

Image Analysis Methods for Biomedical and Industrial Applications

PhD Thesis

Melinda Katona

Supervisor: László G. Nyúl, PhD

Doctoral School of Computer Science

Department of Image Processing and Computer Graphics

Faculty of Science and Informatics

University of Szeged



Szeged
2019

Acknowledgments

First of all, I would like to thank my supervisor, László G. Nyúl, for directing my eyes to digital image processing when I was his student and for his continuous support ever since.

I would also like to thank my colleagues and friends who helped me to realize the results presented here and to enjoy the period of my PhD studies.

I am thankful to Dr. Rózsa Dégi, Dr. Attila Kovács, Dr. Géza Thury, Dr. Flóra Hararits for the OCT image dataset and also thank to Biomedical Application Group for the zebrafish dataset utilized in this research.

Last, but not least, I wish to thank my parents and my sister for their constant love and support.

My work was supported by funds from the Hungarian Government, in particular, by TÁMOP 4.2.4.A/2-11-1-37 2012-0001 National Excellence Program, by the TÁMOP-4.2.2.A-11/1/KONV-2012-0073, by the UNKP-18-3 New National Excellence Program of the Ministry of Human Capacities, by 20391-3/2018/FEKUSTRAT, by EFOP-3.6.3-VEKOP-16-2017-0002 and a grant from Campus Mundi Scholarship. The NFI research was financed by Intersoft Hungary Ltd. within an R&D project and all rights to commercial use of the resulting technology have been transferred to the sponsoring firm.

Abbreviations

2D	two dimensional	INT	intensification operator
3D	three dimensional	INL	inner nuclear layer
AMD	age-related macular degeneration	IPL	inner plexiform layer
BV	blood vessel	IR	ionizing radiation
CCD	charged-coupled device	IRF	intraretinal fluid
CNN	convolutional neural network	IZ	interdigitation zone
D	distortion	J	jaccard similarity measure
DME	diabetic macular edema	kNN	k-nearest neighbors
DNA	deoxyribonucleic acid	LBP	local binary pattern
EZ	ellipsoid zone	MD	medical doctor
FN	false negative	MV	megavolt
FP	false positive	NIQE	naturalness image quality evaluator
GCL	ganglion cell layer	NFI	natural feature identifier
Gy	Gray	OCT	optical coherence tomography
H	Hessian matrix	ONH	optic nerve head
HoG	histogram of oriented gradients	ONL	outer nuclear layer
hpf	hours post fertilization	OPL	outer plexiform layer
HF	hyperreflective foci	ORT	outer retinal tubulation
ILM	internal limiting membrane	PED	pigment epithelial detachment
		ROI	region of interest

RNFL	retinal nerve fiber layer	SLO	scanning laser ophthalmoscope
RPE	retinal pigment epithelial	SRF	subretinal fluid
S	size of structure element	TD-OCT	time-domain optical coherence tomography
SEAD	symptomatic exudate-associated derangements	TP	true positive
SD-OCT	spectral-domain optical coherence tomography	VEGF	vascular endothelial growth factor
SHRM	subretinal hyperreflective material		

Contents

1	Introduction	1
2	Retinal Image Analysis	5
2.1	Introduction to OCT retinal imaging	6
2.1.1	Age-related macular degeneration	7
2.2	SD-OCT image dataset and metrics	9
2.3	Delineation of boundary layers	10
2.3.1	Blood vessel segmentation	11
2.3.2	Algorithm based on vertical projected data	13
2.3.3	Graph-cut based approach	15
2.3.4	Evaluation and results	16
2.3.5	Measurement of retina distortion	18
2.4	Detection of subretinal and intraretinal fluid	19
2.4.1	Segmentation based on projected data	21
2.4.2	Detection and separation using active contour	22
2.4.3	Evaluation and results	23
2.5	Determination of PED and SHRM	25
2.5.1	Abnormality detection around the RPE layer	26
2.5.2	Evaluation and results	28
2.6	Segmentation of outer retinal tubulation	28
2.6.1	Detection using classical image processing operators	29
2.6.2	Localization with neural network	30
2.6.3	Evaluation and results	33
2.7	Concluding remarks	35
3	Quantitative Analysis of Irradiated Zebrafish Embryos	37
3.1	Introduction	38
3.2	Quantitative analysis of morphological changes	39
3.2.1	Preprocessing and delimitation of the well	40
3.2.2	Segmentation of zebrafish embryo	41
3.2.3	Extraction of morphological deteriorations	43

3.3	Evaluation and results	44
3.4	Concluding remarks	49
4	Localization of Visual Codes	51
4.1	Introduction to barcodes	51
4.2	Image data and metric	55
4.3	Mathematical morphology based algorithms	56
4.3.1	Algorithm based on bottom-hat transform	57
4.3.2	Extending the bottom-hat approach	58
4.3.3	Evaluation and results	60
4.4	Detection using template matching	64
4.4.1	1D traditional barcode segmentation	64
4.4.2	QR code localization	66
4.4.3	Evaluation and results	67
4.5	Validating authenticity of visual codes	69
4.5.1	Localization and validation of NFI label	71
4.5.2	Image processing steps of recognition	73
4.5.3	Label identification	77
4.5.4	Evaluation and results	79
4.6	Concluding remarks	82
	Bibliography	83
	Summary	95
	Összefoglalás	99
	Publications	105

Chapter 1

Introduction

Imaging technology is at an advanced level of development nowadays and the processing of large amounts of image data raises many image processing and image analysis problems. In this dissertation, we address the various problems of image analysis in areas of medical imaging based diagnostics, microscopy in biology and localization of visual codes. In cases when the structure of the object is well defined it can be expected that an image analysis method provides acceptable results automatically such as for processing visual codes. In other cases, e.g., the interpretation of medical images can be uncertain even for human experts, thus here we aim to give only assistance that the doctors can use in diagnosis. Classical image processing approaches are well applicable for various detection and segmentation tasks and these do not require a large set of image data which was not available during our work. Besides, in recent years deep neural networks became widespread in image processing. These try to provide more general solutions in many areas. Their disadvantage is that the methods require a large amount of labeled data for training. The methods presented in the dissertation are based on classical image processing operations, but we also applied a deep learning approach to one of the retina analysis tasks.

The human eye is the sense organ of vision in the human body that reacts to the rays of light emanating from environmental objects and maps their shape, spatial position and color. Incoming rays of light pass through the vitreous and focus on the retina. The retina is the sensory membrane and composed of several layers, including photoreceptors which take light focused by the cornea and lens. The small central area of the retina called the macula. This region of the retina is responsible for central vision. Various diseases can affect the eye. Several imaging techniques are available to investigate the retina. Age-related macular degeneration (AMD) is one of the most frequent reasons of acquired blindness in economically advanced countries. Previously, no medical treatment was possible, but today, certain types of disease have the potential to prevent severe visual impairment. In the last decade, optical coherence tomography (OCT) has become a basic imaging technique in diagnosing

and following neovascular AMD and its response to treatment, since it provides high-resolution cross-sectional images of the retina. With OCT we can detect not only the layers of the retina, but we are also capable of identifying the detailed morphology and the effects of the disease, the so-called OCT biomarkers, and AMD has many biomarkers that can be used to describe the condition of the patient. We present the automatic localization and quantitative characterization of these biomarkers. Two solutions are introduced to delimitate the retina ROI on images and determine the position of the disease-induced distortion. There are a large number of publications in the scientific literature that provide an approach to localize the subretinal fluid and cyst that we have also investigated. Besides, we also discuss methods for pigmentepithelial detachment, subretinal hyperreflective material and outer retinal tubulation in Chapter 2.

Radiotherapy is one of the most common methods of treating various cancers, although, unfortunately, its adverse effects on healthy tissues are known. Zebrafish are excellent model for experimental human cancer research. Embryo development is extremely rapid during the first few days post-fertilization while the embryos and larvae are transparent, giving the possibility to study the *in vivo* organ development. Morphological changes in embryos can be observed in a dose-dependent manner after radiation treatment. In Chapter 3 we present a method that is capable of obtaining the various shape characteristics of zebrafish in order to give a quicker picture of the response of fish to irradiation. The statistical analysis shows that there is no significant difference between manual annotation and the values computed by the automatic algorithm.

Visual codes play an important role in automatic identification, which became an inseparable part of industrial processes. Thanks to the revolution of smartphones and telecommunication, it also becomes more and more popular in everyday life, containing embedded web addresses or other small informative texts. While barcode reading is straightforward in images having optimal parameters (focus, illumination, code orientation, and position), localization of code regions is still challenging in many scenarios. Every setup has its own characteristics, therefore many approaches are justifiable. Industrial applications are likely to have more fixed parameters like illumination, camera type and code size, and processing speed and accuracy are the most important requirements. In everyday use, like with smartphone cameras, a wide variety of code types, sizes, noise levels and blurring can be observed, but the processing speed is often not crucial, and the image acquisition process can be repeated in order for successful detection. In Chapter 4 algorithms developed for the automatic detection of visual codes are discussed. We present two methods that localizes barcodes based on mathematical morphology. Both 1D traditional barcodes and QR codes have a specific structure, so template matching is a possible way for detection. We present a method to segment classical barcodes and another one to

localize QR codes. An extensive comparative analysis of code localization methods was performed on publicly available and on our databases. In Section 4 we also present a possible hybrid visual code in which a well-defined artificial identifier (a QR code) and an artificial pattern (formed by dispersed glitters) appear in common image space. New identifiers can be defined using a combination of artificial and natural feature patterns. We can consider it as a kind of “hybrid” visual code which in itself carries the well-defined structure of artificial identifiers and the singularity of natural identifiers. We introduce a method that detects the location of glitters on a specific region of the label and supports the identification of labeled objects on images taken with mobile devices.

Chapter 2

Retinal Image Analysis

The human eye is the sense organ of vision in the human body that reacts to the rays of light emanating from environmental objects and maps their shape, spatial position and color. Incoming rays of light pass through the vitreous body and focus on the retina. The retina is the sensory membrane and is composed of several layers, including photoreceptors which take light focused by the cornea and lens. The small central area of the retina is called the macula, and this region is responsible for central vision. Various diseases can affect the eye. Several imaging techniques are available to investigate the retina and help diagnosis. A modern tool for *age-related macular degeneration* (AMD) examination is *optical coherence tomography* (OCT). AMD is a common eye condition among people above 50 years worldwide, currently affecting 170 million people globally. Existing OCT systems show many features about AMD, such as measurable indicators of some biological state or condition that are so-called OCT *biomarkers*. To improve the treatment procedure, there is a need for more precise measurements, hence our aim was to create automatic methods for quantitative assessment of biomarkers.

Our aim was to segment those biomarkers that are more relevant in diagnosis of AMD or during the treatment. The execution time of the approaches was negligible for our medical doctors because the accuracy was the most important aspect. We introduce retinal imaging in Section 2.1. Details of the used dataset and metrics are discussed in Section 2.2. In Section 2.3 we introduce how to extract outer boundaries and other information that can be used to segment various biomarkares more precisely [87, 92]. We deal with the localization of subretinal and intraretinal fluid in Section 2.4 [87, 89]. The biomarkers and the extension of the *pigment epithel detachment* (PED) and *subretinal hyperreflective material* (SHRM) regions also occur as a consequence of other diseases for which we propose solutions in Section 2.5 [92]. In recent years, it has been reported in the medical literature that ORT can also be a descriptive of AMD disease, so this area has been addressed in Section 2.6 [56, 90].

2.1 Introduction to OCT retinal imaging

The appearance of optical coherence tomography dates back to the early 90s. First, Huang et al. [38] accomplished successful measurements in 1991. The developed technology is based on the Michelson interferometer which enables accurate distance estimation. The first in vivo clinical trials began in 1993 [24] and the technology was patented the following year.

The appearance of the device constituted a supplementary technology in ophthalmic diagnostics so it is no wonder that it is one of the main diagnostic tools in this field. The speciality of OCT is the use of a light source to produce high-resolution, cross-sectional images in non-invasive manner that distinguishes different optical density areas. Other advantages of the device include low cost and robust recording.

The first generation of OCTs was called *time-domain* OCT (TD-OCT). The initial systems were only capable of 2D imaging. These devices produce a few hundred depth profiles called A-scans per second which assemble the 2D OCT image.

As technology advanced, the *spectral domain* OCT (SD-OCT) has been released. SD-OCT is based on the principle of low coherence interferometry. During this, the light reflected from the eye interferes with another beam whose path is unknown (Michelson interferometer). The position of the reference mirror does not change, the interference pattern is distributed by a grid system to frequency components and each component is simultaneously sensed by the detector. The detector is a *charged-coupled device* (CCD) that is sensitive to a certain wavelength range. All sensed frequencies from the signal after Fourier transformation correspond to the given tissue depth [98]. The method largely increases the sampling rate because each point of each A-scan can be detected simultaneously. By collecting adjacent A-scans for consecutive pixels along a transversal coordinate, a cross-section image is obtained called B-scan. The schematic diagram of a SD-OCT device is shown in Figure 2.1.

As mentioned, early devices were only capable of 2D imaging, but as technology advanced, nowadays, they can be used for 3D imaging. However, the two most commonly used acquisition types in the clinical routine are the 2D scan which is a linear scan through the macula or a circular scan around the *optic nerve head* (ONH). The path of the scan is drawn on the scanning laser ophthalmoscope (SLO) image. The SLO images are acquired during the same process as the OCT recording. The latter shows very clearly the layered structure of the retina. Figure 2.2 illustrates a SLO image on the left side of the figure and on the right it is an OCT B-scan and the structure of visible layers.

We note that fundus photography is similar to OCT and there are automatic image processing tools to recognize diseases causing specific distortions based on fundus images [9, 35, 67].

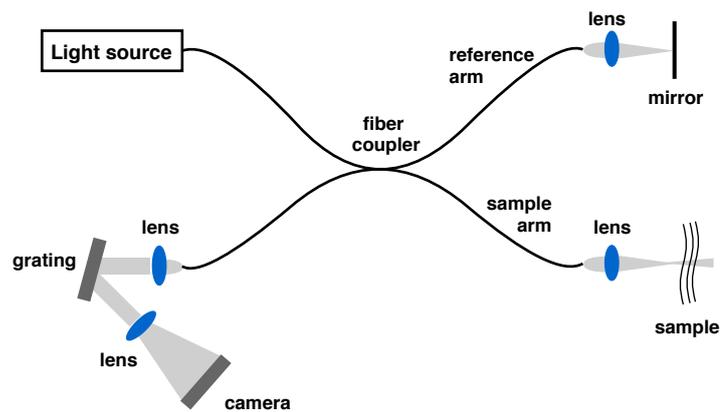


Figure 2.1: Schematic diagram of a spectral domain OCT system. The light from the source is split into a sample and reference arm using a fiber coupler. The beam is reflected by a moving mirror and interfering light is focused on a detector. Lens and objectives also concentrate the beam.

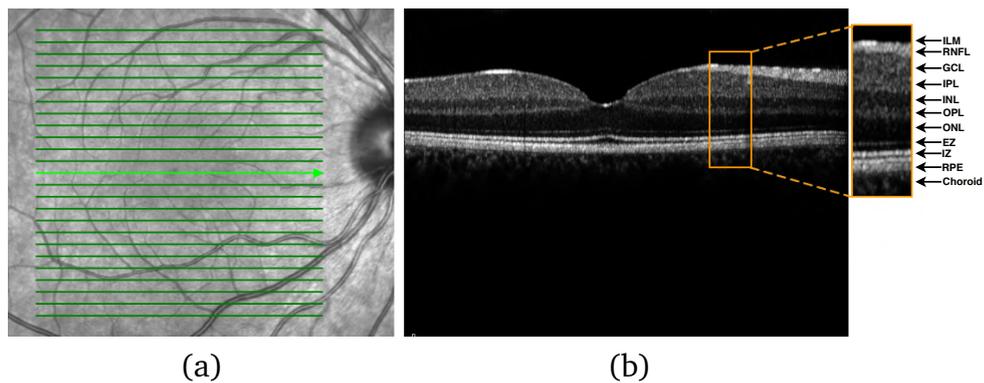


Figure 2.2: Sample SD-OCT image from a normal retina. (a) Scanning Laser Ophthalmoscope (SLO) image. Green, bold line denotes the actual scan path (b) OCT B-scan with denominations of the retinal layers. Abbreviations from top to bottom: internal limiting membrane (ILM), retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer (ONL), ellipsoid zone (EZ), interdigitation zone (IZ), retinal pigment epithelium (RPE).

2.1.1 Age-related macular degeneration

Age-related macular degeneration is one of the most frequent reasons of acquired blindness in economically advanced countries. The constant growing of AMD patient population is more and more challenging. AMD means degeneration of the macula which is the region of the retina responsible for central vision. Since AMD affects only this specific part of the retina, unattended patients lose their fine shape- and

face recognition, reading ability, and central vision [37].

Basically, AMD has two forms: dry and wet form. In the case of dry AMD, drusen, small fatty deposits accumulate at the central part of the retina in the tissues beneath the macula, leading to drying. The extent of vision loss is related to the location and this type of AMD typically has three stages: early, intermediate and advanced.

Wet form causes rapid and serious visual impairment in 10% of the AMD cases [51]. In this more acute, neovascular type of the disease, abnormal angiogenesis causes fluid and blood leakage into the retinal layers thus resulting in photoreceptor lesion. Albeit the exact pathomechanism of the disease is still unclear, it is known that the vascular endothelial growth factor (VEGF) plays crucial role in the pathogenesis [96]. The first choice of treatment in neovascular AMD is anti-VEGF intravitreal injection, a periodic injection into the eye. To follow up the progression of the disease, the Amsler grid will help at an early, treatable stage. Figure 2.3 presents the dry form of AMD in top row, and wet AMD where Amsler grid test also available.

In the last decade, Optical Coherence Tomography has been widely used in the diagnosis of AMD and follow-up therapy. Spectral domain OCT produces 3D volumes of data, which have been useful in clinical practice. Existing OCT systems are partially suited to monitoring the progress of the disease, but OCT shows many features about AMD such as hyperreflective foci (HF), subretinal fluid (SRF), intraretinal fluid (IRF), outer retinal tubulation (ORT), pigment epithelial detachment (PED), subretinal hyperreflective material (SHRM). Figure 2.4 illustrates an SD-OCT B-scan with

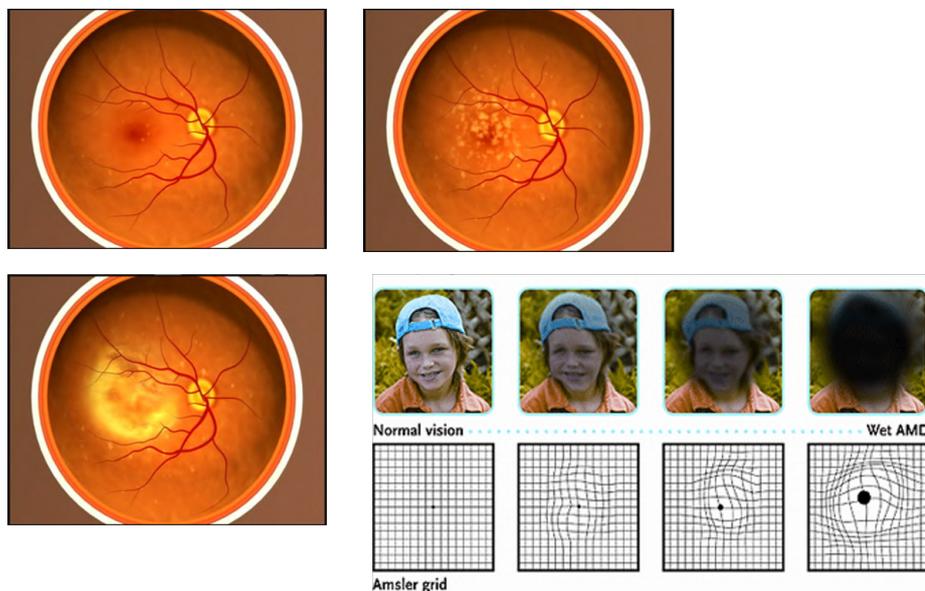


Figure 2.3: Dry (top row) and wet (bottom row) form of Age Related Macular Degeneration on fundus images. The Amsler grid (bottom right) used to test the condition of the eye. (Image source: <http://www.areds2.org.uk/macular-degeneration>)

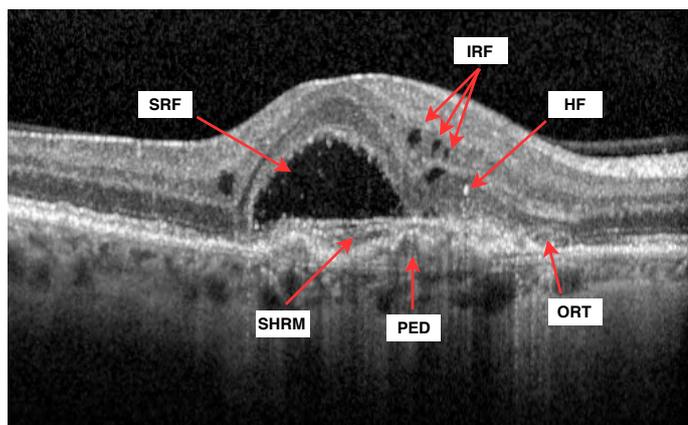


Figure 2.4: Various types of biomarkers of AMD. Subretinal fluid (SRF), intraretinal fluid (IRF), hyperreflective foci (HF), subretinal hyperreflective material (SHRM), pigment epithelial detachment (PED), outer retinal tubulation (ORT).

Table 2.1: Dependencies between the automatically detected abnormalities and biomarkers. For features in the table rows 'X' denotes which biomarker and/or abnormality information is used.

	BV	D	ILM/RPE	PED/SHRM	SRF/IRF	ORT
blood vessel (BV)			x		x	
distortion (D)				x	x	x
ILM/RPE		x		x	x	x
PED/SHRM						
SRF/IRF				x		x
ORT						

biomarkers of AMD. Dependencies between the introduced abnormalities and our approaches are shown in Table 2.1.

2.2 SD-OCT image dataset and metrics

SD-OCT images created by ophthalmologists were used to develop and evaluate the automated methods. The OCT images were acquired using Heidelberg Spectralis OCT (Spectralis, Heidelberg Engineering, Heidelberg, Germany) on wet age-related macular degeneration patients. They were either treatment naive or with anti-VEGF intravitreal injections. Medical experts annotated 35 sequences for validation. The images contained several biomarkers including subretinal fluid, IRF, PED, SHRM, ORT and ILM/RPE boundaries.

These sequences consisted of 49 or 25 slices (ophthalmologists use 25 slices protocol in daily routine). In the case of 49 slices, the pattern size was of 6×6 and the slice distance was $122 \mu\text{m}$. Slices had a resolution of 512×496 pixels with pixel sizes 11.45 and $3.87 \mu\text{m}$ in x and z directions and a quality score above 16 dB. The recordings consisting of 25 slices were taken with a 6×6 pattern size and the distance between slices was $251 \mu\text{m}$. The resolution was the same as in the above case with pixel sizes 11.74 and $3.87 \mu\text{m}$ in x and z directions, and a quality score above 23 dB.

In most algorithms the task is to detect a specific biomarker as accurately as possible to determine the extent, but in some cases the cardinality also provides sufficient information for the doctor. For this reason, we have applied such metrics where we can determine the merit of overlapping between annotation and segmented objects.

We used Dice coefficients to measure the ratio of overlap.

$$\text{Dice coefficient} = \frac{2 \cdot |X \cap Y|}{|X| + |Y|}, \quad (2.1)$$

where $|X|$ and $|Y|$ are the numbers of pixels representing the segmented regions, while $|X \cap Y|$ are the size of overlapping regions in pixels.

