algorithm is a "machine learning" (ML) algorithm originally designed to detect specific regions of the human face, such as the eyes, mouth, and nose [1]. Although this algorithm provides lower face recognition accuracy compared to recent CNN-based techniques, it still has the advantage of running efficiently on resource-constrained devices.

2.1.1. Description and functionality of the Viola-Jones algorithm

The Viola-Jones algorithm was originally developed for face detection, but - later - it was trained to help recognize - from images - different types of objects. Operating speed and great accuracy are advantages of the Viola-Jones algorithm due to the four stages of realization.

2.1.2. Use cases and algorithm performance evaluation

Several applications have been developed using the Viola-Jones algorithm, which can be successfully used for shape detection (face, objects, security systems, road safety elements, etc.).

The Viola-Jones algorithm has also proven effective in real-time shape detection. By using a modular whole-image approach, face detection in video streams has been optimized to meet the real-time processing requirements of at least 30 frames per second [2].

2.2. *Aggregate Channel Features* (ACF) detection algorithm

The ACF detection algorithm was developed to detect rather rigid objects (eg pedestrians, vehicles, faces, etc.). With 30 frames per second for a single core, the ACF detector is a fast and efficient sliding window detector. Compared to the Viola-Jones detector, it provides an approximately 1000-fold reduction in false positives while maintaining the same detection rate [3]. ACF evaluates features as pixel values, not calculating rectangular sums.

2.2.1. Ways to implement the ACF algorithm

The structure in Figure 2.11 illustrates the implementation steps of the ACF detection algorithm, and they are detailed below [3]-[5]:

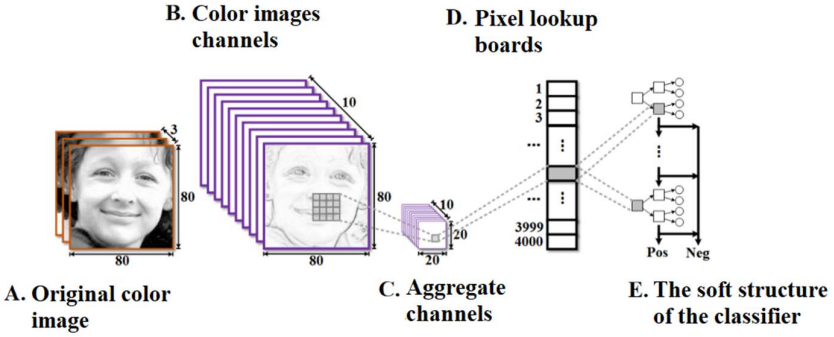

Figure 2.11 *Diagram of the ACF calculation process* [5]

2.2.2. Applicability and performance evaluation of the ACF algorithm

The applicability of ACF is varied, including use cases such as traffic light detection at night [7], vehicles and pedestrians [8],[9], aircraft detection in satellite images [10], object identification in ultrasound [11], and so on.

The detection of the human body, not only the face, is important in the case of pedestrians, people lost on the mountain or fallen for various reasons (injured, people with various medical conditions, etc.).

2.3. Mathematical morphology operators (MMO)

Mathematical morphology is a technique for optimal recognition of shapes in images, but also filtering noise from images to improve its quality. The etymology of the word morphology comes from form (*morpho* in ancient Greek) and research (*logos* in ancient Greek).

Classical morphological operations use a static structuring element (SE - Structuring Element) consistently applied to the image. This approach can be suboptimal, prompting the evolution towards adaptive morphological filtering that adjusts the structuring element for each individual point of the analyzed image. SE is a low-dimensional binary matrix that satisfies the following conditions: the size of the matrix defines the dimensions of SE; the arrangement of ones and zeroes determines its shape, and the origin is -usually- a pixel inside or outside the matrix. The most used forms of SE are illustrated in Figure 2.15 [12]-[15].

Figure 2.15 *Different models of the structuring elements* [13]

The shapes of the structuring elements can be customized to match the shapes of the processed signals.

2.3.1 Basic operations of mathematical morphology: erosion and dilation

Erosion is the mathematical morphology operator by which objects in a binary image are reduced or removed. In most applications, this operation is accompanied by the ***Dilation*** operator.

From a mathematical point of view, morphological operations can be written based on set theory, the sets being formed by all the points that describe the image: background and foreground.

2.3.2 Combined operators of mathematical morphology: opening and closing

The morphological *opening* and *closing* operators use a combination of the basic morphological operations, but in a different order, thus producing different results.

The *morphological opening* operator consists in performing the erosion process followed by the dilation process. The effect on objects is to refine their boundaries, separate narrow connections (arrowed in Figure 2.18) and remove thin extensions (circle in Figure 2.18). Conversely, the morphological closure operator consists in performing the dilation process followed by erosion (Figure 2.19). For each of the two figures, the original image is the one on the left, and the one processed with the morphology operators is the one on the right [16].