In addition, for the further characterization of some procedures, the sensitivity was also calculated.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (2.2)$$

where TP (True Positive) denotes the numbers of overlapping pixels, FP (False Positive) represents the numbers of segmented, but not overlapping pixels.

In some cases, we also measured the recall of the objects.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (2.3)$$

where FN (False Negative) represents the numbers of objects which is present on the annotation but is not overlapping with any segmented object.

2.3 Delineation of boundary layers

The location and thickness measurement of retinal layers provides useful information for detecting pathological changes and diagnosing different type of retinal diseases. In the last decade, numerous approaches have been described about detecting retinal layers based on various techniques.

In the field of image processing, graph theory is used in many fields to solve several problems [2, 16, 29]. In these approaches, the graph nodes usually relate to image pixels, the graph edges are assigned to pairs of pixels, and the edge weights

depend on the intensity differences between the node pixels, and also may depend on spatial distance between the pixels. Image segmentation can then lead to a graph cutting problem, which can be solved using dynamic programming. These procedures are less tolerant to noise, which is a disadvantage, because images are often very noisy, but may give more reliable result than a simple clustering algorithm.

Another basic idea relies on the well-known energy-minimizing active contour method which, unfortunately, also has problems with low contrast and noise. Yazdanpanah et al. [106] suggested a multi-phase framework with a circular shape prior in order to model the boundaries of retinal layers and estimate the shape parameters. They used a contextual scheme to balance the weight of different terms in the energy functional. Hassan et al. [36] used a structure tensor approach combined with nonlinear diffusion process for layer detection. A structure tensor is a second-moment matrix that shows similarities and prominent orientations of the image gradient. Some other approaches use optimized boundary tracking [27] or polynomial smoothing [53].

Machine learning has been widely used in recent years, also for retinal image analysis. Lang et al. [47] used random forest classifier to segment retinal layers. The random forest classifier learns the boundary pixels between layers and produces an accurate probability map for each boundary, which is further processed to finalize boundaries. Ngo [59] et al. nominated a feature-learning regression network without human bias to segment layers. Methods based on active contour or machine learning provide effective solution, but these methods are too time-consuming.

In this section, we introduce two methods to delineate boundary layers of the retina. In the first case, we use information about where the anomalies are in the image due to the shadow of the blood vessels. Second, layer information can be used to determine the start and end points as well as the extent of layer distortion.

2.3.1 Blood vessel segmentation

Major vessels cause more shadowing effect on the OCT slices than the minor capillaries. We do not aim at a perfect segmentation of the whole vessel tree. The segmentation output will be used to identify those positions (bars) on the OCT slices where shadowing may be considerable, since this can invalidate our underlying assumption about large intensity steps indicating layer boundaries. Once we localize these less reliable parts of the image, we can specially treat (or even exclude) these parts from the layer boundary detection and handle it as “missing data” rather than using false information.

Many published methods in the literature [26] deal with the retina vessel segmentation problem and try to reach the most accurate segmentation possible. For our purposes, a rough estimate is sufficient, since the vessels are not the objective of

our studies. It can be seen in Figure 2.5(a), that the central region (the macula and its surrounding) is significantly darker than other parts of the image, therefore, the contrast between blood vessels and their surrounding background differs in the central and the peripheral parts of the image. Our aim is to localize these regions and homogenize. For this, we use a Gaussian-pyramid technique [14] with 4 levels. Figure 2.5(b) depicts one level of the pyramid. Intensity homogenization is performed as a pixel operator according to the following equation:

$$I_{(x,y)} = C \cdot \frac{I_B(x,y)}{I_O(x,y)} \quad (2.4)$$

for each pixel (x, y) in the image, where I_O is the original image, I_B is the blurred image, and C is the maximal intensity in I_O .

Intensity homogenization is followed by a contrast enhancement step, using an adaptive fuzzy contrast stretching method, that is more effective than a commonly used contrast stretching procedure because it suppresses the noise better. Let I denote the input image, I_{max} the maximum and I_{min} the minimum intensity of the image. The linear membership function $\mu_{i,j}$ is defined as

$$\mu_{i,j} = \frac{I_{i,j} - I_{min}}{I_{max} - I_{min}}, \quad (2.5)$$

i.e., the membership value for pixel (i, j) corresponds to the degree of brightness possessed by the gray level intensity of that pixel with respect to the intensity range of the whole image. This is a simple way of fuzzification, i.e., assigning fuzzy membership values to elements of a set (to the pixels, in our case). In a fuzzy processing approach, memberships are manipulated, instead of original values. We achieve contrast enhancement by using the INT fuzzy operator [54]:

$$\mu'_{i,j} = \begin{cases} 2 \cdot (\mu_{i,j})^2, & \text{if } 0 \leq \mu_{i,j} \leq T, \\ 1 - 2 \cdot (1 - \mu_{i,j})^2, & \text{otherwise,} \end{cases} \quad (2.6)$$

where T is an adaptively calculated threshold value. We use statistical mean of the intensities in each window for calculation of T value. Equation (2.6) transforms membership values that are above the threshold to much higher values and membership values that are lower than the threshold to much lower values in a nonlinear manner. The last step here is defuzzification, i.e., generating properties in the original dimensions from the resulted memberships. $I'_{i,j}$ denotes the calculated new intensity for pixel (i, j) , and is obtained by applying the inverse of the transformation used for fuzzification:

$$I'_{i,j} = I_{min} + \mu'_{i,j} \cdot (I_{max} - I_{min}). \quad (2.7)$$

Figure 2.5(d) illustrates the contrast enhanced image. The contrast enhanced image

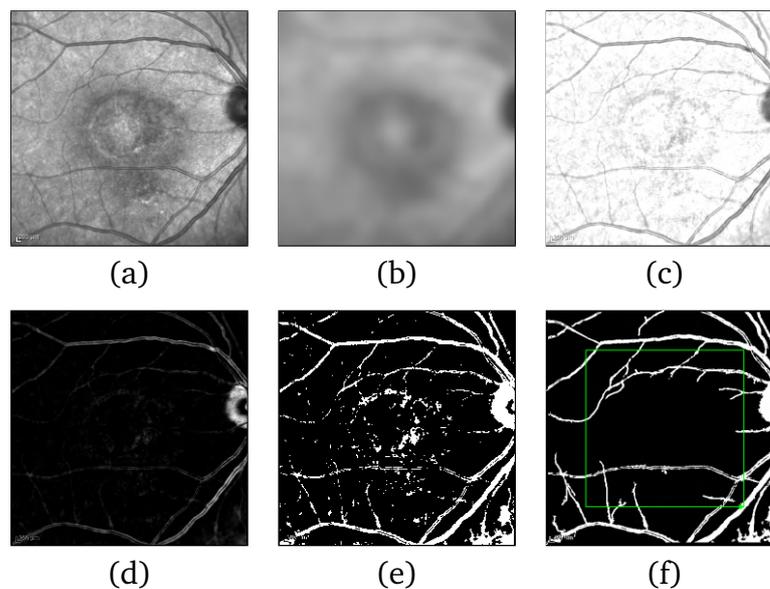


Figure 2.5: Main steps of retinal blood vessel segmentation: (a) original image, (b) one level of the Gaussian-pyramid, (c) homogenization, (d) fuzzy contrast stretching (in complement image), (e) binarization, (f) output image (after removal of false objects).

is binarized to extract possible blood vessel regions, using adaptive thresholding in a sliding window of size 15×15 pixels. In the final step, we eliminate false positive objects, because vessels form a connected object, so smaller segments possibly come from noise. We use morphological closing with rectangle shaped structure element. The size of the element was determined based on empirical observations. After that, we reject all regions whose area is less than 10 pixels and whose bounding box is not rectangle-like. Once we have the vessel mask on the fundus image, we can project this information onto the OCT cross-sections.

2.3.2 Algorithm based on vertical projected data

The quality of the OCT images are usually low, so often requires pre-processing to improve contrast and normalize intensity levels. The B-scans are noisy thus we used noise filtering and contrast enhancement using a fuzzy operator [22]. This operation can highlight major retinal layers. We analyze vertical profiles of the filtered image and large intensity steps in pixel density are assumed to correspond to change of tissue. The function is described by the expression

$$\kappa_{\nu}^* = \frac{1}{1 + \frac{1-\nu}{\nu} \left(\frac{1-x}{x} \frac{\nu}{1-\nu} \right)^{\lambda}}, \quad (2.8)$$

where ν is the threshold, x is pixel intensity and λ denotes sharpness of the filtering. The κ_ν^* function can highlight boundary layers and suppress noise. We determine dynamically the input parameter ν . For this, we use sample intensities from the upper quarter of the image and calculate average value. The λ parameter value is set to 3. Figure 2.6 shows an example of applying the κ_ν^* function.

To define the top and bottom layers, we divided the image into subregions with fixed width size of 10 and we calculate horizontal projections of each bar to determine boundaries.

We observed that the maximal width of the detected blood vessel shadows (Section 2.3.1) in our images is 10 pixels, thus this value does not affect our projection calculation. One of the major steps of our proposed method is to analyze the calculated vertical profiles. This signal is noisy, so there is a need to filter the data and remove outliers. There are many approaches available to filter 1D signals, we used the Savitzky-Golay filter [69] which is a smoothing digital filter. This filter is effective at preserving the pertinent high frequency components of the signal, which is an important aspect for our detection method. Figure 2.7 shows an example of a projection and its filtered version.

The determination of the outer layer boundary is a harder problem than that of the internal boundary, because Choriocapillaris and Choroidal vessels are located under the RPE layer. Intensity in these regions vary, so several peaks appear in the projections. Fortunately, in most cases, the noise points are not prominent thus do not cause problem in choosing the right locations. The algorithm chooses the most salient local minimum from the projected data to identify the possible internal and outer layer.

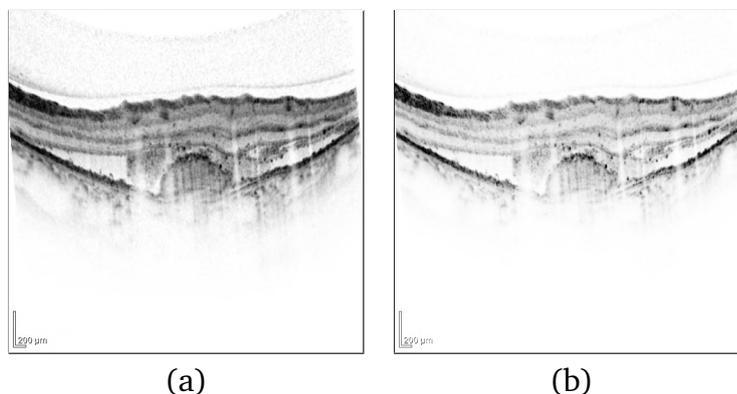


Figure 2.6: Sample OCT image before (a) and after (b) applying the κ_ν^* function. The filter highlighted boundary layers and suppressed the noise.

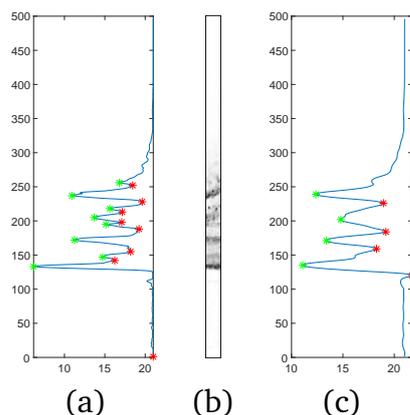


Figure 2.7: Calculated horizontal projection for vertical profile analysis. (a) original data, (b) sample bar, (c) Savitzky-Golay filtered data (polynomial order: 3, frame length: 25)

2.3.3 Graph-cut based approach

The presented method in Section 2.3.2 showed many inadequate results in RPE layer detection hence we developed a new approach. As was discussed in Section 2.3, in the literature, a number of procedures are based on graph cut [21] since it is a more robust method than other techniques. As a result, no preprocessing of the input data has been performed, although in many cases the image is very noisy.

Graph cut is a semi-automatic method that requires seed points. There are several existing approaches to automatically assigning them. We used kernel k-means to automatically determine these points. We map data points in the input space onto a high-dimensional feature space using a Gaussian radial basis kernel function.

The number of clusters is determined by how noisy the image is. So, we measure the image quality using Naturalness Image Quality Evaluator (NIQE) [57] score. This model assess the quality without knowledge of distortion. It is expressed as a distance metric between a trained model statistics and the distorted image. In higher quality images, 5 clusters were isolated empirically, while in the other cases this number was increased. In a cluster image, value 0 represents the darkest region and $k - 1$ is the brightest, where k is the number of clusters. We used graph cut [68] to optimize partitioning. This gives a better classification. We have a priori information about location of PED and SHRM which are near the RPE layer. Figure 2.8 presents a clustered image example.

The ILM elevation can be determined using a simple Otsu thresholding [62] because the foreground and background can be clearly distinguished from that part of the retina in the clustered image. The RPE layer has higher reflectivity in the original image, so it has also higher cluster number in the clustered image. Similarly to the previous method, we calculate vertical projection in every 10th column to determine

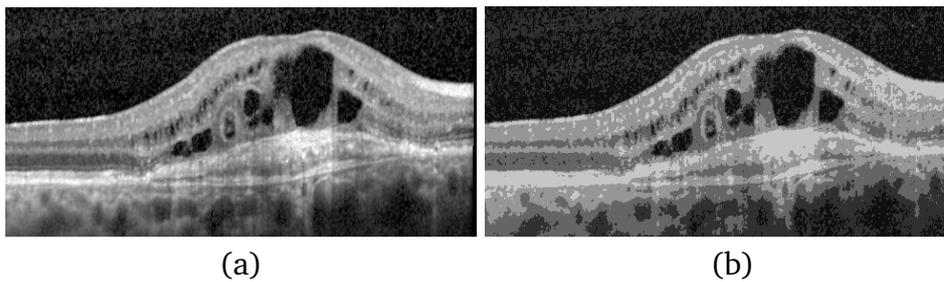


Figure 2.8: Sample clustered B-scan using graph cut with $k=5$. (a) original image, (b) clustered image.

boundary and we choose the local minimum from the projected data.

To determine all pixels which can produce the layer, we fit a curve to the resampled points. Our data may include outliers. The distribution of the points is not normal, due to the distortions, therefore outliers are defined as elements more than 1.5 interquartile ranges above the upper quartile (75%) or below the lower quartile (25%). For fitting, we use shape-preserving piecewise cubic spline interpolation to determine RPE layer.

2.3.4 Evaluation and results

To evaluate and validate our methods, we used 7 annotated sequences from a database which consisted of 49 slices. We evaluated our retinal layer detection algorithm in two different ways. On one hand, to compare the results of our algorithm against the manual delineations, we calculated the mean, maximum and standard deviation of boundary errors for each surface. The 7 curves shown in Figure 2.9 depict the error histogram for those OCT volumes where manual annotation was available. Each curve aggregates the boundary errors in the 49 scans (slices) of a study. It shows that the highest error measure is between 1 and 4 pixels in most cases and Table 2.2 confirms this statement.

As shown in Table 2.2, the maximal distance between manually segmented and automatically detected layer boundaries is 19 pixels (ca. $73.5 \mu\text{m}$). This deflection comes from two sources: substantial jumping between B-scans and layer distortions due to the disease. Unfortunately, we could not exploit 3D information directly to segment layers, because there are some anomalies among slices of the OCT volume, due to the image acquisition and registration processes within the device's software that was out of our control.

On the other hand, as it can be seen in the third sample image in Figure 2.10, there is a large difference between the manually annotated and the automatically segmented outer layer, because our algorithm follows the distorted layers. However,

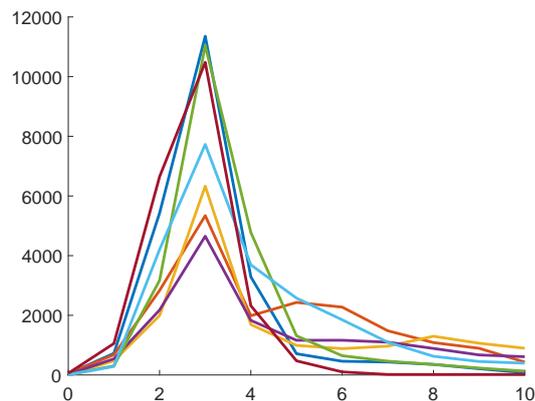


Figure 2.9: Error histogram of ILM and RPE delineation for 7 image sequences. Colors represent different recordings. The horizontal axis denotes the error for a sequence (in pixels) and vertical axis their frequencies.

Table 2.2: Summary of mean, standard deviation, and maximum error (in pixels) between manually segmented and automatically detected layers in 7 annotated OCT image sequences.

	Mean	Standard deviation	Maximum
Seq_02	2.01	1.56	17
Seq_03	2.10	0.69	15
Seq_04	1.44	0.65	15
Seq_05	1.96	0.80	19
Seq_06	2.39	1.63	18
Seq_07	2.17	0.64	15
Seq_08	1.83	0.65	17
All	1.98	0.94	16.57

the recommended possible normal boundary aligns well with the manual annotation, so it avoids the false detection. Nevertheless, in most cases, the mean errors are under 2 pixels and deviations are small between layers delineated by medical experts and the boundaries determined automatically, so the difference is negligible.

The most important advantage of the proposed method is its simpleness, that it uses simple image processing operations that can be highly parallelized, and does not rely on many parameters that are difficult to tune to the application, contrary to, e.g., an energy minimization approach.

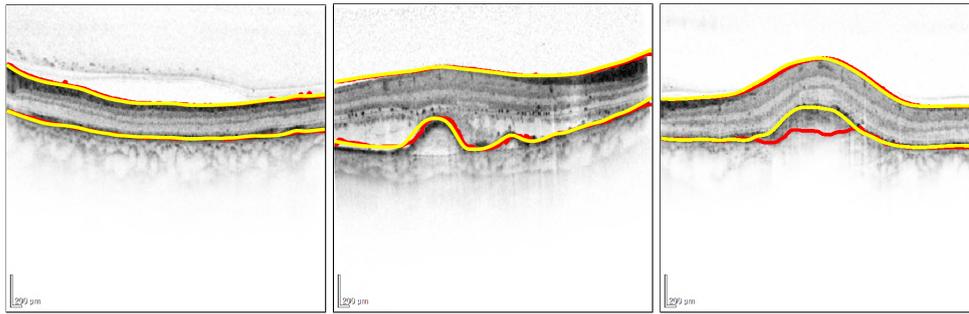


Figure 2.10: Visual comparison of expert annotated (in red) and automatically detected (in yellow) boundary layers.

2.3.5 Measurement of retina distortion

The dysfunction of RPE cells contributes to a series of pathological processes and this layer can locally detach from the underlying Bruch's membrane and fluid can accumulate underneath the RPE and/or within the retina. The detection and determination of start and end points of distorted regions can be useful as a priori information to localize biomarkers of the disease.

The B-scans clearly show that in the vast majority of cases the highest point of the distortion is in the middle of the region. Figure 2.11 summarizes our approach. From previously defined ILM layer points we select the smallest y coordinate. The $(0,0)$ coordinate is at the top left of the image.

The image is split into two parts using this point. Furthermore, the method does not use $\frac{1}{4}$ of either side of the image. This is necessary because the sides are deformed in most cases due to the image registration process during image acquiring image and layer information is not reliable. We also determine the highest y coordinates from the ILM points in the right and left retinal segment to define starting and end point

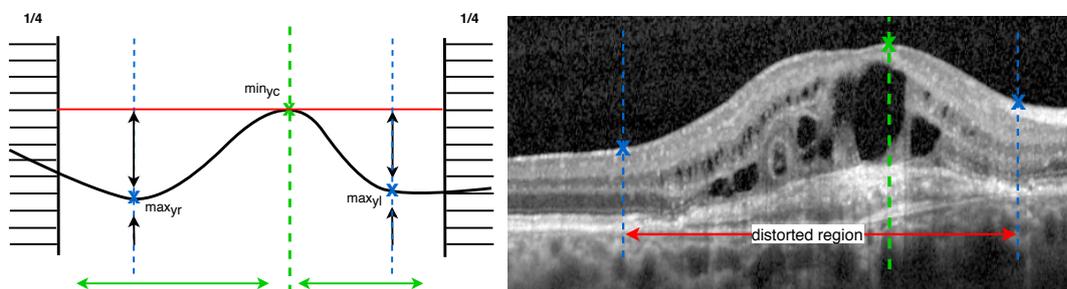


Figure 2.11: Schematic view of the determination of distortion. min_{yc} denotes the smallest y coordinate from ILM points. A line is fitted to this point (Red line). max_{yr} and max_{yl} are the highest y coordinates in the right and left retinal segments.

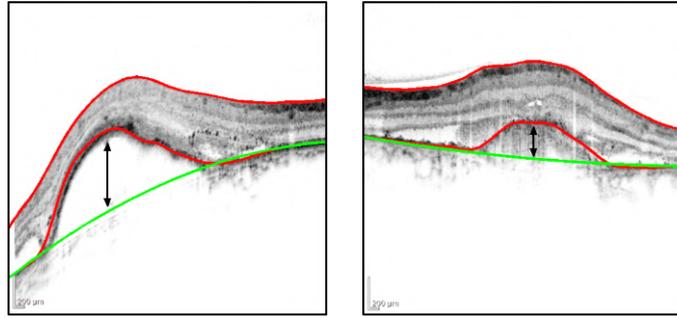


Figure 2.12: Illustration of detected boundaries. Red curve indicates the detected boundaries and green denotes the fitted possible layer. The distorted region can be seen between the fitted layer and detected boundary.

of the distortion. A line is fitted to the assigned minimal y point and we investigate the distance of the ILM layer points and the line. If the distance is smaller than $\frac{1}{5}$ of the actual average retinal thickness, the layer is not distorted.

In addition to the unambiguously specified biomarkers in the case of AMD, the measure of the distortion is also a useful descriptor for medical doctors in diagnosis. Additionally, it is feasible to determine the possible normal layer boundary if we delineate the RPE layer earlier. To define the normal RPE line, a quadratic curve was fitted to data that was assembled as follows:

- The points of RPE layer are considered for each 10 columns.
- Since it cannot be assumed that the retina is visible on the whole image thus we neglect the first two detected points.
- Take the points until the previously detected starting point of the distortion is reached in both sides and fit a curve to these data.

Considering RPE, we know the maximum point of the distorted zone and the inherent minimum point of potential normal layer, so their distance can be defined as height, and the extent can be calculated to the enclosure region. Figure 2.12 shows an example.

2.4 Detection of subretinal and intraretinal fluid

We described in Section 2.3 that in the case of AMD, the RPE layer can locally detach from the Bruch's membrane and in many cases fluid can accumulate underneath the retina. This symptom is also observable in diabetic macular edema (DME). The extent of these markers is an important measure of drug efficiency.

Many research focus on automatic detection of subretinal fluid in OCT images. One of the most popular approaches is based on the Split-Bregman optimization technique [31]. This method is used to segment dark regions between layers¹. These localized regions are considered as possible fluid candidates.

Novosel et al. [60] recommended a locally-adaptive loosely-coupled level set method. The approach exploits the local attenuation coefficient differences of layers around an interface to delineate the fluid. This concept can handle abrupt attenuation coefficient variations and topology-disrupting anomalies. SEAD (symptomatic exudate-associated derangements) segmentation in 3D volumes plays an important role in the treatment of neovascular AMD. The accurate detection is challenging because of the large diversity of SEAD size, location and shape. Xu et al. [104] proposed a voxel classification based approach using a layer-dependent stratified sampling strategy to address the class imbalance problem in SEAD detection. A fuzzy level set method was introduced by Wang et al. [99] for identifying fluid-filled regions. They use the combination of three types of scans (two types of B-scans and a C-scan) to generate a comprehensive volumetric segmentation of the retinal fluid. The remaining artifacts are removed by identifying morphological characteristics and vascular shadowing.

Popular machine learning methods also appeared in this field of the research. In [20, 110] a random forest classifier was trained to distinguish true fluid regions from false segments. Lee et al. [49] and Roy et al. [66] also presented a CNN-based method for detecting SRF and IRF. The Roy's neural network model became known as ReLayNet and is freely accessible and usable by anyone.

In this section, we present algorithms that use different approaches to determine intraretinal and/or subretinal fluid. We briefly describe two existing methods that we re-implemented according to the original papers for comparison. It can be observed that literature methods define boundary layers before the segmentation of specific biomarkers.

Wieclawek's [101] normalized the intensities of the input image to the $[0, 1]$ interval, because images can be acquired with different settings, so that their intensity range may vary. The authors used a non-linear filtering to reduce noise effect on images. Besides, they applied a spatial averaging filtering technique which is based on the real product of complex diffusion. The tools of mathematical morphology were used to delineate specific intraretinal fluid areas, based on the observation that IRF appear as darker segments in the images. Among other operations, they used H-minima transform to suppress local minima to denoise areas of almost constant gradient and highlight IRF like black segments on relatively light background. The transform has a single control parameter that value has been fixed experimentally

¹We note that in clinical routine, our medical experts use that image acquisition settings where the retinal layers are brighter than the vitreous body.

to 30% of the maximum brightness in the image. One of the last steps was the binarization of the obtained image with a given threshold value. Since the result may still contain false regions, they filtered all objects that are above ILM and below RPE layers, however, their determination is not detailed.

By contrast, Wilkins et al. [102] investigated the problem from another point of view. They also determined the major layer boundaries and improved the image quality and filtered noise using the combination of a median filtering and a bilateral filtering during preprocessing. Since the final expected result is a binary image with segmented objects, the image was binarized and determined the boundaries of the remaining possible IRF segments and they defined three conditions based on empirical studies. They examined the extent of the objects, the degree of scattering between the intensity of the pixels in the segment, and whether the object is located between the ILM and RPE layers.

In this section, we introduce an approach to detect subretinal fluid using vertical intensity profiles and another one to localize subretinal and intraretinal fluid areas simultaneously.

2.4.1 Segmentation based on projected data

It can be observed that subretinal fluid regions appear as spots with brighter intensity in the image. The zones of the disease and the intensity of the vitreous body of the eye are almost within the same range if distortions are not considered. Firstly, the region of the interest (ROI) of retina was delineated for processing using the method described in Section 2.3.2.

Sampled points of ILM and RPE layers are used to detect biomarkers which was applied for calculation of mean retinal thickness. This value serves as threshold in a further step of the algorithm.

Observations showed that SRF regions are close to the RPE layer and near/in the distorted area. We investigate the vertical stripes, and based on the above assumptions, we look for minimum locations on the horizontal projections that are closest (and the second closest) to the RPE layer. We consider only the profile between ILM and RPE points.

Using the average layer thickness and vessel shadow information, we can filter out the stripes which are less reliable than the others, and fit a smooth curve to the reliably detected minimum locations, thus approximating the fluid surface in the less reliable positions. In Section 2.3.5, we recommended a possible normal outer layer boundary. This can help us to define the degree of creasing of the layers and also supports outlier filtering. These conditions reduce the dataset to a sufficient point, so after this step, we can fit a curve to the detected points and obtain the subretinal fluid region. We summarize our detection method in Algorithm 1.

Algorithm 1 Pseudocode of subretinal fluid segmentation.

```

1: function SRFLOCALIZATION( $PILM, PRPE, N, P$ )      ▷ ILM points  $PILM$ , RPE points
                                                    PRPE, all sampled x coordinates  $N$ 
2:    $C_{SRF} \leftarrow \square$                         ▷ empty list of candidate SRF points
3:    $thick \leftarrow \frac{1}{N} \sum_{i=1}^N | PILM_i - PRPE_i |$   ▷ thickness of retinal layer
4:   for  $i$  in  $N$  do
5:      $ILM_{act} \leftarrow PILM_i$ 
6:      $RPE_{act} \leftarrow PRPE_i$ 
7:      $P \leftarrow$  calculate peaks between  $ILM_{act}$  and  $RPE_{act}$ 
8:     Let  $P_1$  the minimal and  $P_m$  the maximal y coord of a peak
9:     if  $P$  length  $\geq 2$  then
10:       $SRF_{bot} \leftarrow P_{m-1}$ 
11:       $SRF_{top} \leftarrow P_{m-2}$ 
12:      if  $SRF_{top} < thick \cdot 0.75$  then
13:        add to  $C_{SRF}$  points
14:      end if
15:    end if
16:  end for
17:  Investigate  $C_{SRF}$  candidates and remove isolated data (in this case isolated means a
    point without neighbor(s))
18:  Fit a curve to top and bottom data
19: end function

```

2.4.2 Detection and separation using active contour

The reflectivity of the IRF biomarker is similar to the subretinal fluid and appears as smaller oval object that are located near RPE layer. Based on this a priori information, we have developed a method that can simultaneously segment and differentiate subretinal/intraretinal fluid areas.

We used the method presented in Section 2.3.3 to highlight ILM and RPE layers. An edge-preserving anisotropic diffusion filter [63] was used to eliminate various effects of artifacts. The image noise is reduced using the filter and we quantize the grayscale image into five intensity levels. Our observations showed that each layer of the retina has similar intensity, so this operation facilitates the separation of the major retinal layers. During binarization, we keep the brightest points, because we know that the reflectivity of the IRF is similar to that of the vitreous body. In this step, we create a mask for edge-based active contour process. To achieve the appropriate segmentation result, the input parameters of the model were determined based on empirical studies. The intermediate steps of the method is presented in Figure 2.13.

So far we have identified possible important segments and we need to separate intraretinal, subretinal fluid and false segments from each other. This step may be omitted in some cases, when there is no object in the image that could be detected as a pathological mutation. We developed a condition set for filtering fluid regions.

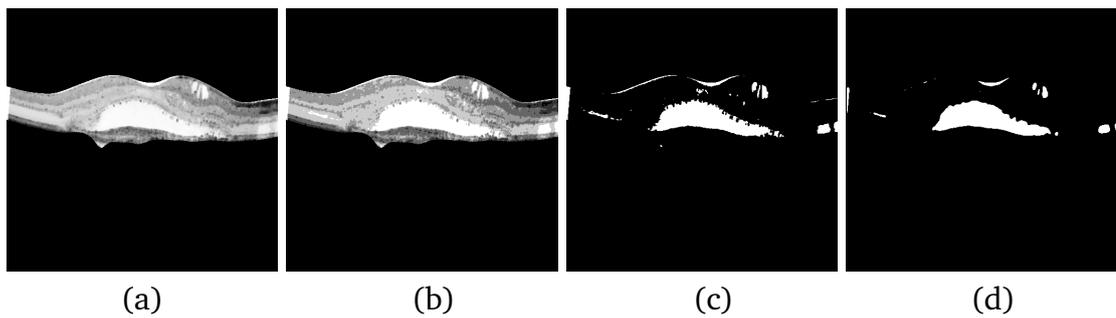


Figure 2.13: Intermediate stages of the active contour based algorithm to detect subretinal and intraretinal fluid. (a) anisotropic diffusion, (b) quantization, (c) binarization, (d) result of active contour.

We test the fulfillment of the three criteria for classification at the object level:

- where the object is located within the retina,
- what is the extent and the shape of the object,
- whether the layer is distorted or not.

In the case of higher distortion of the retina, they may also appear on the left or right hand side. When determining fluid regions, it is also important to examine the cases where there is no creasing in the layer. In the case of IRF, we need to find objects with oval shapes and the observations show that these segments are in the creased zone. Contrary to fluid regions, IRF can be found in higher layers. We distinguish between the symptoms based on these characteristics. False segments may also appear in the image, but these are usually small objects, so we can remove them easily with an area-based filtering.

2.4.3 Evaluation and results

As mentioned in Section 2.4, annotations were not yet available at this stage of the research so we only did visual comparison of the re-implemented methods. However, since the first step of these methods provides ILM and RPE localization it is possible to quantify and compare the results. We used the database that was described in Section 2.3.4.

Firstly, we consider the result of localization of major layers. We calculated the mean, maximum and standard deviation of boundary errors for every surface. It can be observed that the largest error is between 1 and 4 pixels in most cases and Figure 2.9 confirms this statement. Table 2.3 shows the mean error of our algorithm

Table 2.3: Summary of mean, standard deviation and maximum error (in pixels) between manually segmented and automatically detected layers in 7 annotated OCT image sequence. The minimum error of mean, standard deviation and maximum value is highlighted in bold.

		Seq_02	Seq_03	Seq_04	Seq_05	Seq_06	Seq_07	Seq_08	All
Mean	Wilkins	1.90	1.86	2.90	1.88	9.00	1.22	10.91	4.24
	Wieclawek	6.86	10.34	11.20	17.82	6.88	5.42	7.11	9.38
	Proposed	2.01	2.10	1.44	1.96	2.39	2.17	1.83	1.98
Std. dev.	Wilkins	3.34	2.05	2.67	1.62	2.73	0.18	2.93	2.28
	Wieclawek	7.92	5.59	5.92	6.31	5.97	4.78	5.39	5.98
	Proposed	1.56	0.69	0.65	0.80	1.63	0.64	0.65	0.94
Maximum	Wilkins	26	19	17	17	25	11	25	20
	Wieclawek	35	31	24	22	21	20	19	24.57
	Proposed	17	15	15	19	18	15	17	16.57

is less than 3 pixels in all cases that shows this approach is more stable than the re-implemented methods.

For visual comparison of detected biomarkers, Figure 2.14 illustrates segmentation results of the algorithms in some slices. The method developed by Wieclawek detected fewer possible IRF regions, which may be due to the fact that the given threshold only keeps the actual bright points. The disadvantage of this is that important areas may be lost during processing. The other method from literature by Wilkins yields almost the same segmentation results, but in many cases it keeps false objects, because the thresholds are not dynamically defined. In contrary, our method uses dynamic requirements based on a priori information.

We tested the re-implemented methods on images in which subretinal and intraretinal fluid regions may also appear. These approaches can also detect IRF because its reflectivity similarly to the fluid object. Our algorithm can also distinguish these two types of structures from each other as it can be seen in Figure 2.14.

Automatically calculated quantitative descriptors can be visually presented to the medical doctor to aid interpretation of data. The first type of visualization is the traditional slice-by-slice display which provides a good depth context within slice, but no spatial context between slices. In case of colored overlay, anatomical display provides regional context and color encodes quantitative parameters. The masks containing the fluid segments were summarized in each slice and we created a new image with this information whose size was the same as the area examined in the SLO image. Color information was assigned to the values that helps examine the thickness and extent of the subretinal fluid. Figure 2.15 represents restricted subretinal fluid area which results were verified by ophthalmologists and they found the segmentation, quantification and also the visualization technique useful.

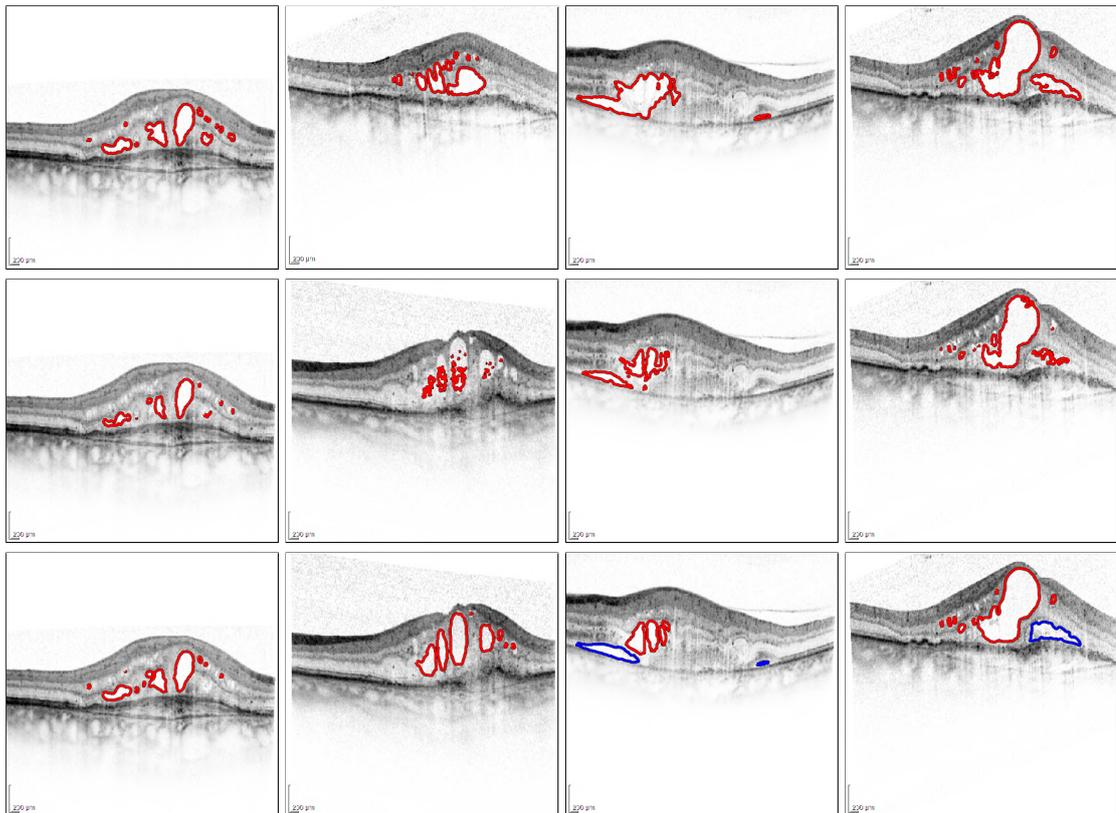


Figure 2.14: Illustration of detected intraretinal (red curve) and subretinal fluid (blue curve) regions by the described algorithms. Rows (from top to bottom): Wilkins et al, Wieclawek, our proposed method. The first two columns contain only IRF, while the last two scan also have SRF areas.

2.5 Determination of PED and SHRM

Automatic detection of abnormal retinal structures remains a challenging task. The retinal morphology and intensity may change acutely resulting from the abnormal structures and the prior knowledge about morphological and optical features clues that are used for normal retinal image segmentation may not be valid. Subretinal Hyperreflective Material, as its name suggests, is likely composed of many components, including fluid, fibrin, blood, etc., and its composition changes over time. So, the reflectivity of SHRM is heterogeneous. In contrast, Pigment Epithelial Detachment has lower intensity, so these can be separated from each other. Nevertheless, in many cases, due to the above introduced reasons, their location and presence are unclear. In some cases, abnormalities are so inseparable that they are managed as one structure.

Significantly fewer literature methods are available to solve this problem and most of them are using a deep neural network. As in retinal layers detection, most

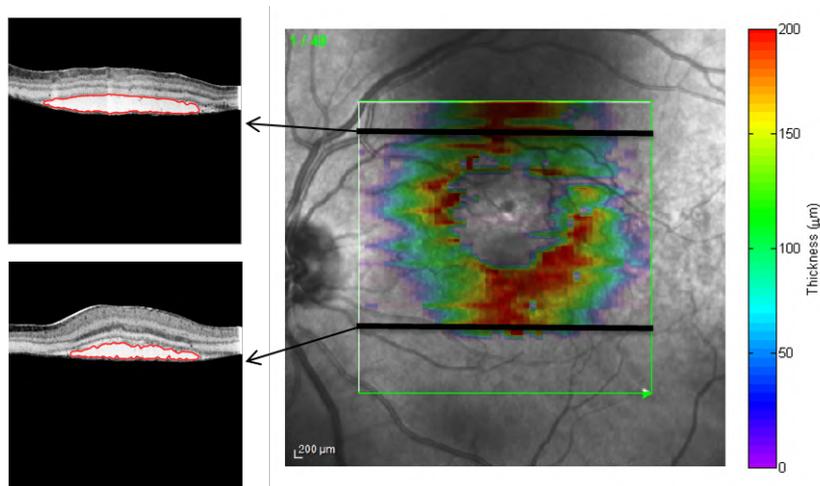


Figure 2.15: Color overlay of the subretinal fluid volume thickness. Left: Red boundaries indicate the subretinal fluid in each slice. Right: The color/hue relates to subretinal fluid thickness.

approaches apply convolutional neural network. Lee et al. [50] constructed a U-net consisting of a contracting and an expanding path. These perform different tasks to help the network to recover detailed spatial information. Xu et al. [105] demonstrated an algorithm that relies on convolutional neural networks (CNN) and deep neural network (DNN) feature extractor network.

However, there are approaches from a completely different aspect for the problem. Ahlers et al. [4] detected the location and measured the volume of PED region using the high-definition OCT tool. Shi et al. [70] introduced a shape-constrained graph cut algorithm that uses foreground and background seed points. They applied AdaBoost method to help remove false segmented regions in the initial segmentation results and mathematical morphology to refine segmentation result.

Our method is also based on graph cut, but the definition of specific segments is determined by different conditions.

2.5.1 Abnormality detection around the RPE layer

To localize abnormalities we use the cluster image which was presented in Section 2.3.3. The pixels of the PED and the SHRM regions are roughly in the same intensity range but in some cases the results of a proper clustering can form a separate region. If these regions are not separable, the medical literature distinguish PED+SHRM label for this cases.

These biomarkers are usually located around the RPE layer, along or near the distortion, so clusters are sampled from these areas. To reduce the clusters to abnor-

mal segments only, we use some a priori information about the position and extent of possible fluid area. To determine the location of distortion and to detect other biomarkers such as intraretinal and subretinal fluid areas the method introduced in Section 2.4.2 is used. So, we only investigate in a specific range over the RPE layer. For this, the threshold was determined as follows:

$$T = \begin{cases} \max(CF_{(x, \max(y))}), & \text{if } \max(CF_{(x, \max(y))}) > 0 \\ RT_x \cdot 0.2, & \text{otherwise} \end{cases}, \quad (2.9)$$

where CF represents the binary image with detected intraretinal and subretinal fluid, x is the actual x position, $\max(y)$ is the actual maximum y position and RT denotes the calculated retinal thickness (from ILM and RPE layers).

All clusters that are above the RPE will be SHRM regions. They may also have amorphous shapes because they are made up of different sub-areas, so their convex shapes is computed to accomplish the final result. The method is summarized in Algorithm 2.

Algorithm 2 Detection and separation of PED and SHRM.

```

1: function PEDSHRM( $GC_{img}$ ,  $CF$ ,  $PRPE$ ,  $DR$ ,  $RT$ ) ▷ cluster image  $GC_{img}$ , fluid
                                                    objects  $CF$ , RPE points  $PRPE$ ,
                                                    distorted region  $DR$ , retinal thickness  $RT$ 
2:    $AR \leftarrow []$ 
3:    $PED \leftarrow []$ 
4:    $SHRM \leftarrow []$ 
5:    $C_{poss} \leftarrow$  get  $GC_{img}$  clusters near  $PRPE$  within  $DR$ 
6:    $T \leftarrow$  threshold value
7:   if  $\max(CF_{(x, \max_y)}) > 0$  then ▷ maximum (x, y) coordinates of objects
8:      $T = \max(CF_{(x, \max_y)})$ 
9:   else
10:     $T = RT_x \cdot 0.2$  ▷ RT value in x position
11:   end if
12:   Add all clusters to  $AR$  that satisfies the condition or below  $PRPE$ 
13:   for  $i$  in  $AR_{CH}$  do
14:     if  $AR_{CH}$  above  $PRPE$  then
15:       add to  $SHRM$ 
16:     else
17:       add to  $PED$ 
18:     end if
19:   end for
20: end function

```

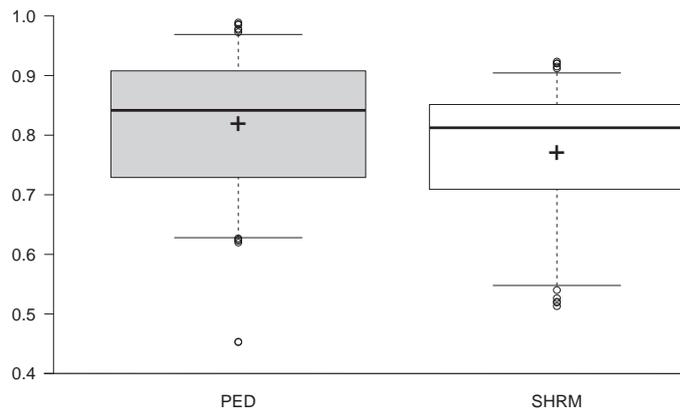


Figure 2.16: Box plots representing the Dice coefficients of PED and SHRM generated by proposed method. In both cases, the mean Dice coefficients were greater than 0.7. Black cross = mean Dice coefficient; empty circles = outliers; PED = pigment epithelial detachment; SHRM = subretinal hyperreflective material.

2.5.2 Evaluation and results

In this case 21 recordings were available for evaluation. The dataset consisted of sequences with 25 and 49 slices. The annotation produced by ophthalmologists were compared with segmented results to characterize the effectiveness. We used the Dice coefficient (Equation 2.1) and Sensitivity (Equation 2.2) metrics.

The location of the possible SHRM has played an important role in determining PED. In many cases, SHRM is not clearly detectable. In spite of all this, as can be seen in Figure 2.16, in both cases, the average Dice coefficients are above 0.75 and the sensitivity is 0.93 for PED and 0.77 for SHRM.

In Figure 2.17 some examples are shown of automatic results and also the manual annotation for a visual comparison.

2.6 Segmentation of outer retinal tubulation

Outer retinal tubulation (ORT) was first described by Zweifel et al. [111] as “hyporeflexive tubular structures with hyperreflective borders within the nuclear layer of the retina” and appear in many retinal diseases (see, e.g. Figure 2.4). Nowadays, in many cases, it is not sufficient for a medical doctor to decide what illness may affect the patient’s eyes, but also the extent, cardinality, etc. of the biomarkers. We did not find any automatic method for ORT segmentation during the literature review.

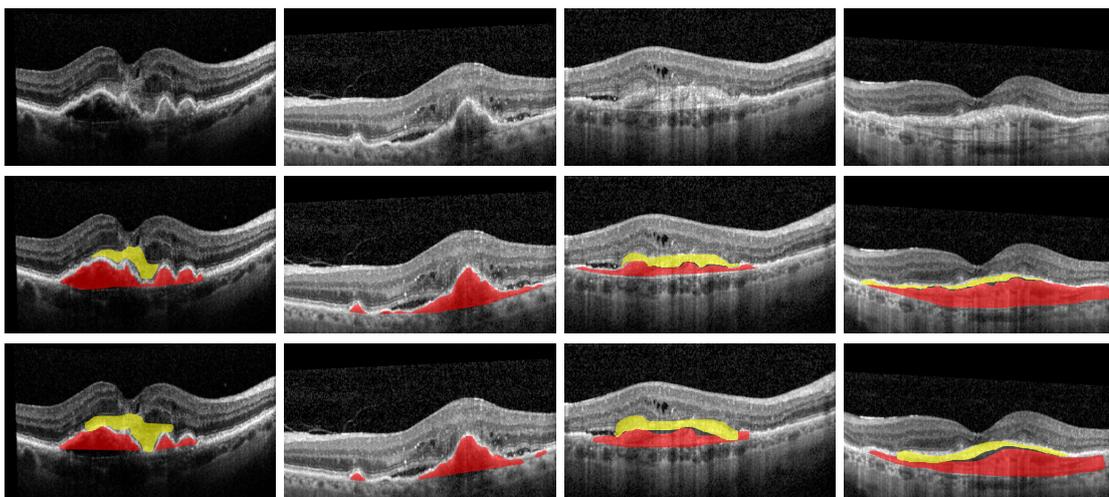


Figure 2.17: Illustration of annotated (in middle row) and detected (bottom row) PED and SHRM. Red regions show PED yellow represents SHRM.