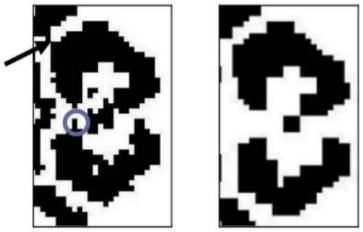

Figure 2.18 *MMO opening* [16]

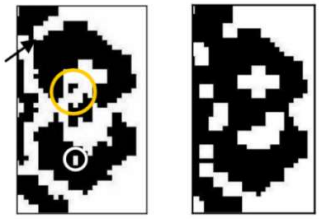

Figure 2.19 *MMO closing* [16]

2.4. Comparative analysis of object detection algorithms

Object detection refers to the precise identification of one or more specific objects in an image, objects that can be characterized by well-defined and distinct specifications. Image quality is crucial for database analysis because image capture environments can

introduce elements that can negatively affect the accuracy of object detection (shadows, fog effects, etc.). Other aspects that can negatively influence the detection in the images can be the angles of the images, various partial or total occlusion of an object by other objects, etc.

2.4.1. Shape detection using conventional methods

Figure 2.20 shows a variety of conventional shape detection algorithms such as Viola-Jones, Canny Edge Detection, Speeded Up Robust Features (SURF), HOG, and Table 2.1 provides key elements used in the case of the conventional approach in object detection.

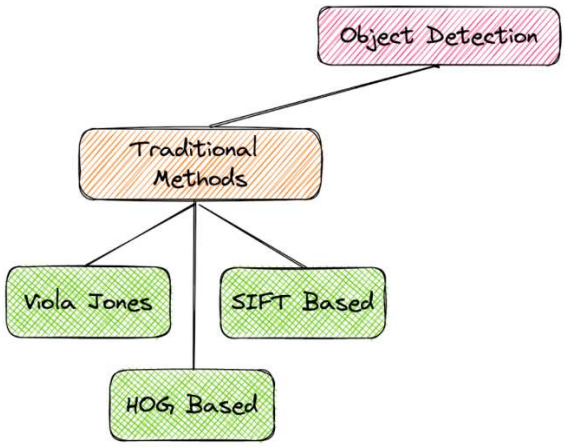

Figure 2.20 *Conventional object detection algorithms* [17]

2.5. Applications of shape detection algorithms in image analysis

Considering the outstanding results obtained in previous research with Viola-Jones, ACF and MMO algorithms, the aim of the research work presented in [PR1] was the detection of wearing a compliant mask on the face using Viola-Jones and ACF algorithms as well as MMO. While most previous studies have focused on comparing the efficiency of Viola-Jones and ACF detection algorithms, this application introduces an innovative approach that uniquely integrates three distinct techniques (Viola-Jones, ACF, and MMO), each serving a purpose precisely, yet collectively demonstrating great effectiveness.

2.5.1. Case study: facial recognition and mask wearing in a pandemic

The innovative methodology proposed in this case segmented the application requirements into distinct phases: detecting faces in one phase, checking the presence of masks on detected faces in another phase, and detecting the correct or incorrect wearing of masks in a distinct phase. This sequential approach was aimed at increasing overall efficiency and reducing errors resulting from image processing.

2.5.2. Case study: reducing fuel consumption in commercial aviation

The analysis in the field of mathematical morphology continued with another scenario, motivated by existing research indicating a strong correlation between an aircraft's fuel consumption, carbon emissions and passenger weight.

Usually, to accurately determine the weight of a passenger, an individual weighing is done as the overestimation of their weight translates into carrying extra fuel, which adds weight to the aircraft, which -in turn- consumes more fuel to support its own weight. The approach to this shortcoming is presented within the method proposed in [PR2].

2.5.3. Conclusions on proposed applications for shape detection in images

The use cases presented in this chapter are results of doctoral research activity in which only the basic mathematical morphological operators (erosion and dilation) were tested. The results obtained are promising and have been validated by the scientific community, being currently published in [PR1] and [PR2]. Based on them, new research opportunities are opened in the vast field of image processing. Thus, the conclusions resulting from the conducted research will be presented below.

Chapter 3

Techniques for filtering noise from images

Advances in the development and optimization of digital image acquisition methods have led to the emergence of numerous applications focused on digital image processing. Noise in digital images can be defined as unwanted extra data that distorts the inherent characteristics of the image. This occurs either from the sensor and circuitry of a scanner or digital camera, or from film grain. Figure 3.1 briefly presents the different types of noise in image processing [18], [19].

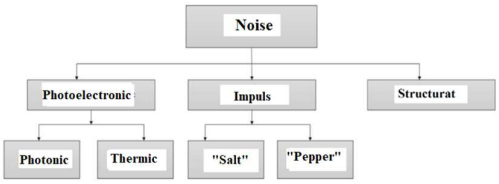

Figure 3.1 *Types of noise in image processing* [19]

3.1 Non-linear filters

Nonlinear filtering is a technique for reducing noise in images in order to shape them for specific applications. Due to the non-linear relationship between input and output signals, non-linear filtering can have high computational complexity and difficult design of filter parameters.

3.1.1. Median filtering

This type of filtering involves selecting a window (kernel) of quadratic shape, and inside the window the value of each pixel is replaced by the median value of the pixels around it. The size of the window is chosen according to the purpose: an image that preserves the fine details requires choosing a smaller window, and selecting a larger window leads to obtaining a more obvious smoothing effect.