2.6.1 Detection using classical image processing operators

The ORT has hyperreflective contour and contains hyper- or hyporeflective points. Our procedure is based on finding hyperreflective points. The effect of image noise is reduced using Wiener filter with a 3×5 pixel kernel. The input image size is 509×496 , so a smaller kernel is sufficient. This filter size does not smooth the reflective dots but reduces noise. Reflective points are localized using a Hessian detector. Hessian

$$H(x, \sigma) = \begin{bmatrix} I_{xx}(x, \sigma) & I_{xy}(x, \sigma) \\ I_{xy}(x, \sigma) & I_{yy}(x, \sigma) \end{bmatrix} \quad (2.10)$$

is a matrix, where $I_{xx}(x, \sigma)$ and $I_{yy}(x, \sigma)$ are second partial derivatives in the x or y direction and $I_{xy}(x, \sigma)$ is the mixed partial second derivative in the x, y directions. σ denotes standard deviation. After that, the determinant is computed:

$$\det(H) = I_{xx}I_{yy} - I_{xy}^2. \quad (2.11)$$

For filtering out the false points, we use some prior information about the biomarkers. Firstly, we calculate the retina thickness and we keep only those points which are located in the lower third of the retina. Since we know that ORT is close to the RPE layer within the distorted retina region or neighborhood, by estimating the start and end of the distortion on the slices, additional points can be removed. The method uses information about the extent and location of the distortion and the position of intraretinal and subretinal fluid areas applying algorithms described in Sections 2.3.5 and 2.4.2. Limiting the specific extent of the ORT in many cases is

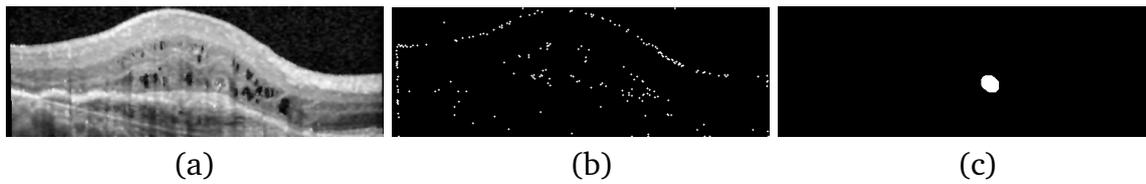


Figure 2.18: Intermediate stages of the simple image processing operations based algorithm to determine ORT structures. (a) Wiener filter, (b) Hessian detector, (c) segmented ORT.

very difficult, because there is no clear distinction between the hyperreflective wall and its surroundings. We perform adaptive histogram equalization in the image so that at least a part of the possible contours become separable by hysteresis thresholding. The two threshold values are given as the lower and upper third of the maximum intensity value in the image. After the last filtering, we keep only those points which are part of an object in the thresholded image. Since we assume that ORT contains hyperreflective points, we compute distance map [18] for the points, which gives how far another point is to a given point. The map is thresholded using a constant value which was determined empirically and finally the convex hull of the objects is computed. Figure 2.18 illustrates some steps of the algorithm.

2.6.2 Localization with neural network

The method presented in the previous section is not robust against more diverse images thus we have developed a more general approach. Nowadays, deep learning based methods are very popular and are widely used solutions for different tasks. These differ from conventional networks in that these may have several hidden layers between the input and output layer. Nowadays, in most areas of research, a variant of a deep neural network is used to solve a problem such as in image processing or language processing. For semantic segmentation tasks the state of the art solution is using convolutional neural networks (CNN). This is a constantly evolving area so new solutions are arising day by day.

It is well-known that deep neural networks need as much training data as possible to achieve better results. In our case 9 sequences of 25 or 49 slices were available which were annotated by two medical experts, overall 16 sequences. We did not use annotated images for training that were not marked with ORT because our aim was that the network learn ORT object-like segments, since their precise localization is less important. For images which have annotations from two experts, the masks were combined using logical OR operator. Hence the masks contain the union of the these marked images. Here the available amount of training data is extremely small, 132 examples, but machine learning techniques profit from more training data and more data can help reduce overfitting thus we use real time augmentation technique during

training. The details of transformations used for data augmentation is described in Table 2.4.

To save computation time, the images were resized to have height and width of 220×256 pixels, respectively. The intensities were normalized to be in range $[0, 1]$ and then standardized to have zero mean and unit variance. The mean and the standard deviation were estimated using the training corresponding set only.

Since our dataset contains sequences from only nine patients, we use nested cross-validation for hyperparameter tuning and measuring the test error. The outer K-fold validation, where $K = 9$, estimates the generalization error and the inner one used for finding the best model settings using $K = 8$. For evaluating one model requires $9 \cdot 8 = 72$ training and evaluation. Note that in this scenario one fold contains the images of one patient.

Here, we consider a slightly modified version of the original U-Net architecture [65] for detecting ORT. There are three differences:

- we reduce the number of layers to speed up training time,
- we double the number of filters in each layer to give larger capacity to the model, since the amount of data augmentation and retorsion in the original images are high, hence we expect the model to be invariant to this transformation,
- to prevent overfitting, we use Dropout [73] between all convolutional layers except the last two ones.

The input is the whole OCT image, hence the images are not splitted into patches with size 32×32 and trained on these subareas as in the basis U-Net architecture. We use stochastic gradient descent (SGD) for optimizing the loss function with one sample per update. The objective function that SGD minimize the Dice loss is

$$\text{DiceLoss}(p, y) = 1 - \text{DiceCoefficient}(p, y), \quad (2.12)$$

where p is the prediction and y denotes the mask label.

Table 2.4: Transformation details used for real time data augmentation to raise training data.

Type	Range
rotation	$[-10^\circ, 10^\circ]$
horizontal mirroring	[yes, no]
shearing	$[0, 0.1]$
vertical and horizontal shifting	$[-0.1, 0.1]$

We use this because our dataset is very unbalanced and previous work [23] shows that the Dice loss in Equation 2.12 performs better in such cases than weighted binary cross entropy which is the most used loss function in a binary classification problem. The exact model architecture is presented in Figure 2.19.

In order to decrease the generalization error, the amount of Dropout is tuned. The average Dice loss for the validation sets can be seen in Table 2.5 for varying dropping rates. The implementation was done in Keras² and the training were running in parallel on four GPUs.

The results in Table 2.5 show that there is no unique parametrization which perform well on every fold. Probably a larger dataset would give a more stable solution. However in most of the cases the 0.4 Dropout rate gives the highest validation Dice coefficient.

Note that the estimation of the test error in every outer fold was done always with the settings that has the highest mean Dice coefficient in the inner cross validation folds and it was retrained with the corresponding hyperparameters on the whole inner dataset for all 8 fold. The used Dropout rates which gave the best results in the inner folds are highlighted in Table 2.5.

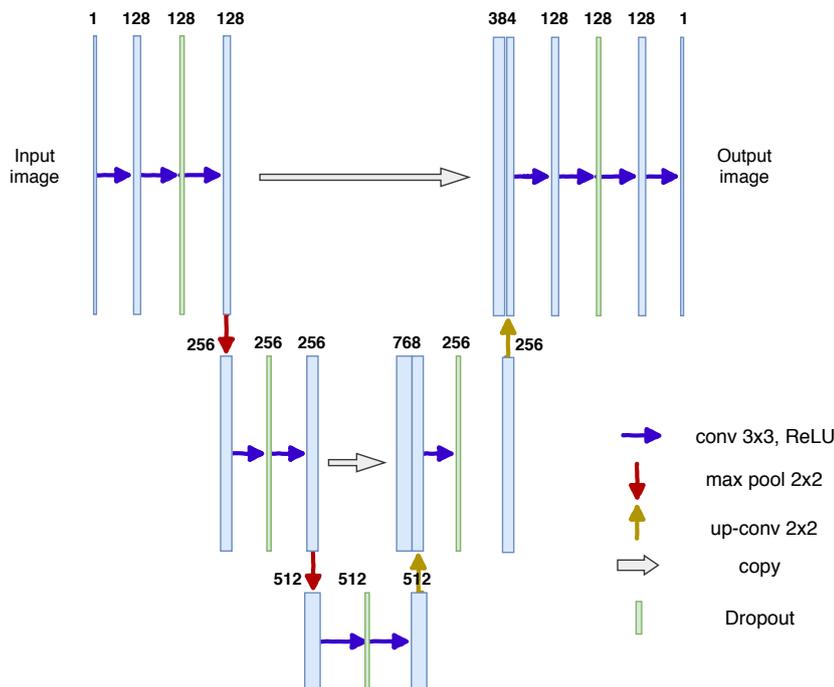


Figure 2.19: Fully convolutional U-Net based architecture. To speed up training time, the number of layer are reduced, filters are doubled, Dropout is used between all layers except the last two ones.

²<https://github.com/keras-team/keras>

Table 2.5: Average Dice coefficients on validation set. K is fold number, r is the dropping rate for Dropout. The best average Dice values are highlighted in bold.

r	0	0.1	0.2	0.3	0.4	0.5
K=0	0.566	0.570	0.597	0.605	0.614	0.586
K=1	0.620	0.599	0.590	0.605	0.612	0.607
K=2	0.519	0.510	0.530	0.515	0.554	0.529
K=3	0.544	0.555	0.583	0.576	0.581	0.517
K=4	0.531	0.553	0.564	0.549	0.589	0.578
K=5	0.546	0.563	0.569	0.564	0.582	0.536
K=6	0.600	0.617	0.627	0.626	0.612	0.594
K=7	0.545	0.556	0.571	0.585	0.553	0.559
K=8	0.547	0.554	0.586	0.553	0.584	0.559

2.6.3 Evaluation and results

In order to achieve widespread validation, we calculated various metrics both for automatic and for manual segmentations. ORT biomarkers appear in variable numbers in images and in many cases only in small size. The marked segments make up only 0.3% of the the whole dataset.

We also measured the average Dice coefficient in object level since there are images with multiple ORTs. When we calculate global and object level Dice coefficient, we considered only those images which contain segmentation.

Dice coefficient measures well how precise a segmentation is, however our main goal is not a perfect segmentation of ORT, only the cardinality of the objects. Since variation of the Dice value among the doctors was moderate, there were cases when either doctor annotated the actual area as ORT. To fulfill this goal, in addition to the Dice coefficient we measured the recall of the ORT objects. These metric describes how many annotated ORT objects were found in the segmentation.

Table 2.6 shows the achieved Dice and Recall values on K-fold validation, the rows are the results of the medical doctors against each other, segmentation result. In case of MD_1 versus MD_2 and our modified U-Net architecture against the union of the masks of MD_1 and MD_2. The results of the selected models can be seen in Table 2.7. The average of the global Dice coefficients on the outer folds is close to the validation scores described in Table 2.5. Hence our model does not overfit the data and performs similarly on unseen examples.

The object level Dice coefficient comparing to the union is better than the global one, which means that where ORT objects were present the model can segment them well. The object level Recall score also supports this statement. In average, our object level Recall reaches the performance of doctors with 0.847 as it can be seen in

Table 2.6: The evaluated performance of the doctors and the proposed convolutional network on K-fold validation.

	Recall (object)	Dice (global)	Dice (object)
MD_1	0.908	0.682	0.665
MD_2	0.812	0.625	0.590
modified U-Net	0.847	0.579	0.583

Table 2.7: Recall and Dice object level and global scores for each sequences.

Patient ID	Recall (object)	Dice (global)	Dice (object)
Seq_1	0.714	0.420	0.412
Seq_2	0.642	0.347	0.390
Seq_3	0.900	0.715	0.707
Seq_4	1.000	0.720	0.742
Seq_5	1.000	0.552	0.626
Seq_6	0.931	0.602	0.627
Seq_7	0.500	0.353	0.361
Seq_8	0.933	0.751	0.664
Seq_9	0.925	0.579	0.633
All	0.847	0.579	0.583

Table 2.6. That means, our model is capable of helping the doctors to identify ORT in OCT images. In Table 2.7 the average scores can be found for each patient. In some cases the model found every annotated objects, and even in the worst case at least half of the objects.

Some example segmentation can be seen in Figure 2.20. The images show that in average cases the model perform quite well both in precise localization and finding every objects. However there are some cases when the Dice coefficient is lower. We can experience this for one patient that has elongated objects which made the Dice coefficients worse. Besides the model segments multiple subregions of the original objects so a professional could easily improve it. Note that the most of the patients have other biomarkers like IRF and SRF. In some cases we experienced that the model segmented the IRF as ORT. The reason behind could be that these markers' reflectivity and shape is very similar in most cases. Training a model to recognize these markers as well probably would improve the capability to distinguish these objects.

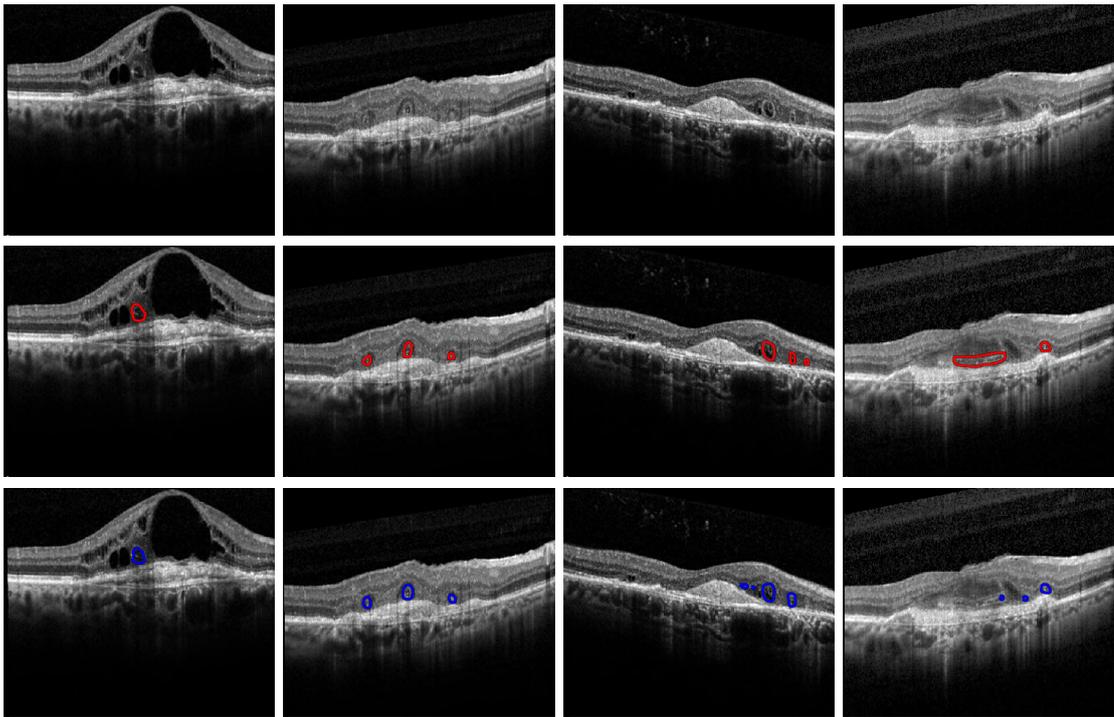


Figure 2.20: Illustration of annotated (in the middle row) and detected (bottom row) ORT biomarkers. The first row contains the original input images. The columns belong to patients Seq_6 (average), Seq_8 (best), Seq_2 (worst), Seq_2 (worst), respectively.

2.7 Concluding remarks

In this chapter numerous approaches are presented to detect biomarkers in AMD and possibly other eye diseases in OCT scans. Two methods were developed to highlight ILM and RPE layers. We have introduced a method that can simultaneously locate and separate the SRF and the IRF and another one to detect and distinguish the PED and SHRM segments from each other. Besides, we segmented ORT using a fully convolutional U-net based architecture.

We computed measures to quantify features of OCT in point of wet AMD patients. This automated localization method can help the quantitative assessment of the OCT biomarkers by providing automatic tools to detect abnormalities and to describe by objective metrics the current state and longitudinal changes during disease evolution and treatment.

Chapter 3

Quantitative Analysis of Irradiated Zebrafish Embryos

Research is ongoing in many areas of life sciences and one of the most important questions is how to treat cancer more effectively. Radiotherapy is one of the most common methods to treat different cancer cells in clinical application despite having harmful effects on healthy tissues. Radiobiological experiments are very important to determine the irradiation-caused acute and chronic effects to define the exact consequences of different irradiation sources. Photon irradiation has been used on zebrafish embryos, a very new *in vivo* and appropriate model system in radiobiology. After irradiation, dose-dependent morphological changes are observable in the embryos.

The advantage of using zebrafish vertebrate model is the biological endpoint assessment on large number of embryos, enabling high power statistical analysis. However, the more sophisticated morphological measurements — beyond the simple survival detection — are extremely time- and labor-intensive. The evaluation process for one study containing hundreds of images might take several weeks for the biologists to complete, taking into account that one image measurement takes ca. 15 minutes. Besides the analysis of the morphological changes, other tasks also occur during processing including the sorting out images of too low quality or not containing a living fish, and improving image quality.

In this section, we introduce a framework to reduce human effort for faster experimentation. The various shape characteristics of zebrafish is obtained in order to give a quicker picture of the response of fish to irradiation. A procedure is presented that is capable of delimiting the area of the zebrafish, localizing the eye or the eyes, determining their diameter, the distance between the head and tail of the fish, and giving the position of the fish. The results are compared with manually annotated data by biologists [91].

3.1 Introduction

Ionizing radiation (IR) is successfully used in both adult and paediatric patients with various primary and metastatic tumours [42, 48]. Approximately 50% of all cancer patients are subject to radiotherapy during the course of their illness with an estimation that radiotherapy contributes to approximately 40% towards curative treatment. Although photon beam therapy is frequently used in the locoregional treatment of cancer, it has also detrimental effects, since radiotherapy induces DNA damage and cell death in tumor cells, but can also induce carcinogenesis in the surrounding healthy tissue of the tumor. Ionizing radiation interacts with matter by excitation and ionization of molecules, producing free radicals and subsequently reactive oxygen and nitrogen species which can attack cell membranes or break chemical bonds in biological molecules, leading to oxidative stress or DNA damage [40].

Zebrafish (*Danio rerio*) embryos have recently been introduced as a novel vertebrate research model for various human diseases and treatments [74]. Zebrafish are excellent model for experimental human cancer research, as they have many key genes involved in cell cycle, oncogenesis, tumor suppression, and DNA repair [55]. Embryo development is extremely rapid during the first few days post-fertilization while the embryos and larvae are transparent, giving the possibility to study the *in vivo* organ development [5]. These features make this animal model appropriate to investigate the effects of ionizing radiation on zebrafish development [30], and this model provides an interim step between the *in vitro* cell culture and rodent systems.

Typically, wild-type adult zebrafish is mated in embryo collection tanks and viable embryos are sorted by microscopic observation (1 embryo/well of standard polystyrene microplates) and maintain under normoxic conditions during the experiments. Figure 3.1 shows an observation setup where the monitoring phase of the embryos is started after 24 hours post fertilization (hpf). Embryos are irradiated with a kind of radiation (e.g., photon) in the specified dose range as determined by the study. After irradiation, changes in the embryos are monitored and acquired in specific intervals for later evaluation. Imaging is supported by the fact that the plate has a known geometry, but there are several problems that can affect the quality of captured microscopic images.

During the imaging, the plate is illuminated from below and each well can be positioned using a stepper motor to take microscopic images. Due to the imprecision of the stepper motor, parts of the nearby well may also be visible, and distortions as a result of uneven illumination may also be observed. Subject distance may vary in different examinations, or even minor variations may appear within a sequence, thus calibration is required for pixel \rightarrow mm conversion.

Irradiation-caused DNA damage causes observable morphological changes in the zebrafish embryos, such as spinal curvature, shortening of the body length, yolk

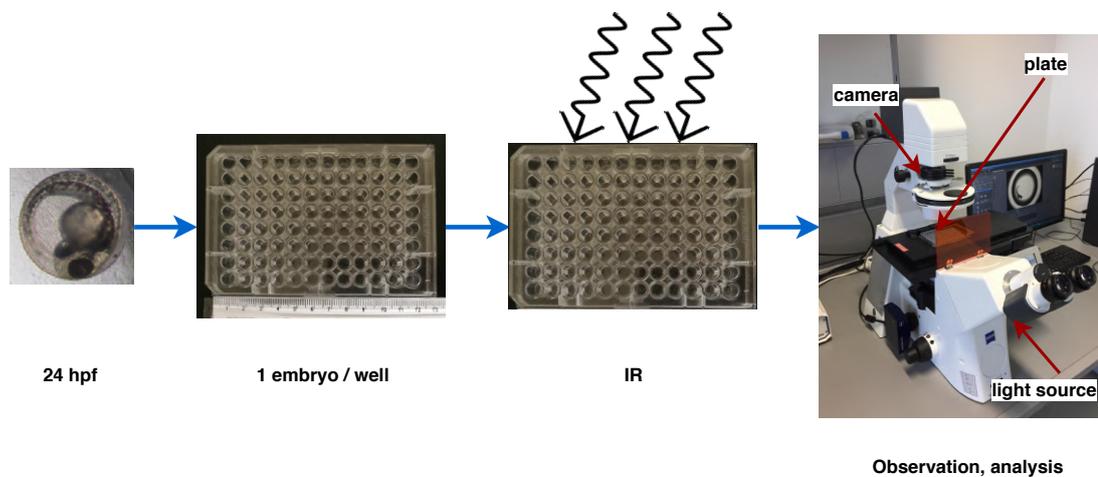


Figure 3.1: Experimental protocol of a zebrafish irradiation study. The 24 hpf zebrafish embryos are sorted in wells and radiated with a type of beam. The morphological changes are monitored daily using a microscope. Representative photomicrographs are taken with a camera.

sac or pericardial edema, abnormality of the eyes (microphthalmia) or abnormal development of the head (microcephaly). The severity of these abnormalities are dose-dependent (see Figure 3.2).

Approaches in the literature mainly aim the detection of zebrafish. Wu et al. [103] used a hybrid active contour model to localize fish. The algorithm of Zhao et al. [109] is based on graph representation. They viewed a graph as a collection of histograms and the main process is similar to “bag-of-words” model. Other publications exist about tracking or detecting zebrafish on videos. Wang et al. [100] used a Gaussian mixture model to tracking multiple zebrafish. Pylatiuk et al. [64] detected heartbeat by digital motion analysis. Ishaq et al. [39] dealt with the deformation of zebrafish tail for drug screening. This process is based on a redefined medial axis generation. However, our literature survey did not reveal any systems that supported biological research by automatically extracting quantitative characteristics and providing morphological analysis.

3.2 Quantitative analysis of morphological changes

The morphological deteriorations introduced in Section 3.1 were measured by biologist researchers taking three weeks of manual work, which was an extremely time demanding process.

The automated framework to monitor irradiated zebrafish embryos involves many tasks to be solved. Due to the varying quality of the images and uneven lightning,

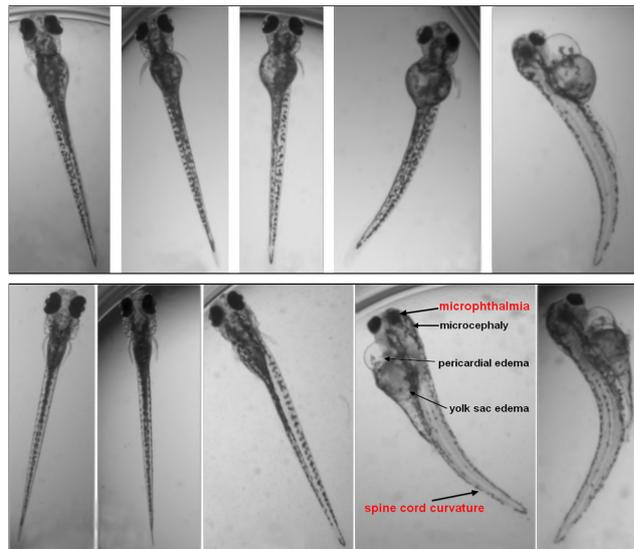


Figure 3.2: Dose-dependent morphological changes after 3 days (top row) and 4 days (bottom row) of (left to right) 0 Gy, 5 Gy, 10 Gy, 15 Gy, and 20 Gy photon irradiation. Significant morphological changes are observable regarding the microphthalmia, pericardial edema, yolk sac edema, microcephaly, spine cord curvature.

preprocessing is required. The acquired images may contain the embryo and in most cases partial regions of the circular well area. To delimit the search region and to be able to make metric measures, the perimeter of the well area must be determined. In order to obtain different morphological features from deformed fish, it is also necessary to segment them and to detect the different anatomical areas to be observed.

In this section, we present a method to detect the abnormal development of zebrafish embryos, to save time for researchers while keeping the consistency and accuracy.

3.2.1 Preprocessing and delimitation of the well

At least some part of the wall of the well is expected to appear in the zebrafish images so in the preprocessing phase, the ROI needs to be limited to the region where the fish might be located. This is important as it makes it easier to distinguish the fish from the segmented false objects and reducing the size of the original input image to the well area that also reduces runtime. The shape of the well is circular, so Hough transformation¹ is applied to the input grayscale images in order to locate wells. We have some a priori information about the size of the image and the possible size of the well appearing, so we can only look for circular objects within a specified radius range. Figure 3.3 shows highlighted well contours in red.

¹As implemented in OpenCV4.

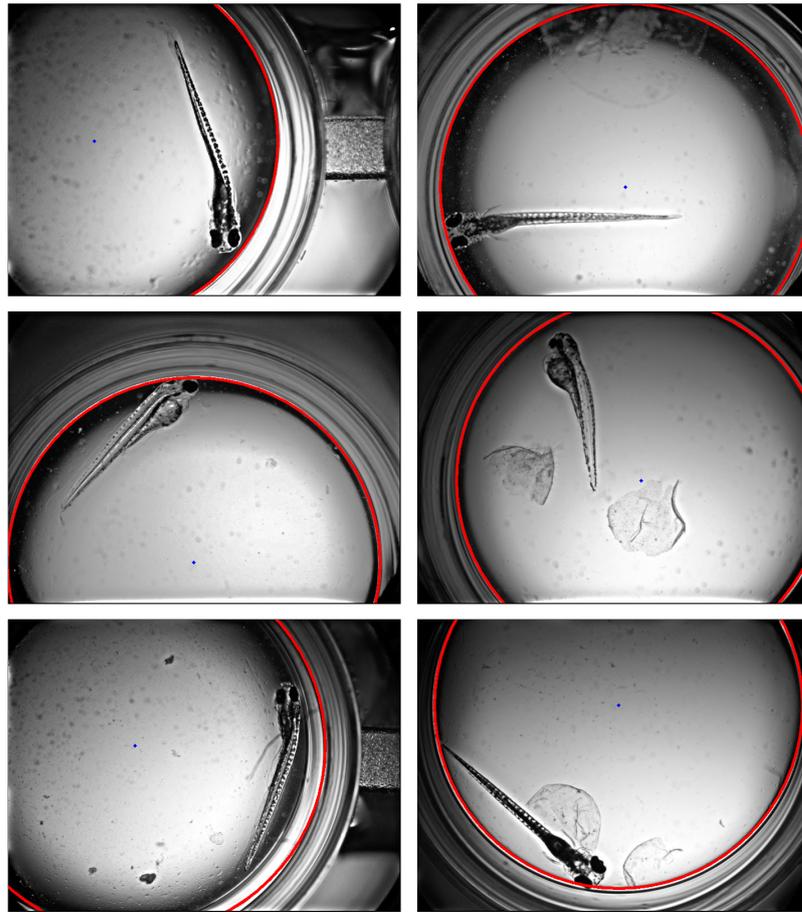


Figure 3.3: Illustration of delineated well wall using Hough-transformation. Red arcs denotes the determined perimeter of the well. The images can contain many circular objects of different intensity levels around the wall making detection of the wall more difficult.

Meniscus effect appears close to the surface of the well due to the curve in the upper surface of the liquid. Thus, illumination correction is applied to reduce this effect [58]. Besides, to improve image quality, the enhancement method from [25] is used.

3.2.2 Segmentation of zebrafish embryo

During the automated segmentation process, we continuously refine the region where the fish could be found. The fish is located within the well area, which may even contain the shell parts of the embryo or other debris present in the fluid. A range filter is used for edge enhancement as a further preprocessing step where each output pixel contains the difference of maximum and minimum value in a 3-by-3 neighborhood

around the corresponding pixel. A binary image is produced using adaptive thresholding after this step.

The embryos are often located close to the perimeter of the well so that their regions can merge during segmentation. Thus, the image is divided into 2×2 equal parts and the average width of the components is computed in each window. To refine the well mask, we dilate each ROI with a structure element of circular shape where its size depends on the previously calculated average value. This mask is used to separate fish and well wall objects from each other. After masking the perimeter of the well and the morphological opening, these areas are reduced in size.

This may segment more objects, thus morphological opening is used to remove the smaller, unnecessary parts. If there are more than two objects in the image, the two having largest areas are kept for further processing. We attempt to filter out false objects from the image by using the following conditions:

- if the eccentricity of object is greater than 0.7 (0 means circle, 1 is a line segment),
- all segments whose area was smaller than 10% of the detected well circle area,
- the major axis length of the object is greater than $\frac{1}{5}$ of the image width.

The intensity of the embryo and other parts of the well zone may be similar, thus before the detection of different parts of the embryo, the bounding box of the object is computed to extend the potential fish region. We used unsharp masking to enhance edges and adaptive thresholding for binarization. Only the largest segment is kept for the post-processing step. At least one eye of the fish is expected to be visible in the images, so the final filter step is to determine whether the object has an eye-like region. We introduce the method in Section 3.2.3. Figure 3.4 illustrates the major steps of the proposed localization method.

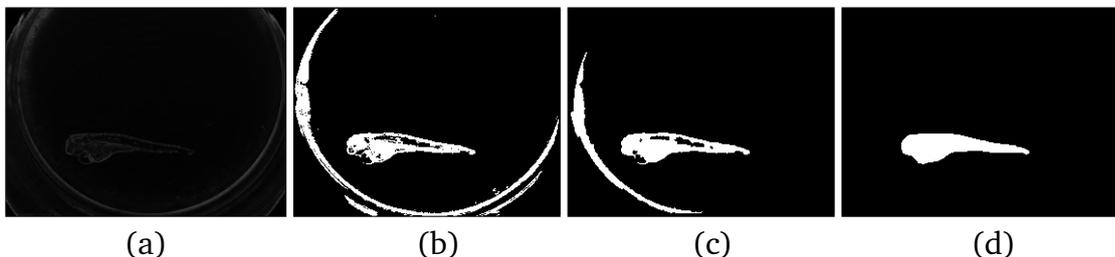


Figure 3.4: Major steps of the zebrafish embryo localization. (a) filtered image, (b) adaptive thresholding, (c) removed possible well zone, (d) segmented fish-like object

3.2.3 Extraction of morphological deteriorations

The morphological changes relevant to radiobiological studies (see on Figure 3.2 in red) are determined after localization of fish. For uniform handling and conditions, all images were rotated by 90° , when the height of bounding box of the fish region is larger than its width. Thus, all examined ROIs are in the same position for further investigation. To determine the tail and head part of the embryo, we divide the embryo ROI vertically and the $\frac{w}{16}$ region of image sides were not taken into account. For the remaining right and left hand side of the ROI, we calculate the mean distance as

$$\text{MeanDist} = \frac{1}{n} \sum_{i=1}^n |z_i - y_i|, \quad (3.1)$$

where n denotes all columns between $\frac{w}{16}$ and $\frac{w}{2}$, z and y denotes object points of the fish. Figure 3.5 represents the detailed determination of the head and tail of the embryo. It is also clear from the figure that the calculated average distance of the ROI part containing the head will be larger, so in each case this side is tagged as head part.

Defining the head of a potential fish makes it easier to locate potential eye regions. The purpose of the previous localization step was to delineate the fish region. The average intensity of the eye region is lower than that of other parts of the the embryo intensity due to the transparent and lit body of the fish. Therefore, we further refine the head part segment at this stage looking for eye-like objects. The information about the intensity of eye regions is used for the adaptive thresholding step. The segmented objects are characterized by circumference and distance from the contour

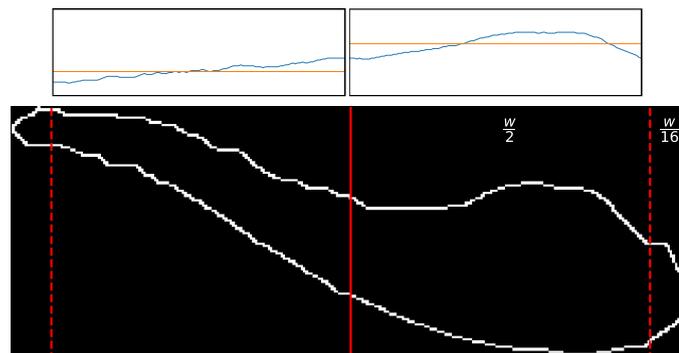


Figure 3.5: Illustration of measures of head, tail determination. w denotes the width of the ROI. Plots represent the distance between contour points for each column. Blue curve denotes the measured values and the orange one the average distance. Higher average value belong to the head part of the embryo.

of the head. If eccentricity of an object is greater than 0.1 and the distance is greater than a threshold, then it is eliminated. Usually we obtain one or two remaining objects, but if there were more, the two largest areas are retained. At higher doses of radiation the fish tend to be oriented its side upwards. This makes the eye regions partially or fully overlap in the microscopic image. Partial overlap must be taken care of to be able to get the correct eye diameter value. Since the eye is a circular object, Hough transformation is used for detection. The part of the detected eye region that is outside of the Hough circle is discarded. Besides, this eye related information also can be used to predict the orientation of the fish.

If the fish is located too close to the wall of the well, the eyes of the fish might not be localized successfully at this stage. Observations showed that in cases of higher doses of radiation the fish head is more likely located towards the wall, so the task is not straightforward considering the dark eye regions and the darker appearance of the fluid due to the Meniscus effect. For this reason, if the results does not contain at least one eye region, the width and height of fish ROI is extended using the object and well wall distance and the fish localization method is run again. The detailed eye detection pseudocode is shown in Algorithm 3.

Extraction of various shape features are obtained by determination of the head and tail parts and the eye region. As next step, the two endpoints of the embryo is determined. The skeleton of the fish object is computed using a thinning method [46] to determine the length of the total body and the endpoints of the embryo. The skeleton may contain spurs that are removed using an iterative method. In every iteration, the shortest spur is removed until the skeleton has only two endpoints. To measure the total body length, the half distance pixel between the eyes or the eye and contour of the fish is used as a skeleton point. A spline is fitted to this point and the previously defined data. The endpoints of the fish are defined as the intersection point of the fish contour and the skeleton.

3.3 Evaluation and results

Investigators irradiated 24 hpf zebrafish embryos with 5 Gy (n=35), 10 Gy (n=35), 15 Gy (n=35) and 20 Gy (n=35) doses using photon beam. Embryos in the control group (n=35) had the same treatment without irradiation. Embryos were irradiated with 6 MV photons, generated by a clinical linear accelerator. The embryos were placed in a well within a 96-well plates with 200 μ l embryo medium, the plates were inserted into a water filled phantom and homogeneously irradiated. The doses were delivered with horizontal beams. Representative photomicrographs were taken with AxioCam MRm at a Zeiss Axiovert 40 CFL (Zeiss, Germany) microscope on the third and fourth day after irradiation to determine the embryos' morphological changes and to measure the eye and spine cord perturbations. According to a pilot study [5],

Algorithm 3 Pseudocode of eye scanning method.

```

1: function EYEXAMINATION( $F, M, B$ )    ▷ grayscale image  $F$ , mean intensity of possible
                                         eye object  $M$ , possible eye mask  $B$ 
2:    $P_{x,y} \leftarrow$  the nearest horizontal perimeter point of well
3:    $C_{x,y} \leftarrow$  centroid of pre-segmented eye object
4:    $dist = |P_{x,y} - C_{x,y}|$ 
5:    $F_{inc} \leftarrow$  increased image size with  $dist$  value both in horizontal and vertical direction
6:    $F_{diff} = |F_{inc} - M|$ 
7:    $F_{dist} \leftarrow$  distance map of  $B$ 
8:    $F_{bin} \leftarrow \emptyset$                                      ▷ empty logical image
9:   for  $i, j$  in all pixels do
10:    if  $F_{diff}(i,j) < th_o$  or  $F_{dist}(i, j) < th_w$  then           ▷  $th_o, th_w$ : threshold values
11:       $F_{bin}(i, j) \leftarrow 1$ 
12:    end if
13:  end for
14:  Hough transformation on  $F_{bin}$  image
15:  if detected circles = 1 then
16:     $H_{mask} \leftarrow$  dilated mask of detected circle
17:    if  $H_{mask} \cap F_{bin} < th$  then                               ▷  $th_a$ : threshold
18:      separate  $F_{bin}$  objects from each other along  $H_{mask}$ 
19:    end if
20:  else if detected circles = 2 then
21:    continue
22:  else
23:    Keep two largest objects
24:     $H_{mask} \leftarrow$  dilated mask of detected circle
25:    separate  $F_{bin}$  objects from each other along  $H_{mask}$ 
26:  end if
27: end function

```

these days are the most relevant time-points in the present experimental setup to determine the abnormal morphological deteriorations. In this study, 393 images were selected from the 525 manually discarding images of poor quality, containing empty well regions or perished embryos.

Embryos were observed without any manipulation in place in the microplates. Morphology was assessed visually and photo documented. The size and shape of the embryo, the spine and the eyes were monitored continuously in the proportion of the living embryos.

Biologists measured manually the morphological changes after three and four days of irradiation to define numerically the dose-dependent changes using ImageJ (Image Processing and Analysis in Java) Software². This application is considered as one of the most appropriate program for different morphological evaluations in cells

²<https://imagej.nih.gov/ij/index.html>

or in zebrafish model. The measured distance is the straight line length in pixels from the head tip to the end of the spine. During the manual measurements, researchers marked the two endpoints of the embryos as a parameter of the body shortening and the furthermore spinal curvature and the diameter of the eyes, what refers to the microphthalmia. Five researchers independently determined the morphological changes, the whole manual processing took 3 weeks.

Data analysis was performed using SigmaStat for Windows statistical software package³. The differences between the manual and the automatically detected measures were analysed applying paired t-test. Mean (M) values and standard deviation (SD) values are given in all reported figures and tables. The statistical analysis shows that there is no significant difference between manual annotations and the values computed by the proposed automatic algorithm. Biologists found significant morphological deteriorations after delivery of dose ≥ 10 Gy, 15 Gy and 20 Gy including a reduction in the distance between the two endpoints of the embryo (Figure 3.6), microphthalmia (Figure 3.7). Table 3.1 presents the measured values.

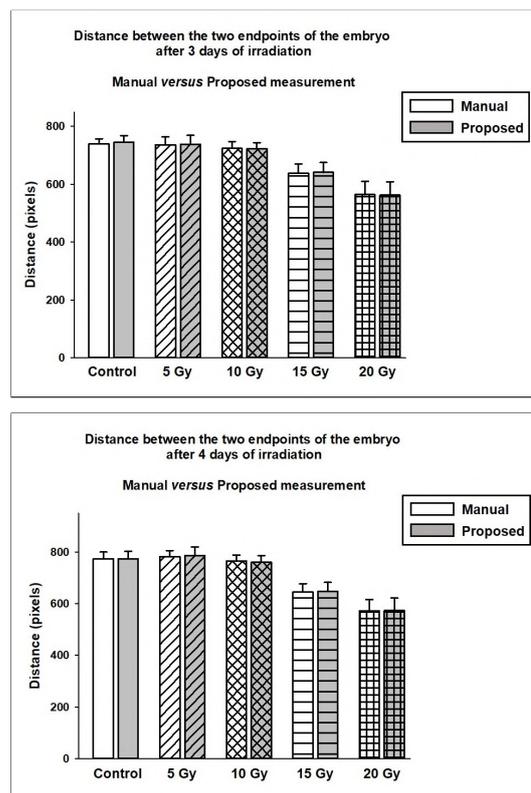


Figure 3.6: Parameter of the length and spinal curvature after three and four days of irradiation.

³Jandel Scientific, Erkrath, Germany

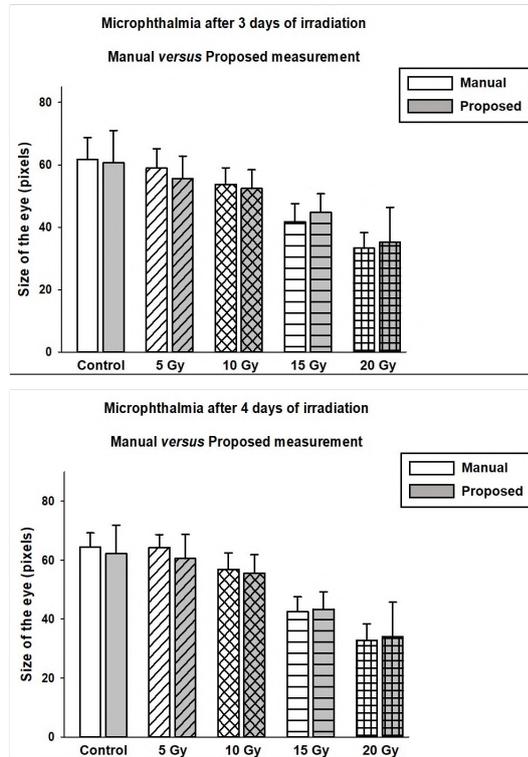


Figure 3.7: Microphthalmia after three and four days of irradiation.

Table 3.1: Results of manual and automatic measurements (mean \pm standard deviation, in pixels).

		Head-to-tail distance		Eye diameter	
		3 days	4 days	3 days	4 days
Control	Manual	739.6 \pm 16.9	774.3 \pm 25.7	61.7 \pm 7.1	64.5 \pm 4.7
	Proposed	744.7 \pm 22.9	773.4 \pm 28.6	60.6 \pm 10.3	62.2 \pm 10.0
5 Gy	Manual	736.3 \pm 27.0	782.6 \pm 23.0	58.9 \pm 6.2	63.6 \pm 4.0
	Proposed	737.6 \pm 31.3	785.8 \pm 34.0	56.7 \pm 6.3	60.9 \pm 8.0
10 Gy	Manual	723.7 \pm 24.1	764.8 \pm 23.4	53.7 \pm 5.3	56.9 \pm 5.6
	Proposed	722.6 \pm 21.1	760.1 \pm 25.1	52.4 \pm 6.0	55.5 \pm 6.3
15 Gy	Manual	637.3 \pm 32.6	645.9 \pm 31.6	41.7 \pm 5.8	42.5 \pm 5.0
	Proposed	641.0 \pm 34.0	647.8 \pm 34.9	44.7 \pm 6.0	45.0 \pm 10.3
20 Gy	Manual	563.8 \pm 45.5	572.2 \pm 43.8	33.3 \pm 5.0	32.8 \pm 5.6
	Proposed	562.0 \pm 46.6	573.6 \pm 48.2	33.8 \pm 8.6	34.1 \pm 11.7

Table 3.2: Execution time of stages of the algorithm for all images expressed in seconds (mean values \pm standard deviation).

	Well localization	Zebrafish detection	Feature extraction	All steps
Execution time	2.8 ± 0.54	2.06 ± 0.42	0.57 ± 0.48	4.92 ± 1.08

The automatic method can save a tremendous amount of time for biologist researchers while providing comparable results. The average execution time of our reference implementation takes about 5 seconds per image. Table 3.2 shows execution times of the stages of the algorithm.

The well area detection using the Hough transformation was successful for almost all the cases. Only 7 of the 393 cases exhibited larger distances between the detection and the real contour. The bottom row of Figure 3.3 illustrates the cases with the largest errors and Figure 3.8 shows segmentation results of the algorithm for some selected cases.

Top row shows correctly localized zebrafish embryos and their extracted features. The automated method might perform poorly in cases where the body of the embryo

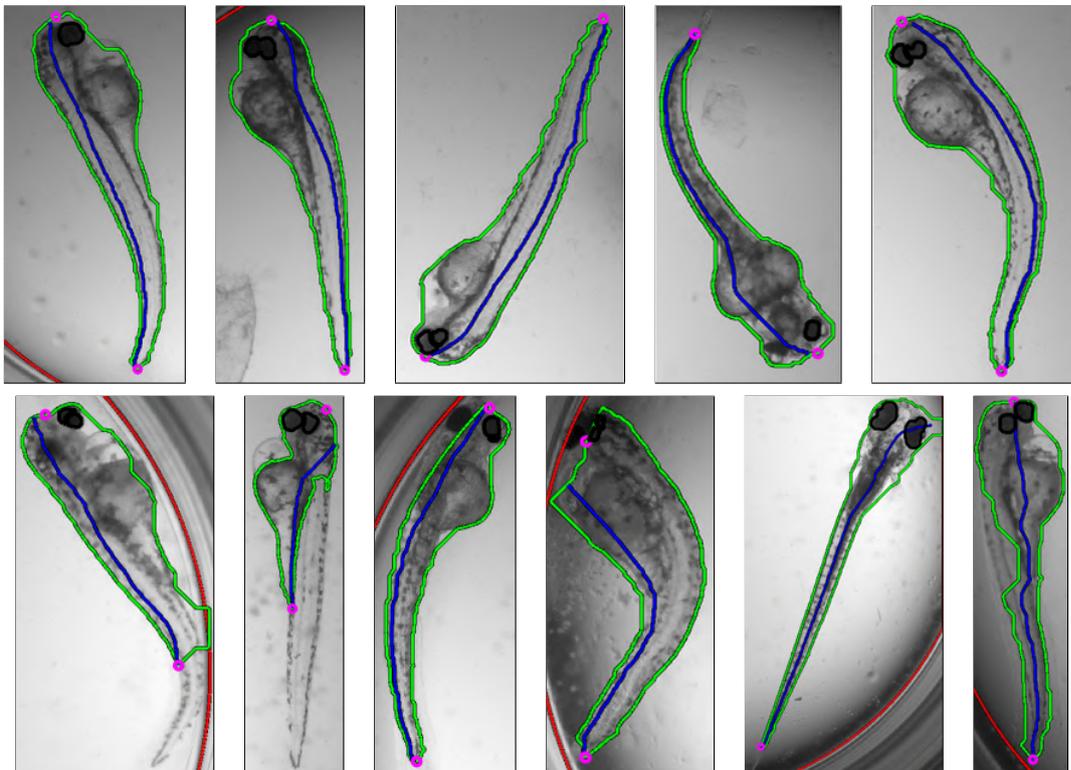


Figure 3.8: Illustration of detected zebrafish embryos with their eye/eyes and end points.

is very bright thus becoming too similar to its surroundings, and when the embryo is located near the well thus becomes difficult to localize because of the dark Meniscus effect. Some incorrect detections are shown in the bottom row of the figure.

3.4 Concluding remarks

In this chapter, a reliable quantitative morphological analysis of dose-dependent organ malformations using an *in vivo* vertebrate system has been presented. The zebrafish embryo model proved to be appropriate for complex evaluation of the irradiation-caused damages, and the most relevant morphological parameters could be defined for later radiobiological experiments.