3.1.2. Bilateral filtering

Compared to median filtering, bilateral filtering removes noise from images but keeps its edges intact. Edge preservation is because the filtering algorithm considers how different the pixels are in terms of color, but also how close they are. The calculation algorithm for obtaining a new image involves adjusting the pixel value at a given location to the average of similar pixel values nearby.

3.1.3. Filtering based on mathematical morphology

According to [20] (Figure 3.5), MM filters are used for contour detection and can also filter images by combined morphological operations. The basic MMOs described in Chapter 2 have the role of enlarging the image which means shrinking the background (through the dilation operation), respectively shrinking the image and enlarging the background (through the erosion operation). The combined operators of mathematical morphology have the roles of removing small objects to restore the original image (morphological opening) or to smooth contours, fill gaps and remove small holes (morphological closure).

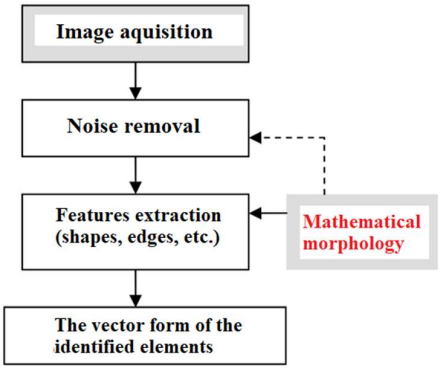

Figure 3.5 *Facilities of mathematical morphology* [20]

Optimized conventional MM filters have demonstrated their effectiveness in complex implementations, producing high-quality restored images. A fuzzy filter calculates the value of a pixel using the values of neighboring pixels, while maintaining the essential characteristics of the image. The main advantage of this filter is the use of uncertainty characteristics for successful noise removal [21], [22].

3.2. Linear filtering

Through *linear filtering*, the images are improved in appearance, having modified certain components of the frequency spectrum, a modification made with a linear filter. At the output of the filter, a combination of the input image and the weight function of the respective filter is obtained through the convolution operation.

In general, linear filters perform poorly in the presence of noise, and - as a result - soften sharp edges, eliminate lines and spoil fine image details.

3.2.1. Filtering using mediation

It belongs to the category of conventional filtering methods, and, after filtering, the effect is image *blurring* if the window chosen for filtering is small, or good noise filtering if the window chosen is large. The blurring phenomenon is evident in areas of the image where the variation of pixel intensities is sudden.

3.2.2. Filtering with the Gaussian characteristic filter

Compared to the averaging filter, the Gaussian feature filter does not use a uniform average value for all pixels of the window, but the pixel of interest (center) receives the highest weighted value, and neighboring pixels receive lower weighted average values.

3.2.3. Wiener filtering

The Wiener filter reduces Gaussian noise from a damaged image by adjusting the smoothing level based on noise levels, applying less smoothing to high-contrast regions and more to low-contrast areas, often outperforming other filters in performance and holding account for the deterioration in specific frequency bands [23], [24].

3.3. Evaluation criteria of filters in image processing

The quality of a filter can be evaluated by determining parameters such as [25]:

• **Peak Signal-to-Noise Ratio** (PSNR) is a parameter that provides a measure of the degree of error between the original image and the one reconstructed with the help of filtering methods. Mathematically it is defined in equation (3.7):

$$PSNR = 20log_{10}\left(\frac{N}{MSE}\right) \tag{3.7}$$

where N is the maximum possible value of a pixel in the original image and MSE is the mean squared error.

3.4. Practical implementation of the MM filter in image processing

3.4.1. Detection of knots in pine wood using MM and RGB filters

The application of mathematical morphology has been extended to the detection of knots in pine wood, a defect with significant implications for the use of wood in construction, including houses, beams and doors.

To address this problem, the solution developed in the research activity [PR3] uses an MM filter and an RGB filter to highlight pixels that differ from the typical color values of pine wood.

3.4.2. Conclusions on the effectiveness of the MM filter in the proposed case study

The detection of knots in pine wood involved the testing of two MMOs (erosion and expansion) in contrast to the cases presented in Chapter 2 where only one MMO was used. The results obtained are promising and have been validated by the scientific community, being currently published in [PR3]. Based on them, new research opportunities are opened in the field of image processing.

Chapter 4

Contributions to the optimization of image processing solutions by integrating MMO and CNN

To evaluate the effectiveness of the MMO fusion solution with Convolutional Neural Networks (CNN), a sophisticated and current solution in today's technological applications, a series of use cases that we designed for this purpose will be highlighted in this chapter. Along with the optimization of the proposed solutions and the thorough examination of the potential of the MMO fusion technique with CNN, personal contributions to highlight the benefits of the MMO-based filtering technique will also be presented.

4.1 Introductory aspects related to DL-type algorithms

A brief presentation of the DL method and a comparison with traditional image processing methods is given in Chapter 2. In this chapter, the DL method is presented in detail to make the transition to DNNs, CNNs and to identify the main implementation ways for image processing.

The DL method falls under the broader category of artificial intelligence (AI), being a division within machine learning (ML -Machine Learning). Essentially, while all ML and DL are forms of AI, not all AI involves ML.