It is doubtlessly necessary to develop automatic evaluation algorithms which can make the measurements fast and accurate, consequently saving time for the researchers.

Chapter 4

Localization of Visual Codes

Visual codes play an important role in automatic identification, which became an inseparable part of industrial processes. Thanks to the revolution of smartphones and telecommunication, it also becomes more and more popular in everyday life, containing embedded web addresses or other small informative texts.

While barcode reading is straightforward in images having optimal parameters (focus, illumination, code orientation, and position), localization of code regions is still challenging in many scenarios. Every setup has its own characteristics, therefore many approaches are justifiable. Industrial applications are likely to have more fixed parameters like illumination, camera type and code size, and processing speed and accuracy are the most important requirements. In everyday use, like with smartphone cameras, a wide variety of code types, sizes, noise levels and blurring can be observed, but the processing speed is often not crucial, and the image acquisition process can be repeated in order for successful detection.

In our work, our main goal was to detect different types of codes using classical approaches. We introduce visual codes and their types and also demonstrate related works in Section 4.1. Details of the used dataset and metrics are discussed in Section 4.2. In Section 4.3 we present methods based on morphological operations to localize traditional barcodes [81, 82]. Using the specificity of the barcode structure, we suggest pattern matching algorithms in Section 4.4 [88]. A possible industrial application is described in Section 4.5 [85].

4.1 Introduction to barcodes

Item identification using visual codes is popular in our everyday life, and there are several methods available for the process to be fast and reliable. The retrieval of the embedded data takes place in two steps. First, we have to find the visual code object within the acquired sensor data or image (localization step), then we have to use the symbology of the code and recognize the embedded data (decoding step). Decoding

is widely studied, so we can use many approaches from the literature or public APIs like the ZBar library¹.

It should be emphasized that decoding is far more straightforward, while the issue of localization is similar to object recognition and is still not fully solved. For localization of the code object, most algorithms use segmentation techniques using different features. Several industrial applications simply ignore the localization step by adding a fast rotating laser that scans in many directions. Also, false positives are not acceptable, but the checksum digit (barcode) and the error correction (QR code) make false positives very unlikely in practice.

The use of visual codes has a reputation of more than 50 years, however, in the past, the localization process required many conditions to fulfill. The first barcodes were in a fixed position on railway trucks and were read by a fixed sensor gate. As technology progressed, PoS terminals appeared, still requiring human intervention to perform code reading. In the '70s, new algorithms have been developed that could localize codes having various orientation and position within the image. The first approaches were very simplistic, they imitated the laser scanners of the barcode reading device. From the '90s, machine learning provided some more sophisticated solutions for the issue. Methods providing automatic code localization are usually slower, but more accurate than their predecessors. Accuracy and processing speed are conditions that can hardly be fulfilled simultaneously, and most approaches aim to find a balance between these. Some machine learning algorithms make an exception, and they are capable of a quick evaluation after a significantly slower learning process, provided that the features can be computed efficiently and there is sufficient amount of training data available.

In industrial applications, accuracy is more crucial, since missed codes may lead to loss of profit and in those cases, speed is a second desired attribute. Contrary, in smartphone applications accuracy is not as critical, because the user handles the device interactively, and repetition of the image acquisition is possible and relatively easy. In addition, these classical visual codes can also be produced in a way that they become unique and thus can be used to validate originality or authenticity.

Barcode formats

Visual codes are not meant to be readable for humans, they are decoded by specific devices and GS1 manages several types of barcode². The most popular 1D barcode subtypes are the EAN-13 and UPC standards. These are widely used in commerce, e.g., on wrapping of products, and they help quickly obtaining the information on, e.g., the producing country, types of entities of products. The flow of information is

¹Publicly available at <http://zbar.sourceforge.net/>

²<https://www.gs1.org/standards/barcodes>

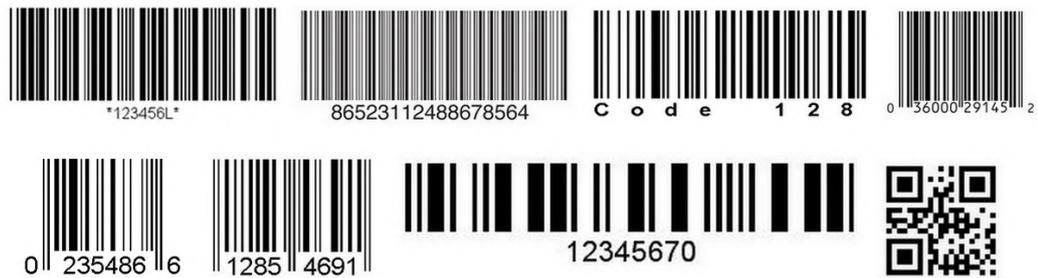


Figure 4.1: Popular barcode types (from left to right). Top row (1D codes): Code39, Codabar, Code128, UPC-A; Bottom row (1D codes, 2D code): UPC-E, EAN-8, I2of5, QR code.

greatly boosted using visual codes, which provide decoding of the embedded data by electronic devices. Some types include features that also help their localization. The traditional 1D barcode structure is simple: a sequence of parallel bright and dark bars of varying thickness represent information.

The literature sometimes refers to 2D codes as “barcodes”, however, they do not necessarily consist of “bars”. They carry the embedded data along two axes, and their most popular types are QR code and Data Matrix. Some 1D and 2D codes are presented in Figure 4.1.

Related works

There are numerous methods for the localization of visual codes in digital images. Some approaches imitate the classical laser scanner. Adelman et al. [3] introduced such a barcode recognition and information system to detect and read EAN-13 barcodes. This system works as a mobile application and localizes localization traditional and widely used barcodes. After preprocessing steps, Ohbuchi et al. [61] used a scanline based procedure to detect QR and EAN codes.

Mathematical morphology has been used in many approaches in the literature. Bodnár et al. [11] proposed using simple detectors such as combination of different morphological operators and distance map to detect barcodes efficiently. Texture analysis [10] can also achieve great efficiency. Lin et al. [52] demonstrated a fast and effective method that can simultaneously detect 1D and 2D barcodes. Their method is based on a modified run length smearing algorithm. Kong [44] defined regions of interests that may contain QR codes in synthetic images with the mix of Harris corner detector and convex hull. In addition, they recommended a solution to correct for geometric distortion. Belussi et al. [8] introduced a machine learning method that is based on the locator pattern of the QR code. They proposed a cascade of weak classifiers using features from the Haar wavelet family. Although it is fast, it provides

a noticeable amount of false positive code candidates. Bodnár et al. [13] proposed an improvement on that, using visual descriptors such as local binary pattern (LBP) and histogram of oriented gradients (HoG) features as an extension of the training step on the full code object. Sörös et al. [71] aimed to localize 1D and 2D codes using edge and corner maps, even considering the saturation channel in HSV images. Their algorithm is optimized for images suffering from heavy directional smoothing [72]. The method has high accuracy, in cases, however, where the code object is surrounded by text, their approach provides oversized bounding boxes. Text filtering can help get rid of this problem, considering the availability of surrounding text as a priori information. Szentandrás et al. [75] also worked with edge and corner maps and HoG features. Their method works locally on square image cells, similarly to convolution. This approach enables parallel execution and it is also highly accurate.

Yun et al. [107] introduced an orientation histogram-based method. They used a histogram to separate the principal orientation components from the entire image and calculate the local entropy of the orientation to generate a saliency map. Bodnár et al. [12] presented a method based on distance transformation. The algorithm also considers local image blocks and evaluates the distance map of the edge map. It takes into account the mean and standard deviation of the distance values within each block, then makes a binary decision whether or not the block contains a barcode part. While this feature can be computed efficiently, it has weak classification power, therefore it is not sufficient for use alone for the localization step. In their work, the authors tried to overcome this attribute using morphological operations.

Many recent papers use machine learning methods to solve various image processing problems. Hansen et al. [34] used a deep learning object detection algorithm, namely You Look Only Once (YOLO) model. Their network is based on a pre-trained Darknet19 model with 6000 epochs. The most common architecture for semantic segmentation is the U-Net that has different variants for each task. Ventsov et al. [97] divided the input image into 128×128 blocks, extracted statistical characteristics for each block and trained a convolution neural network. A Region-based Convolutional Neural Network (R-CNN) model was proposed by Ban et al. [93] for detecting diversified barcodes under complex scenes. For experiments, they used two pre-trained models, ImageNet and VGG16.

Our proposed approaches can effectively detect barcodes using classical image processing operations. Similarly, some of the introduced methods in the literature use classical methods, they also take advantage of the specificity of the barcode. In some cases, the basis idea is similar to ours but major steps of these algorithms are different. The efficiency of the described neural network methods may be higher than the classical ones, but these requires a large amount of training data, which is a disadvantage.

4.2 Image data and metric

For evaluation and comparison, we used datasets of real and synthetic images that contain one or more barcodes. Several research groups [17, 28, 34, 71, 76, 97, 107, 108] evaluated their algorithms on the WWU Muenster dataset, therefore we also decided to use that set for evaluation.

We have created an artificial test set from the barcode examples presented in Figure 4.2. Various distortions and levels of noise were applied. The generated images were rotated from 0° to 180° by 15° steps. Gaussian smoothing was applied with 3×3 kernel and 11 different σ values in range $[0, 5]$ with step of 0.5. Also, Gaussian noise was added with variance of 0.01 and $[0, 0.5]$ mean with step of 0.1. In total, we created 12 orientations from 8 types of barcodes, using 12 different smoothing and 6 different noise levels, with perspective distortions, counting as cca. 15 000 images. We also collected 100 real-life example images. Figure 4.2 illustrates some examples from our artificial and real-life datasets.

For the test set of QR codes, we used the database of Sörös et al. [71] consisting of 1400 real images. Besides, we used our synthetic database containing 10 000 test images. The latter set was generated similarly to the 1D barcode set. Figure 4.3 shows some samples from that data set.

An other public database by Dubská et al. contained two similar sets of QR code images, surrounded with text in a scene having low saturation in general. The first set contains 410 high-resolution (2560×1440 px) images with uneven lighting conditions, high grades of geometry or intensity distortion and blur. The second test set has 400 low-resolution (604×402 px) images with smaller grades of distortion and more even illumination, but having less light in general, thus producing darker images.

To measure the efficiency of the proposed and reimplemented literature algo-



Figure 4.2: 1D barcode samples with different distortions, synthetic (top row) and real images (bottom row).

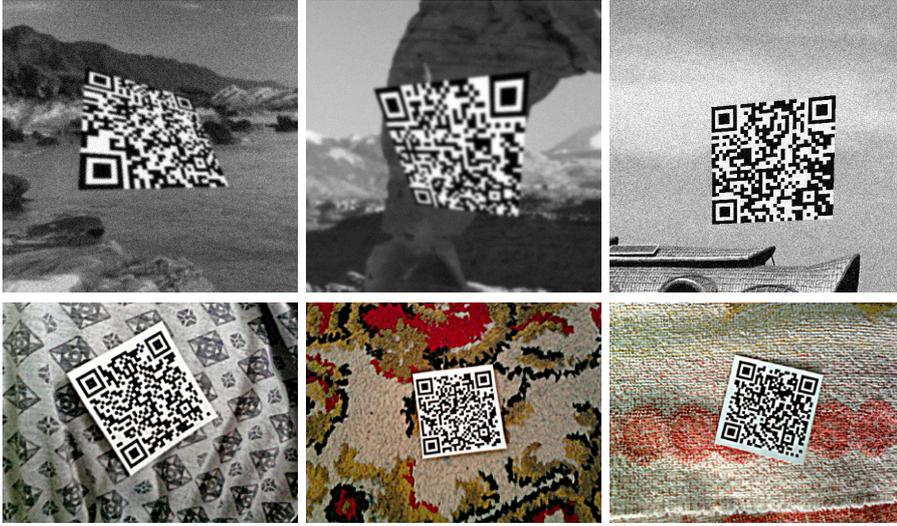


Figure 4.3: Synthetic (top) and real images (bottom) with QR code.

rithms, the overlap of the segmented codes is more important than the number of founded barcodes, so we evaluated this at pixel level for each object using the Jaccard similarity measure, defined as

$$J = \frac{TP}{TP + FP + FN}, \quad (4.1)$$

where TP (true positive) denotes the correctly detected pixels, FP (false positive) is the number of not valid code pixels and FN (false negative) is the number of not localized code pixels. Note that, ground truth regions have a tight fitting bounding polygon highlighting the code object without numbers and “quiet zones” as a border. A successful detection is where $J > 0.5$, according to the work of Szentandrásí et al. [75].

4.3 Mathematical morphology based algorithms

As described in Section 4.1, the traditional barcodes have a specific structure that can be used to localize them with simple and effective approaches. Mathematical morphology is useful for the analysis and processing of geometrical structures in images. The basic morphological operators are erosion and dilation.

We compare the effectiveness of the proposed method with the following approaches from the literature. The algorithm of Tuinstra’s [94] relies on that the intensity difference between the stripes is high within barcode region. Sobel kernel was used to estimate the gradient in the x and y directions and they selected pixels with high gradient and performed various morphological operations on the binary

image and in the last phase solidity test was used to remove false barcode segments.

The procedure of Tekin et al. [77] was designed for visually impaired or blind people to facilitate their everyday lives. The approach can be splitted into two main phases: preprocessing and detection. They used Gaussian smoothing to reduce noise level and also calculated gradient magnitude image. The image was scanned in four directions (horizontally, vertically, and in direction of the diagonals ($\pm 45^\circ$)) in the detection phase in the last step. They calculated entropy that characterize the texture of the image within a given neighborhood around each pixel.

Juett et al. [41] introduced a method based on bottom-hat filtering. The bottom-hat transform is defined as the difference between the original image and its closing. Closing operation is a dilation followed by an erosion, using the same structuring element for both operations [7]. The non-ideal image was improved by simple contrast stretching. They used morphological operations to find the possible barcode regions in the image and eliminated false candidate objects using area based thresholding. Our algorithm also relies on bottom-hat filtering and use morphological operations, but the processing steps build up a dissimilar system for detecting.

4.3.1 Algorithm based on bottom-hat transform

Similarly to the previously described methods [41, 77, 94], our algorithm also consists of two main phases. The input image is converted to grayscale in the preprocessing phase, because although barcodes may be printed in various colors, the pattern of dark bars on a light background is equivalent.

To reduce the image noise, we used Gaussian smoothing. Most of the barcode localization methods calculate gradient image to highlight edges. Instead of this, we used bottom-hat filtering that is also based on intensity differences. Although bottom-hat filtering is slower than other non-direction edge enhancement operations, its accuracy is higher. In order to detect possible barcode regions, the edge-enhanced image is binarized with global thresholding using 95% of the maximal intensity as threshold value.

The binarized image is suitable for finding possible barcode segments, but so far may contain false regions. We take advantage of the structure of the barcode, the fact that it consists of parallel lines at about the same distance. The entire image is analyzed using a horizontal scan whether the white pixels are located at appropriate distances at the given direction. The maximal distance parameter depends on the image resolution. This value is 5 pixels in the case of 800×600 images.

There might be small connected components which satisfy the criteria but are not barcode regions. These false regions are eliminated using an area threshold. We set the threshold to half the size of the largest component. A too high threshold value could remove small barcodes from high-resolution images, therefore in this step we

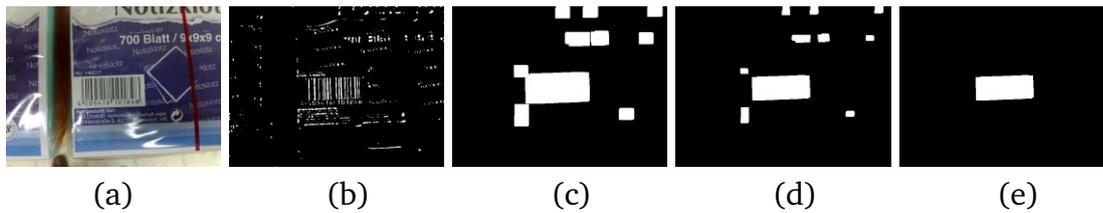


Figure 4.4: Intermediate steps of the algorithm based on bottom-hat transform at key phases: (a) original image, (b) binarization, (c) scanning, (d) morphological operations, (e) final output.

only remove the smaller false regions.

Our observations showed that for the final detection steps mostly dense text regions might remain along with the supposedly barcode areas. Since a barcode consists of a sequence of parallel bars and the bars are located at varying distance from each other, they do not compose a connected component. Therefore in this phase we use dilation to merge these patterns. The size of the square shaped structure element was defined as

$$S = \max(40, \text{width of the widest bar} \cdot 3). \quad (4.2)$$

For calculation of the size of the structure element, we used a constant that depends on the size of the image. In this case, this value was chosen empirically for the used dataset. Nevertheless, this dilation may also thicken and merge unwanted non-barcode regions as well. To cope with this problem, we also use the dual operation, i.e., erosion. Here, the structure element is also square shaped, and size is chosen less than (cca. 1/3 of) the one used for dilation, that is matching the width of the widest bar. Of course, there still may be false-positive objects which are removed in the last step on the basis of their size. Similarly to the previous step, objects smaller than half the size of the largest object are dropped. The pseudocode of the method is shown in Algorithm 4 and Figure 4.4 illustrates the key steps.

4.3.2 Extending the bottom-hat approach

In order to improve the method introduced in Section 4.3.1, we use another aspect to localize barcode regions after binarization since we observed worse results under poor lighting.

The preprocessing steps are retained thus we use bottom-hat filtering to highlight the bars of barcode(s) and a simple global thresholding operation to binarize the grayscale image.

Also we take the advantage of the structure of the barcode, the fact that it consists of approximately regularly spaced parallel stripes. Here, for each pixel, we calculate

Algorithm 4 Pseudocode of simple operations based 1D code localization algorithm.

```

1: function BARCODELOCALIZATION( $I_{gray}$ )                                ▷ input grayscale image  $I_{gray}$ 
2:    $C \leftarrow (I_{gray} \bullet Se_1) - I_{gray}$                                 ▷  $\bullet$ : morphological closing,  $Se_1$ : structure element
3:    $B \leftarrow []$                                                     ▷ empty image
4:    $B_{result} \leftarrow []$                                            ▷ empty image for output
5:   for  $i, j$  in all pixels do
6:     if  $C_{i,j} > th$  then                                           ▷  $th$ : threshold value
7:       if  $C_{i,j} < \max_d$  then                                       ▷  $\max_d$ : maximum distance
8:          $B_{i,j} \leftarrow 1$ 
9:       else
10:         $B_{i,j} \leftarrow 0$                                          ▷ no edge seen within specified maximum distance
11:      end if
12:    else
13:       $B_{i,j} \leftarrow 0$ 
14:    end if
15:  end for
16:   $B_{dil} \leftarrow B \oplus Se_2$                                        ▷  $\oplus$ : morphological dilation,  $Se_2$ : structure element
17:   $B_{er} \leftarrow B_{dil} \ominus Se_3$                                    ▷  $\ominus$ : morphological erosion,  $Se_3$ : structure element
18:  for all components do
19:    if actual  $B_{er}$  object area  $> \min_a$  then                       ▷  $\min_a$ : minimum area
20:      Record this object as a barcode segment to  $B_{result}$ 
21:    else
22:      Discard this object
23:    end if
24:  end for
25: end function

```

the Euclidean distance of the pixel from the nearest nonzero pixel. Using the distance map, objects that are far from other objects can be easily dropped, and only nearby objects (sort of a cluster of bar segments) will be kept. We calculate the average distance value for each row and use the minimum as the threshold value of the distance map. Regardless of the orientation of the barcode, where code segments are present, the distance values may be smaller than in other cases thus possible code segments can be found using threshold by average value.

Since a barcode consists of a sequence of parallel bars that are located at varying distance from each other, they do not compose a connected component. Therefore, we used dilation with a square structuring element whose size is defined as in Equation 4.2 to merge fragmented objects.

To remove unwanted objects, we used the same steps as in the original procedure, but with different structure elements and thresholds. We used a linear structure consistent with the size of the 1D barcodes. The size was less than the one used for dilation, i.e., it matches the width of the widest bar. Relying on the size of the objects, we remove false positive objects. If the image size was higher than 800×800 , objects

smaller than half the size of the largest object was removed, otherwise objects smaller than $\frac{1}{4}$ of the size of the largest object was rejected. We summarize the algorithm in Algorithm 5 and intermediate key stages in Figure 4.5.

Algorithm 5 Improved barcode detection method based on simple operations.

```

1: function IMPBARCODELOCALIZATION( $I_{gray}$ )           ▷ input grayscale image  $I_{gray}$ 
2:    $C \leftarrow (I_{gray} \bullet Se_1) - I_{gray}$            ▷  $\bullet$ : morphological closing,  $Se_1$ : structure element
3:    $d_{avgs} \leftarrow []$ 
4:    $B \leftarrow []$ 
5:    $B_{result} \leftarrow []$ 
6:   for  $i, j$  in all pixels do
7:     if  $C_{i,j} > th$  then                           ▷ th: threshold value
8:        $C_{i,j} \leftarrow 1$ 
9:     else
10:       $C_{i,j} \leftarrow 0$ 
11:    end if
12:  end for
13:   $D_{map} \leftarrow$  euclidean distance map of  $C_{i,j}$ 
14:   $d_{min} \leftarrow$  minimum of row averages on  $D_{map}$ 
15:  for  $i, j$  in all pixels do
16:    if  $D_{map_{i,j}} > d_{min}$  then
17:       $B_{i,j} \leftarrow 0$ 
18:    end if
19:  end for
20:   $B_{dil} \leftarrow B \oplus Se_2$                        ▷  $\oplus$ : morphological dilation,  $Se_2$ : structure element
21:   $B_{er} \leftarrow B_{dil} \ominus Se_3$                  ▷  $\ominus$ : morphological erosion,  $Se_3$ : structure element
22:  for all components do
23:    if actual  $B_{er}$  object area  $> min_a$  then       ▷  $min_a$ : minimum area
24:      Record this object as a barcode segment to  $B_{result}$ 
25:    else
26:      Discard this object
27:    end if
28:  end for
29: end function

```

4.3.3 Evaluation and results

In Section 4.3 we introduced approaches from the literature. Here, we compare our methods described in Section 4.3.1 (Proposed431) and Section 4.3.2 (Proposed432) with these methods.

The performance of the algorithms do not seem to depend on the applied stretching in the test images and also seem to be insensitive to rotation as well. We used our synthetic 1D dataset and 100 real-life examples for evaluating algorithms.

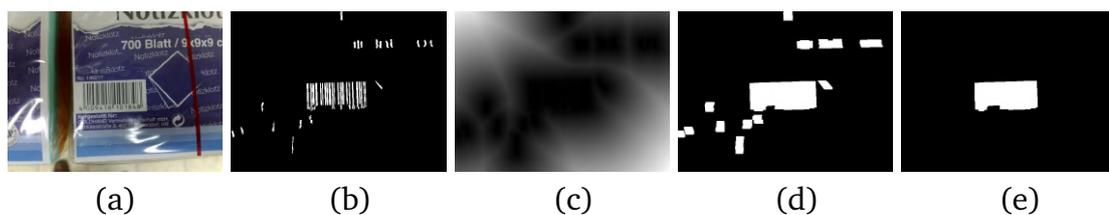


Figure 4.5: Intermediate steps of the improved barcode localization method based on simple operations at key phases: (a) original image, (b) binarization, (c) Euclidean distance, (d) morphological operations, (e) final output.

We can conclude that all five algorithms are capable of effectively localizing barcodes. The method of Tekin [77] is fast, but it loses efficiency when there are several barcodes in the image and noise is also present. The algorithm of Juett et al. [41] is slower than the others, because in the detection phase the image is scanned in 16 different orientations. However, this brings better accuracy to the process at the expense of more computation.