4.2 Convolutional Neural Networks (CNN)

CNN is the most prominent DL technique for image processing, using convolutional layers to extract complex features, thus enabling accurate predictive capabilities. To achieve the highest accuracy in recognizing shapes and objects in images, a CNN must be pre-trained on an extensive dataset of annotated images, each containing the target

shapes or objects for detection. During the training phase, the CNN systematically calibrates its ability to map extracted features to their corresponding labels through iterative backpropagation and advanced optimization algorithms. Upon completion of training operations, the network can extrapolate predictions for unknown images by propagating them through the network and selecting the label associated with the maximum predicted probability [26].

4.3 VGG16 and VGG19: implementations, advantages and disadvantages

The classical architectures of VGG16 and VGG19 networks, together with the identification of layers, are illustrated in Figure 4.4 (VGG16), Figure 4.5 (VGG19) [27].

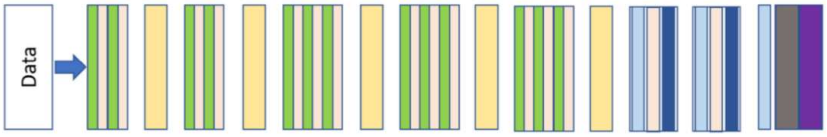

Figure 4.4 *Organization of layers in the network VGG16* [27]

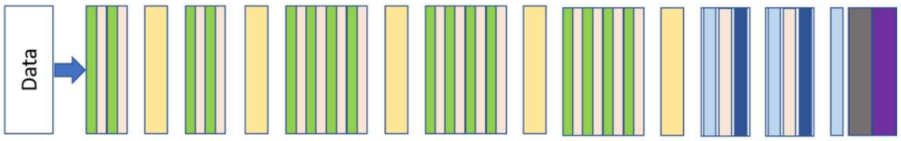

Figure 4.5 *Organization of layers in the network VGG19* [27]

Both architectures contain a significant number of convolutional units as well as repetitive layer structures (convolutional, ReLu and maximized pooling). To optimize the results, additional ReLu layers have been introduced in VGG19 that help to improve the network's ability to recognize and detect objects.

4.4 Practical fusion of MMO and CNN: case studies

4.4.1. Detection of handwritten mathematical symbols affected by noise

Mathematical symbols (MS) are universal, and accurate recognition of handwritten SMs is essential for verifying the correctness of equations that combine both

symbols and letters. This application focuses on the recognition of handwritten SM affected by noise on the whole image acquisition process. *The architecture of the proposed algorithm models a DNN network for the purpose of removing noise from a database using MM filters*. After removing the noise, it is checked whether the image processed with the help of filters can be identified and classified.

As described in Chapter 3, the noise superimposed on the images can be Gaussian noise (Figure 4.8 b)) or *"salt and pepper"* noise (Figure 4.8 c)) [PR4].

a) b) c)

Figure 4.8 *a) Original image b) Image affected by gaussian noise c) Imagine affected by „salt and pepper" noise* [PR4]

4.4.2. Handwritten letter recognition

Encouraged by the promising results obtained within the previously proposed application, I advanced the idea of recognizing the handwritten letters that make up the Romanian language alphabet because the Romanian alphabet has 5 special letters (written with diacritics). Correct use of these special letters is vital to correct word formation because misuse can change the meaning of words and sentences.

Handwriting is vital in education, everyone's style – which varies in size, width and spacing, bringing unfamiliar elements to pattern recognition. Handwriting recognition is a key innovation that transforms handwritten text into digital form through image acquisition.

To implement the algorithm proposed in this case study, in Figure 4.15 we detailed the workflow [PR5].

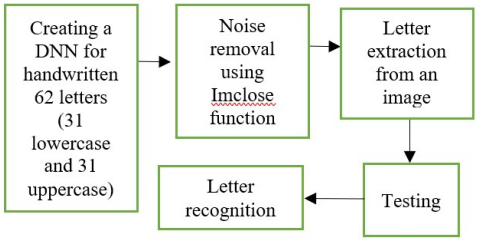

Figure 4.15 *Workflow of the proposed algorithm for recognition letters of the Romanian language alphabet* [PR5]

4.4.3. Generating synthesized images with augmentation algorithm

Image analyzes by humans can lead to different results due to different levels of knowledge, but computer analysis provides more consistent and complex results. Computer vision techniques automate pattern recognition processes for object detection and classification by analyzing high-quality image databases and creating a diverse dataset. A robust model in image processing is obtained using image augmentation, and the realization methods are [28]:

- circular shift: image augmentation method independent of its inherent properties, applicable to any types of images in the database.
- rotation. Rotation increments involve rotating the original image around an axis by 1º to 359º.
- translation. The resulting augmentations shift the images in any direction, preventing data displacement problems.

The personal approach involved analyzing the effects of digital image augmentation procedures with MMO. The working method is illustrated in below Figure 4.21 [PR6].

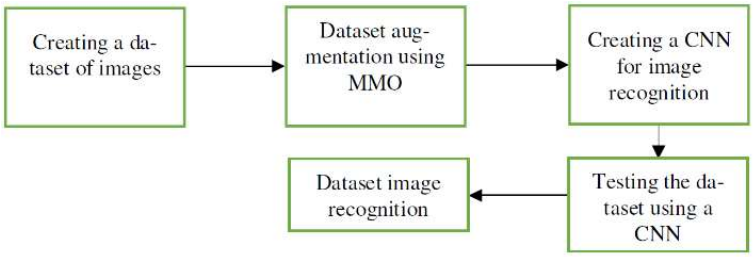

Figure 4.21 *Image augmentation workflow* [PR6]

4.4.4. Determination of vessel patterns using MMO and CNN

The approach we proposed in [PR7] aimed to contribute to the field of boat classification from images using a CNN trained on images of different types of boats (surface images of private ships), using mathematical morphology to reduce noise in images.