Our proposed algorithms are also very fast. In most cases, faster procedures locate barcodes less accurately or not at all, as if loss of accuracy were the price for the speed gain. However, in our case, speed is not coupled with a significant cost in accuracy. From the evaluation output we can conclude that the proposed procedures have second best running efficiency for all test images and in many cases it also shows best accuracy. The weakness of the method appears when the image is very noisy or when there are such image areas which are similar to a barcode.

Table 4.1 presents the execution time of the detection methods for images containing 1 or 3 barcodes. Here, one can easily discover the significant differences between the different approaches.

Subsequent tables show how the algorithms operate on images which contain either one or three code pieces of various 1D and 2D barcodes types. The structure of barcode types varies, which also has an effect on how well the algorithms can perform. Accuracy of the detection methods for images containing a single or 3 pieces of codes for various code types is presented in Table 4.2. Tuinstra et al., Tekin, and Juett et al. reached quite low performance for Code128, however our

Table 4.1: Execution time of the code localization algorithms. Mean values (\pm standard deviation) (expressed in seconds) for all test images of a certain class.

		Tuinstra et al.	Tekin	Juett et al.	Proposed431	Proposed432
1 code	1D	0.23 \pm 0.13	0.67 \pm 0.22	0.81 \pm 0.43	0.13 \pm 0.10	0.16 \pm 0.08
	2D	0.21 \pm 0.14	0.77 \pm 0.47	0.73 \pm 0.44	0.12 \pm 0.12	0.14 \pm 0.09
3 codes	1D	0.54 \pm 0.35	1.73 \pm 0.66	1.86 \pm 0.78	0.18 \pm 0.09	0.35 \pm 0.12
	2D	0.46 \pm 0.21	1.70 \pm 0.54	1.68 \pm 0.71	0.18 \pm 0.10	0.28 \pm 0.07

Table 4.2: Accuracy of the algorithms for various types of codes, for images containing 1 or 3 codes. Mean values (expressed in percent) for all distorted test images with 1D barcodes (above) and for all test images with 2D codes (below). The highest efficiency is highlighted in bold.

		Codabar	Code128	Code39	EAN-13	EAN-8	I2of5	UPC-A	UPC-E	all
1 code	Tuinstra et al.	99.0	50.4	98.7	98.8	98.7	98.6	99.1	82.7	91.8
	Tekin	98.9	63.6	94.8	94.7	99.6	84.6	98.2	92.6	86.4
	Juett et al.	99.2	70.9	86.3	91.7	84.6	84.4	94.9	76.0	78.9
	Proposed431	99.9	95.7	99.4	98.9	99.6	93.1	86.4	99.6	96.6
	Proposed432	98.6	93.7	100.0	99.3	98.2	92.5	100.0	100.0	96.8
3 codes	Tuinstra et al.	83.8	83.6	83.3	83.3	83.3	78.8	74.6	98.7	74.0
	Tekin	56.2	95.0	90.5	91.4	97.2	35.5	65.5	77.5	74.7
	Juett et al.	56.2	79.6	67.3	76.6	81.8	93.8	92.6	96.7	77.4
	Proposed431	92.8	80.1	94.7	99.3	99.4	99.2	99.9	99.3	95.6
	Proposed432	87.5	84.8	90.5	99.2	95.9	84.9	95.4	92.5	89.2

		Codablock	PDF417	all
1 code	Tuinstra et al.	87.0	81.7	85.0
	Tekin	99.8	99.6	99.7
	Juett et al.	83.3	79.7	82.5
	Proposed431	87.7	93.9	90.8
	Proposed432	100.0	97.6	99.1
3 codes	Tuinstra et al.	86.6	83.3	85.0
	Tekin	100.0	100.0	100.0
	Juett et al.	80.0	83.9	82.4
	Proposed431	90.7	91.7	91.2
	Proposed432	94.2	90.8	92.5

Proposed431 and Proposed432 algorithm handle this variable-length code as well as the fixed-length types. For the 2D codes, Tekin shows exceptionally good accuracy, but the other methods also perform well on these stacked barcodes. There is more variance in performance on images with 3 codes, and some methods (Tuinstra et al., Proposed431, Proposed432) show a solid performance while the others very much depend on the type of code present in the image.

The accuracy of the algorithms degrades on distorted images. Next, we analyze the effect of distortions on efficiency. Table 4.3 shows the methods' efficiency with respect to the level of blur applied to the test images. Image smoothing does not change the performance considerably. Although there is a noticeable inverse relation between values of σ and accuracy, this is not significant. It is interesting to note that in some cases the barcode detection accuracy is higher in images with three barcodes than for images with a single code present. This can be due to the fact that relative size (w.r.t. image size) of individual codes are smaller in the 3-code examples. Tekin, Proposed431 and Proposed432 methods show outstanding accuracy in images which contain 2D barcodes and there are no considerable differences between the precision values.

Table 4.3: Accuracy of the algorithms for different blur levels, for images containing a single code and for those containing 3 codes. Mean values (expressed in percent) for all distorted test images with 1D barcodes and for all test images with 2D codes. The highest efficiency is highlighted in bold.

		1D				2D			
		0	1.5	2.5	all 1D	0	1.5	2.5	all 2D
1 code	Tuinstra et al.	100.0	90.4	90.4	90.5	100.0	87.7	83.6	87.5
	Tekin	100.0	88.5	90.4	89.1	100.0	100.0	100.0	100.0
	Juett et al.	100.0	89.4	87.5	89.6	100.0	82.6	82.6	83.3
	Proposed431	100.0	96.8	97.1	96.7	100.0	95.7	95.7	95.8
	Proposed432	100.0	95.4	93.3	94.3	100.0	100.0	100.0	100.0
3 codes	Tuinstra et al.	100.0	91.7	82.7	82.7	100.0	86.7	82.0	82.7
	Tekin	100.0	86.7	82.0	77.9	100.0	100.0	100.0	100.0
	Juett et al.	100.0	89.1	90.2	90.6	100.0	82.5	74.5	82.9
	Proposed431	100.0	97.0	97.0	97.5	100.0	100.0	100.0	100.0
	Proposed432	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 4.4: Accuracy of the algorithms for different noise levels, for images containing a single code and for those containing 3 codes. Mean values (expressed in percent) for all distorted test images with 1D barcodes (above), and for all test images with 2D codes (below). The highest efficiency is highlighted in bold.

		1D				2D			
		0	30	50	all 1D	0	30	50	all 2D
1 code	Tuinstra et al.	100.0	89.4	93.3	90.5	100.0	87.7	77.8	85.8
	Tekin	100.0	86.2	84.2	88.6	100.0	100.0	98.6	99.5
	Juett et al.	100.0	79.2	69.9	83.5	100.0	79.3	69.6	77.3
	Proposed431	100.0	97.4	94.9	96.8	100.0	94.2	92.3	94.8
	Proposed432	100.0	92.5	91.6	95.4	100.0	100.0	100.0	100.0
3 codes	Tuinstra et al.	100.0	84.6	73.8	87.6	100.0	90.4	84.9	86.4
	Tekin	100.0	86.3	82.7	87.3	100.0	100.0	100.0	100.0
	Juett et al.	100.0	82.7	64.4	84.7	100.0	73.0	71.4	73.3
	Proposed431	100.0	96.7	90.6	95.6	100.0	100.0	85.7	95.6
	Proposed432	100.0	95.8	91.6	95.1	100.0	100.0	100.0	100.0

In a similar manner, we analyzed how the algorithms perform on images having different levels of noise added. We can see from Table 4.4 that each algorithm is somewhat sensitive to noise. The introduced literature methods lose much of their accuracy as the noise level increases, however, our methods degrade to a lesser extent only.

We also tested the methods on a set of 100 real-life images. These images contain 1D barcodes each. The accuracy of the algorithms are 92 % (Tuinstra et al.), 91 % (Tekin), 92 % (Juett et al.), 94 % (Proposed431) and 96 % (Proposed432).

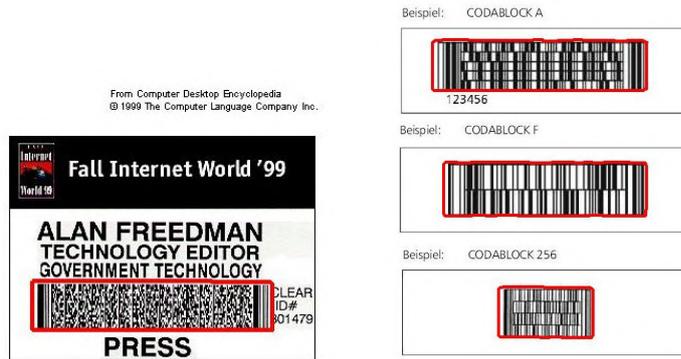


Figure 4.6: Examples of detection output by the proposed algorithm for PDF417 (left) and Codablock (right) code samples.

In Figure 4.6 we show two result images of our algorithm for 2D barcodes. Stacked barcodes contain data all parts of the barcode so we have to detect the total code and there is a much smaller visual redundancy contrary to 1D barcodes.

4.4 Detection using template matching

Pattern matching is an approach to find the location of a template image. Since classical 1D barcodes show similar patterns and in 2D codes we can also observe specific features, thus pattern matching may be a possible solution for detection.

4.4.1 1D traditional barcode segmentation

During processing, the input images can be of different sizes. We reduce the height of the image to a fixed size of 500 pixels in order to make them more uniform, easier to handle, and make processing faster. Empirical experience has shown that this is the smallest image size where smaller area code regions can still be localized in the used databases. We did not use color information during the process, so the input RGB image was converted to grayscale. Input images are often blurry, therefore we also use sharpening (Figure 4.7(b)).

The detection process is based on binary images, so the image is binarized using a global threshold. In our case, this value was 4% of the maximum intensity (Figure 4.7(c)). This threshold was chosen empirically, based on the observation that white parts of the code more often fall into the gray intensity range because of dust, cheap quality labels or shapes being present because the packaging does not necessarily use white color for the bright parts of the code. This is not a robust solution, but selecting valid regions from false regions is easier than finding a missing

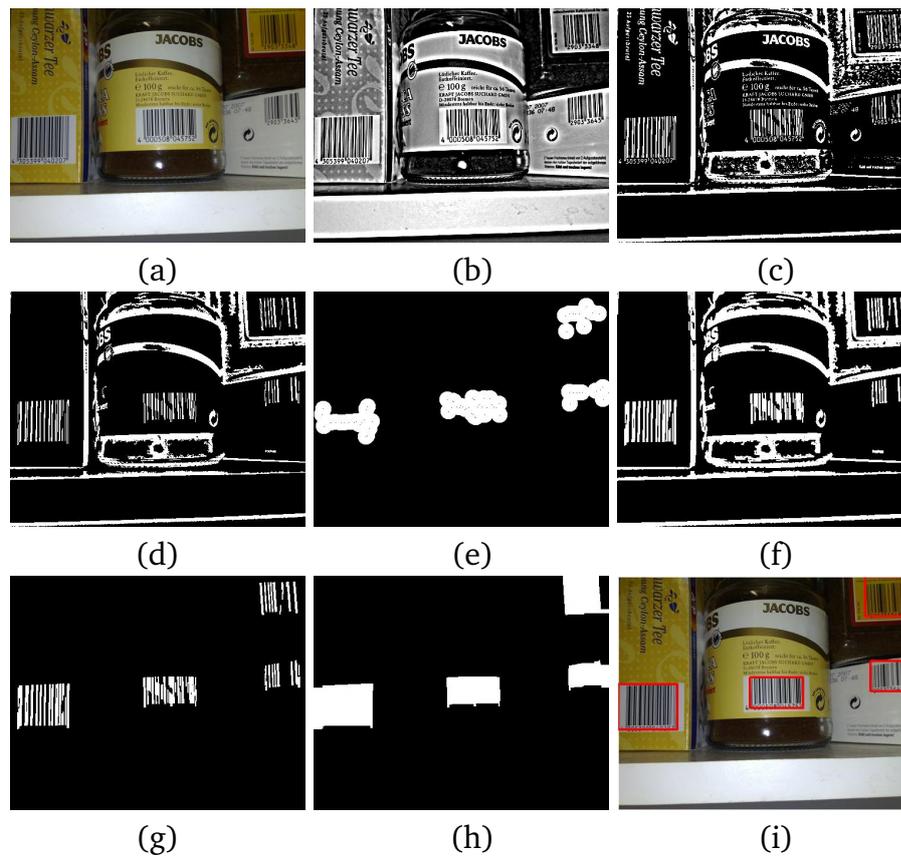


Figure 4.7: The proposed method for 1D barcode localization. (a) input image, (b) deblurring, (c) binarization, (d) filtering by rectangularity, (e) thresholding using the number of cluster points, (f) result of (d) after dilation, (g) matching objects with (e) on image (f), (h) opening, (i) code candidate boxes overlaid onto (a) in red

part during a post-processing process. Obviously, a filter step is needed to reduce the number of false regions where different noise, etc., can occur. The shape of the bars of a barcode are rectangular, so we examine the shape of each object formed by connected binary components. If the shape of the object is not approximately rectangular, we do not consider it as a candidate region. The input for pattern matching is illustrated in Figure 4.7(d). As barcodes consist of parallel “bars”, the template consists of two parallel lines. Traditional barcodes consist of a plurality of parallel lines, so a similar part of the image may be suitable for template matching. We also know the maximum and minimum distance between bars for each type of code. The template image was selected based on this information.

The process of template matching occurs in the frequency domain using Fast Fourier Transformation. We rotate the template image in every 10° , up to 170° , and compute the sum of the pointwise multiplication of the frequency representation of the original image and the rotated template images. The 10° step was empirically

found to be sufficient. Thanks to the symmetric nature of the matched template, it is sufficient to examine only the aforementioned rotations. The summation image is converted into image space using inverse Fourier transformation and thresholded with the mean intensity of the inverse image.

Next, we use only the center of the objects that are being obtained. Pixels belonging to a specific cluster are well-separable like the bars of the barcode that are close to each other. We used the well-known kNN clustering method to separate connected objects from each other. The set of points from template matching can be well separated, so we have chosen $k=3$. To keep valid regions, we use a priori information that a barcode consists of at least 8 bars, so only those clusters are kept that have at least 8 points (Figure 4.7(e)).

We apply morphological dilation on the binary version of the original image, using a 3×3 structuring element (Figure 4.7(f)). We investigate objects between the dilated image and filtered cluster points and we only keep overlapping regions (Figure 4.7(g)). In order to determine the whole barcode region, we use morphological opening with a square-shaped structuring element. The size is defined based on the maximal distance between the stripes, which provides that every barcode will have its own connected region (Figure 4.7(h)).

4.4.2 QR code localization

Similarly to the procedure described in Section 4.4.1, we also work with images with specific size during the localization and convert the images to grayscale. In order to highlight the barcode areas, we use standard deviation based adaptive filtering method with 3×3 neighborhood [33]. The resulting image is heterogeneous, so we calculate the number of object points for every kernel with a fixed 7×7 size and remove the actual segment from possible barcode regions where this value is under the half of the kernel size.

In order to remove false small regions and then merge the connected regions, we apply a morphological opening. The shape of the QR code is ideally a square, so we use a square shaped structuring element for the morphological operation. Based on empirical observations, we binarize the image obtained in the previous step with the threshold value of $\frac{7}{8}$ of the maximum intensity. Since global thresholding is not an overly robust operation, but the bars of the barcode have low-intensity usually, thus we can eliminate numerous false segments with a low-intensity value when determining the global threshold value. In the last step, we validate the code segments by pattern matching. For this, we used a small QR code as a sample. We do the template matching on the original image and we investigate overlapping with the opened binary image similarly as described in Section 4.4.1. Valid QR code regions are available after the validation step. The major steps of the algorithm and also the

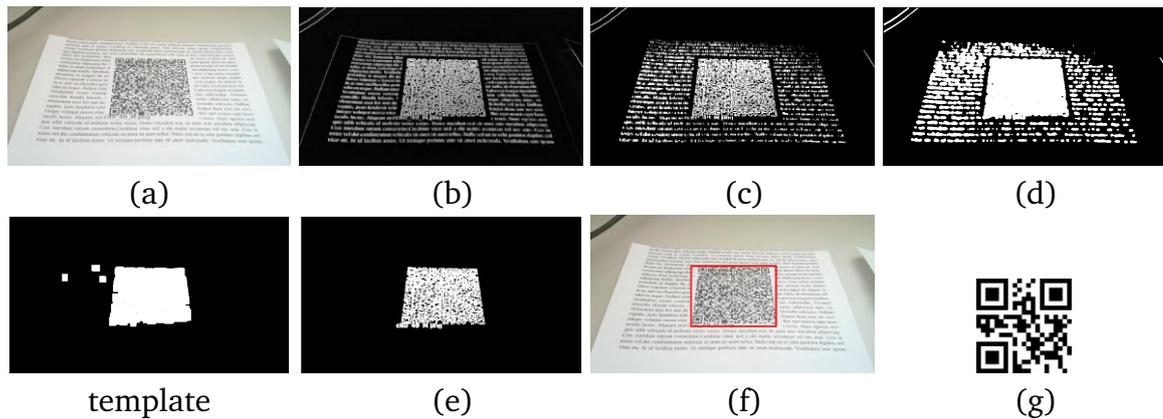


Figure 4.8: Proposed method for QR code localization. (a) input image, (b) σ filtering, (c) binarization, (d) density calculation, (e) morphological opening, (f) post-processing, (g) detected QR code

used template are illustrated in Figure 4.8.

4.4.3 Evaluation and results

Our approaches were compared against other algorithms in the literature using the Muenster database as a benchmark. For the comparison, we selected algorithms based on various features, like edge and corner maps [71], or deformable templates [28]. Zamberletti et al. [108] work with the popular localization method that is also implemented in ZXing barcode reading framework, while Creusot et al. [17] have an approach that uses MSER. We note that refers to our template matching based approach PROP-TM. Hansen et al. [34] reported 0.87 Jaccard value to result in their article, but we did not use this result in our comparison, because there were some missing information about training parameters, so we were unable to re-implement and evaluate the procedure.

Most of the presented methods highlighted the bars of the barcode (gradient calculation, bottom-hat filtering, etc.) for further processing in the first main step. The algorithms that are based on simple image operations and logic are examined on the synthetic image database under different conditions. We investigated and found that the procedures are not sensitive to rotation. We also examined the results with different noise and blur levels. In both cases the generated images had 6 different levels for those parameters, modifying the σ value of Gaussian smoothing and a γ weight for weighted addition of uniform noise to the original image. We present the obtained Jaccard values in Table 4.6. It shows that the efficiency of the procedures is hardly reduced by increasing the level of noise and smoothing. However, in all cases, the methods do not behave significantly differently on ideal cases. We used 8 differ-

Table 4.5: Comparison of various localization algorithms on the Muenster 1D barcode dataset. (Mean and standard deviation of Jaccard index.) Best performing method is typeset in bold.

Algorithm	\bar{J}	st.dev.
Zamberletti et al. [108]	0.6950	N/A
Creusot et al. [17]	0.7990	N/A
Gallo et al. [28]	0.7089	0.3542
Tekin et al. [76]	0.8122	0.2562
Proposed431 [81]	0.5200	0.2967
Sörös et al. [71]	0.6647	0.2277
Yun et al. [107]	0.4716	0.2240
PROP-TM	0.8430	0.1876

Table 4.6: Accuracy of the methods for different blur (σ value) and noise levels (in percent) on 1D synthetic images (mean Jaccard index). The highest efficiency is highlighted in bold.

Blur	Noise	Gallo	Yun	Sörös	PROP-TM
	0	0.74	0.64	0.75	0.82
0	30	0.75	0.66	0.76	0.82
	50	0.75	0.65	0.77	0.81
3	0	0.77	0.64	0.74	0.82
	30	0.75	0.67	0.76	0.81
	50	0.73	0.63	0.76	0.78
5	0	0.76	0.65	0.74	0.82
	30	0.74	0.66	0.76	0.81
	50	0.73	0.63	0.75	0.78

ent codes to generate the images. The procedures were less successful in localization task for Code-39, I2of5, UPC-E types as shown in Table 4.7. Results are shown in Table 4.5. Only the algorithm of Tekin et al. [76] has better mean value, however, with higher variance. We also present some qualitative results of the aforementioned approaches on challenging images that are illuminated, geometry distorted and contains several objects similar to a barcode. See Figure 4.9 for details.

Efficiency of our template matching approach is compared with methods introduced by Ohbuchi [61] and Lin [52] in Section 4.4. Our algorithm (PROP-TMQR, Section 4.4.2) with 2D QR codes performed well on the Dubská datasets, with $\bar{J}_{set1} = 0.8315$ and $\bar{J}_{set2} = 0.8102$. Hansen et al. [34] published 0.73 average Jaccard value on Dubská sets (Table 4.8).

Table 4.7: Accuracy of the algorithms for various types of code on 1D synthetic images (mean Jaccard index). The highest efficiency is highlighted in bold.

	Gallo	Yun	Sörös	PROP-TM
Codabar	0.81	0.71	0.80	0.87
Code-128	0.82	0.73	0.83	0.92
Code-39	0.57	0.52	0.59	0.72
Ean-13	0.82	0.69	0.81	0.83
EAN-8	0.81	0.70	0.81	0.83
I2of5	0.69	0.60	0.69	0.72
UPC-A	0.78	0.66	0.78	0.80
UPC-E	0.69	0.64	0.74	0.73

Table 4.8: Comparison of various localization algorithms on the QR synthetic database (mean Jaccard index). The highest efficiency is highlighted in bold.

input	Ohbuchi [61]	Lin [52]	PROP-TMQR
Dubská set1	0.79	0.81	0.83
Dubská set2	0.77	0.79	0.81

We also compared the results on synthetic images along different blur and noise levels (Table 4.9) and present some results on challenging images in Figure 4.10.

4.5 Validating authenticity of visual codes

One of the most important and widely used technique for automatic identification is the use of visual codes. Identifiers encoded in various symbols and patterns make electronic reading possible, that greatly helps and speeds up processing, e.g., at cashier lines, warehouse transactions, high speed processing places, production lines. The common codes designed using geometric patterns usually identify types or entities. However, patterns can be produced that, by their nature, are unique and thus can be used to validate originality or authenticity.

Natural feature identifiers are closer to biometrical identifiers, such as fingerprints, iris or retina pictures, that are also unique, basically irreproducible, and inseparable from the objects that they uniquely identify. Labels with printed identifiers can be usually easily reproduced and, depending on the technology, removed and reattached to a different object. Techniques exist for producing non-detachable labels and when combined with natural features, these may become a highly secure



Figure 4.9: Qualitative results for the compared methods on some challenging 1D images. From top to bottom: Gallo et al., Yun et al., Sörös et al. and Proposed.

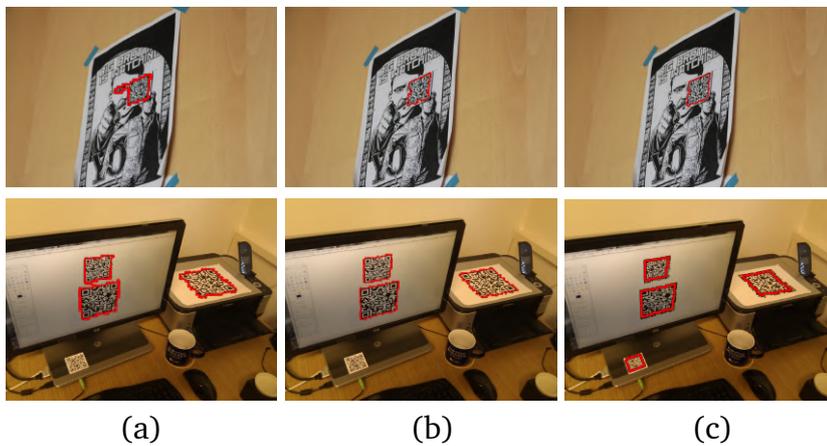


Figure 4.10: Qualitative results for the compared methods on some challenging QR images. (a) Ohbuchi, (b) Lin, (c) Proposed

means for identification.