4.4.5. Optimizing datasets using MM techniques

Considering the aspects mentioned above, this case study aims to evaluate the impact of the MMO technique on improving the accuracy and precision of the VGG16 and VGG19 networks in earthquake crack detection and classification.

Table 4.16 contains values for accuracy, running time, positive and negative estimates, all obtained over a number of epochs equal to 2 and 3, respectively.

Table 4.16 *Metrics for estimating the performance of the proposed solution* [PR8]

	Epochs number = 2				**Epochs number = 3**			
Model CNN	**Accuracy**	**Compilation time**	**PPV**	**NPV**	**Accuracy**	**Compilation time**	**PPV**	**NPV**
VGG16	98.83%	704 minutes	99.10	98.10	98.83%	621 minutes	98.95	98.90
VGG19	98.88%	504 minutes	99.60	97.75	92.20%	600 minutes	99.35	87.20

4.4.6. Evaluation of earthquake-induced structural defects

Despite the many studies on this topic, this study is motivated by the results in subchapter 4.4.5, where the opening MMO significantly improved the detection accuracy with the VGG16 and VGG19 networks. Our main goal is to improve the accuracy of pre-trained neural networks with noise-affected images in civil engineering. We used a dataset with three classes of images: cracks, no cracks and noise.

Determining the accuracy values for the fusion of VGG16, VGG19 and MMO, it was observed that the best values and the lowest running time are obtained using open MMO (Table 4.20).

Table 4.20 *Evaluation of accuracy and running time* [PR9]

CNN model used	**Accuracy [%]**	**Compilation time**
VGG16	74.64	916 minutes
VGG19	78.63	1016 minutes
VGG16 and MMO dilation	76.74	821 minutes
VGG16 and MMO erosion	77.60	1279 minutes
VGG16 and MMO opening	**80.83**	**763 minutes**
VGG16 and MMO closing	77.37	763 minutes

4.4.7 TRIB Crack Database: Useful Tool for Automatic Crack Detection in Road Infrastructure

In image processing, the reliability and accuracy of the results largely depend on the quality of the processed images. Thus, while my previous research was primarily based on processing images from freely available online sources, developing a dedicated database tailored to a specific case study became essential. Starting from the concept of identifying structural faults caused by earthquakes, described in subsection 4.4.6, this section details the methodology for building an image database that captures existing cracks in road infrastructure, how to process them, and their utility.

Chapter 5

Conclusions

5.1 OBTAINED RESULTS

The doctoral thesis presents, in an integrated form, the research activity carried out during the doctoral studies, research that consisted both in the in-depth study of the necessary theoretical elements, of the current state of research in the field, but also in the publication of scientific articles that address from a personal perspective, the current topic of shape detection, shape recognition, but also their classification.

Chapter 1 provides a comprehensive overview of the chosen field, establishing the motivation for choosing the research topic. The scope of the topic addressed in the doctoral thesis, the objectives of the study and the detailed summary of the content of the thesis are presented.

In Chapter 2, the theoretical study related to shape detection algorithms is presented. Considering their proven efficiency in various studies, the focus is on detailing Viola-Jones algorithms, ACF and mathematical morphology operators.

Chapter 3 presents the theory related to the main methods of filtering noise from images, while addressing the methods of detection and identification of geometric structures of objects present in images. Aspects related to nonlinear filtering are detailed, and their classification (into averaging, bilateral, and MM filtering) is described.

In Chapter 4, aspects related to DL algorithms were introduced, to advance the doctoral research activity towards current image processing techniques. Performance metrics (accuracy, precision, detection rate) were defined and a classification of DL into three broad machine learning categories: supervised, unsupervised and hybrid was performed.

The last sections of the thesis briefly present the relevant results obtained and presented in chapters, the personal contributions identified in the research activity presented in the thesis, the list of publications of the author of the thesis, as well as his own vision for the continuation of research activities in this field.

5.2 PERSONAL CONTRIBUTIONS

Personal contributions to the research activity presented in this thesis consist of:

1. Detailed analysis of shape detection methods, analysis focused on the effectiveness of Viola-Jones algorithms, ACF and mathematical morphology operators (MMOs) for accurate detection (with minimal resources and processing time) of faces and objects in images in [PR1].
2. The effective application of erosion (fundamental MMO) was achieved in the development of a method for determining the weight of the human body to reduce fuel consumption and carbon emissions in civil aviation. Considering the numerous studies stating the dependence of an aircraft's fuel consumption on the load on board, the proposed scenario aimed to develop a method for determining human body weight through an alternative method to the classical weighing method. Erosion was the proposed MM filtering method for the correct detection of a passenger's silhouette by removing noise from the tested images [PR2].
3. Carrying out a thorough review of the image processing literature and classifying noise into three categories, allowing a detailed study of filtering techniques relevant to each of these types of noise in [PR3].
4. The superiority of MM filters over median filters was also demonstrated in the application described in [PR4], where we combined MM filters with a DNN network for accurate recognition and classification of digital images by choosing an appropriate filter to remove *"salt and pepper" noise"*.
5. Based on the results of [PR4], an innovative approach represented the detection and recognition of letters from the Romanian alphabet, the main challenge being the detection of letters with diacritics. The optimized solution using MM filters included the implementation of Adam augmentation and optimization techniques and analysis based on the number of epochs [PR5].
6. An important contribution regarding the image augmentation technique was highlighted in [PR6] by comparing the results obtained by data size reduction and the aperture MMO filter.
7. Our own approach, detailed in [PR7], provides a viable alternative to RGB cameras and satellite data for boat recognition. By integrating MM filters with CNNs, our solution demonstrates significant advantages over similar methods.
8. Other relevant contributions in the field of image processing through the fusion of MMO with CNN were made in the scenarios related to the detection of cracks in civil infrastructure presented in [PR8] and [PR9]. The optimal design of data for training the VGG16 and VGG19 networks, the choice of appropriate optimization techniques (augmentation and the Adam algorithm), as well as extensive testing of all combinations of CNN and MMO networks (erosion, dilation, opening and closing) facilitated, on sets of different data, for different numbers of epochs, reaching values for the accuracy of the proposed methods above 98% in the case of the fusion between MMO and CNN.

9. In article [PR10] I presented the major contribution we made in the field of image processing by creating our own TRIB Crack database, a database that contains images with cracks from different points of the road infrastructure in Bucharest. In order to demonstrate its validity, we performed operations of separating and resizing them (from 2160 x 2160 pixels to 256 x 256 pixels), we successfully used for the classification of images (with or without cracks) two CNNs (respectively Darknet-19 and Darknet -53), and the removal of noise from the images as a result of the segmentation operation was performed with three MMOs placed in series. In addition, we proposed a way to determine the sizes of the cracks in order to be able to emphasize the usefulness and importance of these processes in the management of the road infrastructure.

5.3 List of original publications

My personal concern for shape detection algorithms in images (with detailed analysis of mathematical morphology operators), optimization of global performances of algorithms analyzed together with convolutional neural networks is demonstrated by the 9 published scientific articles of which: *1 article in ISI indexed journal Q2 [PR9], 2 international conferences indexed by ISI [PR6] and [PR7], 3 conferences currently being indexed by ISI, this year currently being indexed IEEE Xplore and ACM [PR3]- [PR5], and two papers presented at conferences indexed in other BDIs (SPIE)*. Additionally, in [PR10] the data of the article currently under evaluation at a journal indexed Q2 is presented. It is important to note that I am the first author of all published articles.

Next, I will list the publications resulting from the doctoral research activity:

[PR1] **D. Abrudan**, A.-M. Drăgulinescu, R.-O. Preda, and N. Vizireanu "*COVID-19 approved mask detection using mathematical morphology*", Proc. SPIE 12493, Advanced Topics in Optoelectronics, Microelectronics, and Nanotechnologies XI, 124932O, 2023, https://doi.org/10.1117/12.2642041 (BDI: SPIE Digital Library)

Results from article [PR1] were included in Section 2.4.

[PR2] **D. Abrudan**, A.-M. Drăgulinescu, R.-O. Preda, and N. Vizireanu "*Fuel burn reduction in commercial aviation using mathematical morphology*", Proc. SPIE 12493, Advanced Topics in Optoelectronics, Microelectronics, and Nanotechnologies XI, 124932Y, 2023, https://doi.org/10.1117/12.2643287 (BDI: SPIE Digital Library) (**1 citation** in proceeding paper SpringerLink)

Results from article [PR2] were included in Section 2.4.

[PR3] **Abrudan Dumitru**, A. Drăgulinescu and N. Vizireanu, "*Pinewood Knot Detection Method Using Color Analysis*", 2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest,

Romania, **June 29-30, 2023**, pp. 01-05, doi: 10.1109/ECAI58194.2023.10193948 (BDI: IEEE Xplore)

Results from article [PR3] were included in Section 3.4.

[PR4] **Abrudan Dumitru**, A.-M. Drăgulinescu and N. Vizireanu, "*Recognition of Handwritten Mathematical Symbols Affected by Noise using DNN*", 2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, **June 29-30, 2023**, pp. 01-04, doi: 10.1109/ECAI58194.2023.10194115 (BDI: IEEE Xplore)

Results from article [PR4] were included in Section 4.4.1.

[PR5] **Abrudan Dumitru**, A. -M. Drăgulinescu and N. Vizireanu, "*Handwritten letter recognition using mathematical morphology*", 2023 IEEE 29th International Symposium for Design and Technology in Electronic Packaging (SIITME), Craiova, Romania, 18-22 **October 2023**, pp. 20-23, doi: 10.1109/SIITME59799.2023.10431387 (BDI: IEEE Xplore)

Results from article [PR5] were included in Section 4.4.2.

[PR6] **D. Abrudan**, I. M. Marcu, N. Vizireanu, "*Digital images augmentation using mathematical morphology operators*", In: Yang, XS., Sherratt, S., Dey, N., Joshi, A. (eds) Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems, vol 1012. Springer, Singapore. https://doi.org/10.1007/978-981-97-3556-3_45. **Feb 2024**) (WOS: 001326995600045)

Results from article [PR6] were included in Section 4.4.3.

[PR7] **D. Abrudan**, A.-M. Drăgulinescu, N. Vizireanu, "*Boats Imagery Classification using Deep Learning*", In: Rocha, Á., Adeli, H., Dzemyda, G., Moreira, F., Poniszewska-Marańda, A. (eds) Good Practices and New Perspectives in Information Systems and Technologies. WorldCIST 2024. Lecture Notes in Networks and Systems, vol 987. Springer, Cham. https://doi.org/10.1007/978-3-031-60221-4_43 **May 2024** (Indexed in *SpringerLink*) URL: https://link.springer.com/chapter/10.1007/978-3-031-60221-4_43#citeas (WOS: 001267237000043)

Results from article [PR7] were included in Section 4.4.4.

[PR8] **D. Abrudan**, I. M. Marcu, A.-M. Claudia Dragulinescu, N. Vizireanu, "*Images data set improvement using mathematical morphology operators*", CACML '24: Proceedings of the 2024 3rd Asia Conference on Algorithms, Computing and Machine Learning, Pp. 450 – 455, https://doi.org/10.1145/3654823.365491 **May 2024** (BDI: ACM Digital Library)

Results from article [PR8] were included in Section 4.4.5.

[PR9] **D. Abrudan**, A.-M. Claudia Dragulinescu, N. Vizireanu, "*Applications of mathematical morphology operators in civil infrastructures*", EARTH

SCIENCE INFORMATICS, ISSN: 1865-0473, https://doi.org/10.1007/s12145-024-01379-3 **June 2024** (**ISI, Q2, IF: 2.7, WOS: 001254810900001**)

Results from article [PR9] were included in Section 4.4.6.

[PR10] **D. Abrudan**, "*TRIB Crack Dataset: Automatic Recognition System for Road Cracks Detection*", sent for reviewing to journal EARTH SCIENCE INFORMATICS, ISSN: 1865-0473, **October 2024** (**ISI, Q2, IF: 2.7**)

Results from article [PR10] were included in Section 4.4.7.

5.4 Perspectives for further developments

Shape detection in both static and real-time images presents significant challenges in designing effective solutions for various applications. Based on research conducted during the PhD program, we have identified several key areas for future development:

- for the correct and precise identification of the contours of human silhouettes from static and moving images, it will be exploited after the doctoral research activity in as many different scenarios as possible and on images acquired in real time. In addition, more complex analysis will involve testing new adaptive filtering methods as well as combinations thereof.
- Improving the MMO technique for more accurate weight estimation, considering factors such as loose hair and clothing that affect overall accuracy and including weight estimation from images captured by airport video cameras.
- Application of the combined MMO and CNN solution for detection, recognition and classification in color images, as well as the analysis of other image processing algorithms.
- Examining methods for recognizing the size of earthquake-induced cracks, using those sizes as a criterion for other analyses, as well as exploring other methods to improve the accuracy of the studied methods.

Bibliography

[1] P. Viola, M. Jones, *Rapid object detection using a cascade of simple features*, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

[2] H. T. Ngo, R. C. Tompkins, J. Foytik and V. K. Asari, *An area efficient modular architecture for real- time detection of multiple faces in video stream*, 6th International Conference on Information, Communications & Signal Processing, Singapore, 2007, pp. 1-5, doi: 10.1109/ICICS.2007.4449885.

[3] P. Dollár, R. Appel, S. Belongie and P. Perona, *Fast Feature Pyramids for Object Detection*, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 8, pp. 1532-1545, Aug. 2014, doi: 10.1109/TPAMI.2014.2300479.

[4] B. Yang, J. Yan, Z. Lei and S. Z. Li, *Aggregate Channel Features for Multi-view Face Detection*, In IEEE International Joint Conference on Biometrics, IEEE, (29 September 2014 - 02 October 2014, Clearwater, FL, USA), DOI: https://doi.org/10.1109/BTAS.2014.6996284.

[5] L. Bourdev and J. Brandt, *Robust object detection via soft cascade*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 236-243 vol. 2, doi: 10.1109/CVPR.2005.310.

[6] M.B. Jensen, M.P. Philipsen, T.B. Moeslund, M. Trivedi, (2016). *Comprehensive Parameter Sweep for Learning-Based Detector on Traffic Lights*, In: Bebis, G., et al. Advances in Visual Computing. ISVC 2016. Lecture Notes in Computer Science(), vol 10073. Springer, Cham. https://doi.org/10.1007/978-3-319-50832-0_10

[7] M. A. S. Sajat, H. Hashim and N. M. Tahir, *Detection of Human Bodies in Lying Position based on Aggregate Channel Features*, 16th IEEE International Colloquium on Signal Processing Its Applications (CSPA), pp. 313-317, Feb. 2020.

[8] J. Hua, Y. Shi, C. Xie, H. Zhang and J. Zhang, *Pedestrian- and Vehicle-Detection Algorithm Based on Improved Aggregated Channel Features*, in IEEE Access, vol. 9, pp. 25885-25897, 2021, doi: 10.1109/ACCESS.2021.3057616.

[9] A. Zhao et al., *An Effective Method Based on ACF for Aircraft Detection in Remote Sensing Images*, in IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 5, pp. 744-748, May 2017, doi: 10.1109/LGRS.2017.2677954.

[10] F. Astuti Hermawati, H. Tjandrasa, N. Suciati, *Combination of Aggregated Channel Features (ACF) Detector and Faster R-CNN to Improve Object Detection Performance in Fetal Ultrasound Images*, International Journal of Intelligent Engineering and Systems, Vol.11, No.6, 2018 DOI: 10.22266/ijies2018.1231.07

[11] J. Zhang, Z. Zeng, L. Zhang, Q. Lu, K. Wang, *Application of Mathematical Morphological Filtering to Improve the Resolution of Chang'E-3 Lunar Penetrating Radar Data*, Remote Sens. 2019, 11, 524. https://doi.org/10.3390/rs11050524

[12] B.A. Landman, I. Lyu, Y. Huo, A. J. Asman, *Chapter 6 - Multiatlas segmentation (6.2.6.4 Morphology correction)*, Handbook of Medical Image Computing and Computer Assisted Intervention, The Elsevier and MICCAI Society Book Series, 2020, Pages 137-164

[13] N. Efford, *Digital Image Processing: A Practical Introduction Using Java*, Publisher: Addison Wesley, ISBN-10: 0201596237, 2000 (Chapter 11 "Morphological image processing")

[14] A. Yousuf, H. Khawaja, M. S. Virk, *Conceptual design of cost-effective ice detection system based on infrared thermography*, Cold Regions Science and Technology, Volume 215, Nov. 2023, 103941 https://doi.org/10.1016/j.coldregions.2023.103941

[15] A. R. Kof, *Revised Mathematical Morphological Concepts: Dilation, Erosion, Opening and Closing*, Kwame Nkrumah University of Science and Technology, Kumasi, 2014 Online: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=14b6f54a6358ae3b052eaa54583e1ad528fcdc3a

[16] N. Jamil, T. M. T. Sembok and Z. A. Bakar, *Noise removal and enhancement of binary images using morphological operations*, 2008 International Symposium on Information Technology, Kuala Lumpur, Malaysia, 2008, pp. 1-6, doi: 10.1109/ITSIM.2008.4631954.

[17] *Introduction to object detection with deep learning*, Online: https://www.superannotate.com/blog/object-detection-with-deep-learning, Sept 2023 Accesat la 03.08.2024

[18] D.Y. Pimenov, L.R.R. da Silva, A. Ercetin, et al., *State-of-the-art review of applications of image processing techniques for tool condition monitoring on conventional machining processes*, Int J Adv Manuf Technol 130, 57–85 (2024). https://doi.org/10.1007/s00170-023-12679-1

[19] A. Kaur, G. Dong, *A Complete Review on Image Denoising Techniques for Medical Images*, Neural Process Lett 55, 7807–7850 (2023). https://doi.org/10.1007/s11063-023-11286-1

[20] M. Kowalczyk, P. Koza, P. Kupidura, J. Marciniak, *Application of Mathematical Morphology Operations for Simplification and Improvement of Correlation of Images in Close-Range Photogrammetry*, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B5. Beijing 2008

[21] H. H. Kareem, R. Husain A. Ali, G. A.Hafedh Jaber, *Noise Removed by Processing the Lightness and Chromatic Components Basic on YCbCr Color Space*, Journal of Babylon University: Journal of Applied and Pure Sciences, Volume 24, Issue 9, December 2016, Pages 2332-2345

[22] S. V. Vaseghi, *WIENER FILTERS* ,Advanced Digital Signal Processing and Noise Reduction,Second Edition, John Wiley & Sons Ltd ISBNs: 0-471-62692-9, 2000

[23] *Integral Image*, Online: https://www.mathworks.com/help/images/integral-image.html, Accesat la data de 24.07.2024

[24] S. Ehsan, A.F. Clark, N. ur Rehman, K.D. McDonald-Maier, *Integral Images: Efficient Algorithms for Their Computation and Storage in Resource-Constrained Embedded Vision Systems*, Sensors (Basel). 2015 Jul 10;15(7):16804-30. doi: 10.3390/s150716804. PMID: 26184211; PMCID: PMC4541907.

[25] R. Mohan, S.P. Chokkalingam, Kirupa Ganapathy, A. Rama, *Comparative Image Quality Analysis of Spatial Filters for Pre-processing of CT Abdominal Images*, Webology, Volume 18, Special Issue on Computing Technology and Information Management, September 2021

[26] M. Krichen, *Convolutional Neural Networks: A Survey*, Computers 2023, 12, 151. https://doi.org/10.3390/computers12080151

[27] M.Shaha, M. Pawar, *Transfer learning for image classification*, Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018) IEEE Conference Record # 42487; IEEE Xplore ISBN:978-1-5386-0965-1

[28] K. Zhang, Z. Cao and J. Wu, *Circular Shift: An Effective Data Augmentation Method For Convolutional Neural Network On Image Classification*, IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020, pp. 1676-1680, doi: 10.1109/ICIP40778.2020.9191303