In the topic of identification/recognition of objects, algorithms are needed to automatically locate and decode the identifiers attached to the objects. Very different

Table 4.9: Accuracy of the methods for different blur (σ value) and noise levels (in percent) on QR synthetic images (mean Jaccard index). The highest efficiency is highlighted in bold.

Blur	Noise	Ohbuchi	Lin	PROP-TM
0	0	0.96	0.50	0.99
	50	0.96	0.50	0.99
	100	0.96	0.50	0.99
1.5	0	0.57	0.51	0.77
	50	0.52	0.49	0.78
	100	0.49	0.40	0.75
3	0	0.47	0.42	0.83
	50	0.48	0.45	0.80
	100	0.45	0.37	0.76

techniques are required for natural biometric IDs, such as fingerprints [43], iris [15] or retina patterns [6]. Decoding the information content is more complex when the identifier pattern is the result of natural processes and thus can be considered almost random and continuous as opposed to traditional well-structured graphical codes.

4.5.1 Localization and validation of NFI label

Label identification in this case contains a combination of artificial and natural feature identifiers. A standard QR code of a given size is in the center of the label area. This code can encode any content relevant in the application context, e.g. a serial number or a key into a database. Figure 4.11 shows a schematic drawing of the label layout and an image of a real piece of a prototype label.

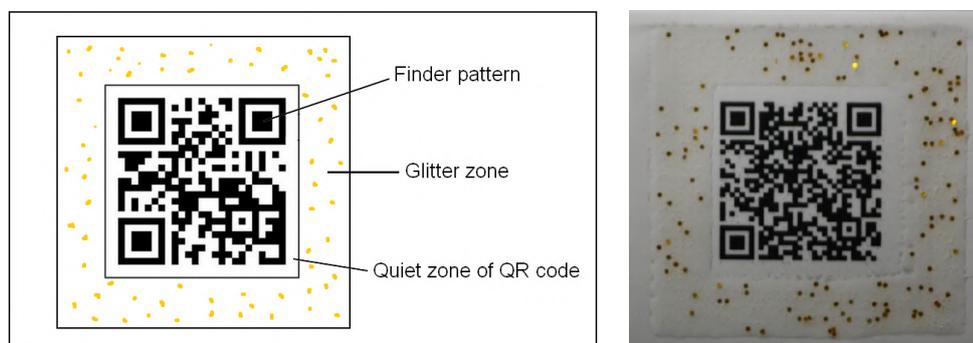


Figure 4.11: A sketch of the NFI label layout (left) and an image of a real prototype label (right).

The QR box is surrounded by a random set of particles or glitters in a window frame of given width and background color. The number, color, size, shape, and position of the glitters can all be considered random variables. Particles and the pattern can be made of various kinds of material using different technologies. In the examples shown, glitters are made of metallic flakes and are fixed in thin layer of transparent glue or under a transparent foil. Even glitters made by the same technology may reflect light differently which is mostly caused by the small differences in their surface angle w.r.t. the light source and camera position. In the right hand side image in Figure 4.11, a couple glitters (in the top-right quadrant) appear very bright, while most of the others are dark, but there are also a few of faint color.

In the presumed application context each label uniquely identifies a piece of product/parcel/object. After creating/manufacturing the label, it is supposed to be tracked through its lifetime and after it has been attached to an object, the uniquely identified object is tracked as well. Counterfeiting by duplicating the visual appearance of a known valid label would signal the tracking system if an identifier shows up at a place where it is not supposed to be.

In our envisioned setup, a reference photo is taken of each printed label under controlled positioning and lighting condition as part of the production process. (Instead, a few reference photos could also be taken and their information fused to provide better reference features for the matching.) Our task is to locate the label in an image taken by a camera of a mobile phone or tablet, find and decode the QR code in the label, extract the NFI features (glitters and their properties) from the image, and match them with the features extracted from the reference image(s) to reliably identify objects and detect fake labels. The algorithm is expected to work on images taken by average mobile phone cameras³ (say, e.g. 5 megapixels) under non-perfect conditions (e.g. uneven lighting, shadows, water drops, frost, or other surface artifacts) and feature extraction should also run on a standard mobile phone or tablet. It is also important to note here, that matching/verification is done purely based on the features extracted from the images, without any image transfer between the mobile client and the server infrastructure, neither the reference image nor the one being checked.

The following steps can be identified in the process:

1. image acquisition,
2. image processing,
3. identification.

Standard tools exist for the image acquisition step. Here, we focus on the image

³This work was done in 2013, so we developed the approach to process images taken by average cameras at that time.

processing part and with the last part we only deal to the level necessary to present the usability of the image processing part.

4.5.2 Image processing steps of recognition

Depending on the label specification, different image processing approaches are required. If the label specification allows for a good contrast between the glitter background and the surroundings e.g. a dark background with bright glitters and labels are applied to white/bright surfaces, the outer frame, the glitter zone can be easily identified based on color contrast, thresholding, and simple morphology. When the label specifications do not use dark glitter background (as in our prototype setup), the bright glitter zone is hard to differentiate from the white surroundings so we decided to indirectly determine its bounds. Thus, our image processing pipeline consists of the following major steps:

1. locate the QR code (position and orientation),
2. delineate the glitter zone,
3. find the glitters,
4. extract features from the glitter pattern.

The QR code in the middle of the label has a very specific pattern and many algorithms exist to detect QR codes in an image. In our previous work, we proposed an algorithm to detect QR code which is described in Section 4.4.2. Besides, other approaches from the literature are also introduced in Section 4.1. Unfortunately, most QR reader applications available on the mobile platform work only if the QR code occupies the major part of the field of view, it is well (centrally) positioned in the image, and it is properly oriented. Assistance is sometimes provided in the application to aid the user in finding the position, zoom, and focus and perhaps give feedback (or even more, actually take the picture without requiring the user to hit a button) when everything seems appropriate to take the picture for successful QR code decoding. As we are not only interested in the QR code, but the surrounding glitter zone is also (even more) important, the algorithm must be more robust to image acquisition differences and artifacts.

Since we need an algorithm that works on images coming from a wide range of mobile cameras, we decided to standardize the image size before further image processing takes place. We rescaled the input image to a given width such that the aspect ratio is not changed. The purpose of this size normalization is twofold: 1) to reduce the processing time to allow almost realtime applications, and 2) to design an algorithm that works with a single parameter setting on a wide range of images

coming from various mobile cameras. After the initial tests, we settled at 450 pixel wide images that turned out to be sufficiently detailed but at the same time not too large, and basically all mobile cameras have at least such a resolution.

The native RGB images (Figure 4.12(a)) are converted to the device-independent $L^*a^*b^*$ color space, and in further processing we use the L^* (luminosity) channel. In our case L^* channel usage turned out to be more effective than other simple weighted grayscale conversions of the RGB space such as the V channel from HSV.

Since input images are distorted by uneven (or even non-sufficient) lighting, we enhance the luminance image by contrast stretching (Figure 4.12(b)). This seems a valid step considering the assumption that the QR code elements show high contrast w.r.t. the background.

In the following step we exploit the known fact that a QR code is square shaped, and apply grayscale morphological opening with a square shaped structuring ele-

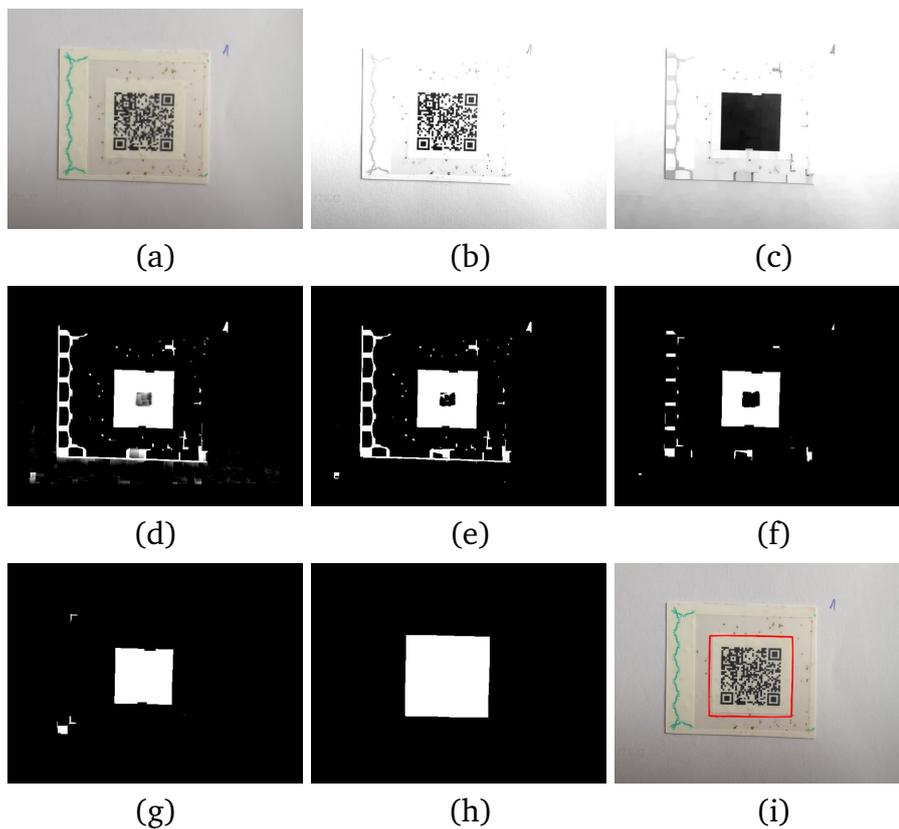


Figure 4.12: Intermediate stages of the QR code localization pipeline. (a) input image, (b) image after contrast stretching, (c) resulting image after morphological opening, (d) LoG filtered image, (e) binary image after LoG filtering threshold, (f) separate unconnected components with morphological erosion, (g) after eccentricity thresholding, (h) after major axis length thresholding, and (i) overlaid result

ment (Figure 4.12(c)). This is followed by LoG filtering [32] (Figure 4.12(d)) and binarization using global thresholding (Figure 4.12(e)). The resulting image so far may contain many false positive regions, many small connected components which satisfy the criteria so far but do not belong to barcode region. A morphological erosion with a smaller structuring element helps removing the majority of these artifacts (Figure 4.12(f)). Further filtering is done by using thresholds on eccentricity (Figure 4.12(g)) and major axis length (Figure 4.12(h)) and in the result only the QR code component remains.

There is a quiet zone around the QR pattern but the detected component does not include this, so we use morphological dilation to increase the object to extend the detected QR code region to include the quiet zone as well. This helps producing a square like component with smoother boundaries (without this the detected component would have ragged edges) which does not interfere with the NFI label since the quiet zone separates the QR pattern from the NFI zone.

Once the QR code position is known, orientation can be determined. Remember, we want a robust algorithm that does not depend on proper alignment of the camera when taking the picture. Nevertheless, knowing the orientation of the QR code is required for the QR decoder and also forms the basis for describing the glitter pattern in a comparable way.

The corner points of the extended QR code region are detected as intersection points of the lines fitted to the set of contour points of the segmented mask [95]. The projection transformation is determined based on these corner points assuming that the original pattern is a square and distortions are due to the varying positioning and angles of the camera taking the picture. Then the image can be rectified using the inverse of the found transformation.

In addition to rectification, orientation of the QR code also needs to be determined. Rectification and proper orientation of the image are required to be able to define a standard coordinate system for the glitters. Without this, images taken of the same label by different cameras at different times could not be compared to each other or to the reference. We should store all extracted features in the reference database with respect to a QR code with standard orientation (see Figure 4.11).

The square shaped locator marks in the corners of the QR code can be used to determine proper orientation. To detect the locators we perform the following operations on the segmented QR zone part of the image. The grayscale QR image (Figure 4.13(a)) is binarized by adaptive thresholding using 11×11 kernel and the calculated median as the threshold value (Figure 4.13(b)) and hole filling (Figure 4.13(d)) is applied to the complement (inverted) image (Figure 4.13(c)). A distance map is calculated (Figure 4.13(e)), grayscale morphological closing is performed on the map (Figure 4.13(f)) to reduce small and dark regions, and the result is thresholded (Figure 4.13(g)). In order to determine the direction of the QR code, it is necessary to

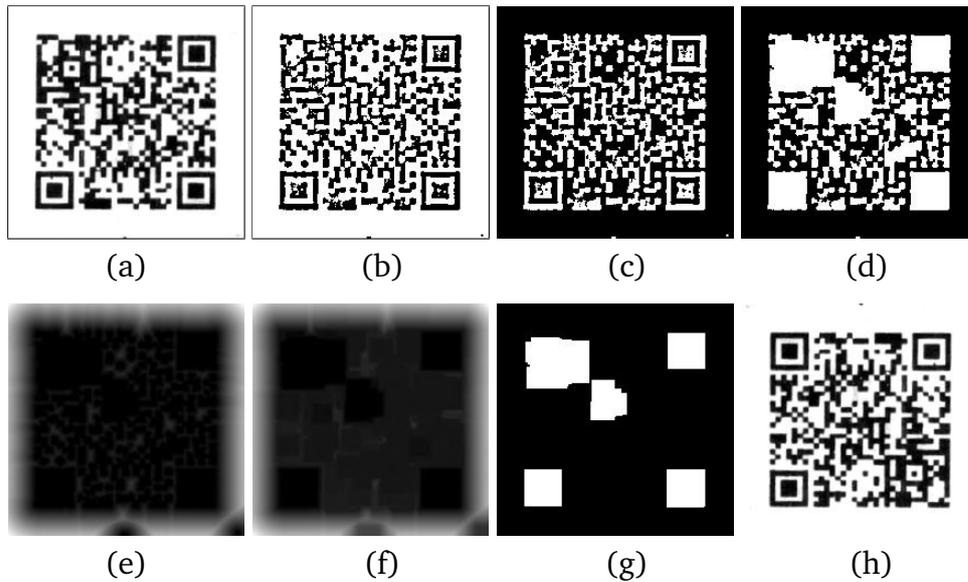


Figure 4.13: Intermediate steps of QR code orientation detection, from (a) initial QR frame, through (b) binarization, (c) complement, (d) hole filling, (e) distance map, (f) grayscale morphological closing, (g) thresholding, to (h) the reoriented frame.

localize the locator markers from the binary objects. We assume that the markers are similar in shape and their distance from the side of the image. The center of each object is determined and the minimum and maximum x and y coordinates for all objects are calculated. Using this information, we examine the location of each center point and filter out possible false regions. In most cases, only the appropriate three marker objects remain in the image. In case there are more segments in the image, those three regions are retained that the closest to the image sides, because the code surrounded by a bright border in original image so there is not any marker in that zone. Since we now work only on the segmented QR region, several priors can be used. The centroids of the locator marks shall have minimal or maximal position with respect to the others, and one of the locators in a pair along a line has minimal and the other has maximal position.

The above process is based on the assumption that the input image is not distorted to an extent that would prevent the determination of the locator mark positions. This, however, is a valid assumption, since in a real application, QR decoding is also needed and thus a certain level of image quality is required. If the application fails to detect the locators, it may report it to the user and request a new photo to be taken after repositioning the camera.

Given the segmented QR code, the NFI zone (containing the glitters) can be easily determined since the size of the QR code, the quiet zone, as well as the size and shape of the NFI zone, and even the colors used in the printing process are defined in the

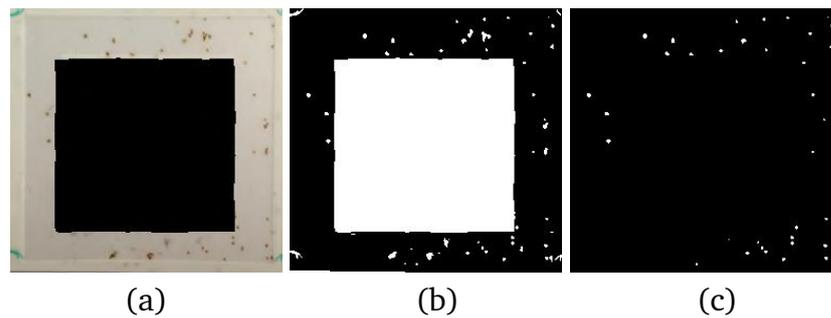


Figure 4.14: (a) limited glitter zone, (b) adaptive thresholding result, (c) segmented glitters which get into identification

specification. Based on these information about the label geometry, we can create a mask for the glitter zone and restrict glitter detection to this region (Figure 4.14(a)).

The last step of the image processing task is the detection of glitters within the NFI zone, for which we use adaptive thresholding using the same operation as in the case of QR code (Figure 4.14(b)). The wrongly detected objects are filtered out using various morphological features, such as area and circularity (Figure 4.14(c)).

Since input images can be of different resolution, it is absolutely necessary to define a normalized geometry in which the detected (glitter) features can be compared between the actual and the reference images. In our prototype setup we defined a standard coordinate system to be centered at the center of the QR code (and the NFI zone frame), aligned according to the zone edges (after rectification and standard QR orientation is applied). The scale is defined so that the glitter zone spans -1.0 to 1.0 both horizontally and vertically. In the later process, the glitter positions, sizes, and other shape features, geometrical distances are all expressed with respect to this standard coordinate system.

4.5.3 Label identification

We outline the process of identification or NFI recognition that can be performed using features extracted in the previous image processing phase which introduced in Section 4.5.2. Each of the following subtasks can be solved in various simple or more sophisticated ways:

1. find matching points,
2. calculate similarity measure,
3. decide about acceptance or rejection.

A wide range of features can be extracted from the images and be used in the above subtasks. Below is a non-exhaustive list of such features:

- number of detected glitters,
- number of matching (paired) glitters (glitter in the image paired with a glitter in the reference),
- number of outlier (not paired) glitters,
- distance between glitters (as pixel regions) or their center points,
- glitters' shape features (area, eccentricity, orientation, shortest axis, longest axis, ...),
- glitters' appearance features (grayscale/color distribution, histograms, cross-correlation, ...).

Regardless of an established standard orientation and normalized coordinate system, the extracted features are always overloaded with some noise, due to imaging artifacts and segmentation inaccuracies. These must be considered in the algorithms for identification or recognition.

In our prototype implementation we used only one image-based features, solely the normalized position of the centroid of the detected glitter. At this stage of the development this was enough for an appropriate validation. Regardless of the simplicity of this setup, the results reported below show the feasibility of such NFI technology using mobile applications. Also, we only had a small number of images of a couple of labels for this work. The evaluation is planned to be performed once the label specification is settled, and a larger number of labels and images will be available.

Similarly, we keep at a very simple matching strategy, using geometric distances between the glitter centroids. For each detected glitter we find a matching pair in the reference image such that we consider a glitter matching if it is the nearest glitter in the reference image and its distance is smaller than a threshold value. We average the pairwise distances between matched glitters and their pairs and accept an NFI if this average distance is below a given threshold. This latter threshold is smaller than the threshold used for matching individual glitters.

Note that here we only consider the task of validating an NFI label, i.e. to check whether the taken image matches its reference. So, this is not about decoding a particular glitter pattern and finding it in a large database. No specific information content is encoded in the glitters, just the pure natural features coming from the label production. Note also, that validation is done purely using the features extracted from the images, and the reference image is not used directly. The particular reference (i.e. the features extracted from the reference image(s)) can be easily selected for validation using the QR code in the middle of the NFI label.

During the parameter determination phase the input images were compared with all images in the small reference database and the thresholds were set to maximize

the true acceptance rate while minimizing false detection. In the end, the first threshold (for paring up individual glitters) were set to 0.025 and the label acceptance threshold was set to 0.0065 (both with respect to the standard coordinate system).

4.5.4 Evaluation and results

The prototype NFI label specification was the following. In an image there are two concentric and aligned square areas. The sides of the outer square are 18 mm and those of the inner square are 12 mm. The inner square is filled with a QR code in black-and-white. The outer frame is monochrome, but there are 30-60 particles made from light-reflective material (glitters), with a maximum size of 400 micron each. The glitter locations are random within the window. (See sample images in Figure 4.15).

For testing, we had 80 images captured by 6 different cameras, so image quality (including resolution, color balance, geometric distortions, and lighting conditions) varied considerably.

Images were taken by an iPhone and a Sony Xperia mobile phone. The reference images were taken by a Canon camera and the same references were used for both mobile sets. Table 4.10 shows basic settings of the cameras.

In Table 4.11, the sequential numbers in the column and row labels correspond to individual NFI labels. Elements in the major axis of tables correspond to true matches, i.e. comparison of two images taken by different cameras of the same NFI label. The values shown in the table are the average pairwise distances of the matched glitters. The values for image pairs that the algorithm accepted are shown in various boldface type and for rejected pairs are shown in normal typeface. Those values marked ***boldface italic*** are true positives (real pairs accepted by the algorithm) and those marked in **boldface underlined** are false positives (non-pairs accepted by the algorithm as a valid pair). In this rather small set of images the other type of error



Figure 4.15: Sample images of prototype NFI labels taken under various conditions. The wavy pattern on one side of the piece is an artifact of the manufacturing has nothing to do with the NFI. (These prototypes are simply not fully finished and cut out from the printing sheet.)

(false negatives, i.e. a true pair that is rejected) did not occur. In this sample the algorithm indicated only 1 mistaken identities for each camera set and 5 (all) true matches.

Table 4.12 shows the execution time of the algorithm for all images. The method is implemented in MATLAB. Execution time is dominated by the image processing operations and since matching is only to be done to a particular reference image, the identification time is basically independent of the size of the reference database.

Figure 4.16 shows image pairs and the relative positions of glitters plotted in the same reference coordinate system.

Table 4.10: Type and image settings of the devices used for the evaluation.

	Canon	iPhone	Xperia
Brand	Canon	Apple	Sony Ericsson
Model	PowerShot SX220 HS	iPhone 4S	LT15i
ISO Speed Ratings	1600	50	800
Flash	Flash not fired	Flash not fired	Flash not fired
Exposure Mode	Auto	Auto	Auto
White Balance	Auto	Auto	Auto
Image size	4000 x 3000	3264 x 2448	1920 x 1080

Table 4.11: Matching scores for 5 labels using the reference images (taken by a Canon camera) and images taken by two different cameras (iPhone (top table), Xperia (bottom table)). True positive matches are typeset in ***boldface italic*** (in the diagonal), false positives are shown in **boldface underlined** font, and the rest is correctly considered non-matches under the used parameter settings.

	iPhone_1	iPhone_2	iPhone_3	iPhone_4	iPhone_5
Canon_1	<i>0.0065</i>	0.0073	0.0075	0.0099	0.0066
Canon_2	0.0099	<i>0.0017</i>	0.0066	0.0080	0.0067
Canon_3	0.0093	0.0078	<i>0.0042</i>	0.0077	0.0066
Canon_4	0.0086	<u>0.0057</u>	0.0073	<i>0.0034</i>	0.0067
Canon_5	0.0093	0.0079	0.0072	0.0093	<i>0.0054</i>

	Xperia_1	Xperia_2	Xperia_3	Xperia_4	Xperia_5
Canon_1	<i>0.0062</i>	0.0076	0.0069	0.0083	0.0096
Canon_2	0.0112	<i>0.0027</i>	0.0086	0.0096	<u>0.0051</u>
Canon_3	0.0097	0.0069	<i>0.0045</i>	0.0105	0.0072
Canon_4	0.0094	<u>0.0062</u>	0.0087	<i>0.0058</i>	<u>0.0065</u>
Canon_5	0.0100	0.0083	0.0072	0.0097	<i>0.0043</i>

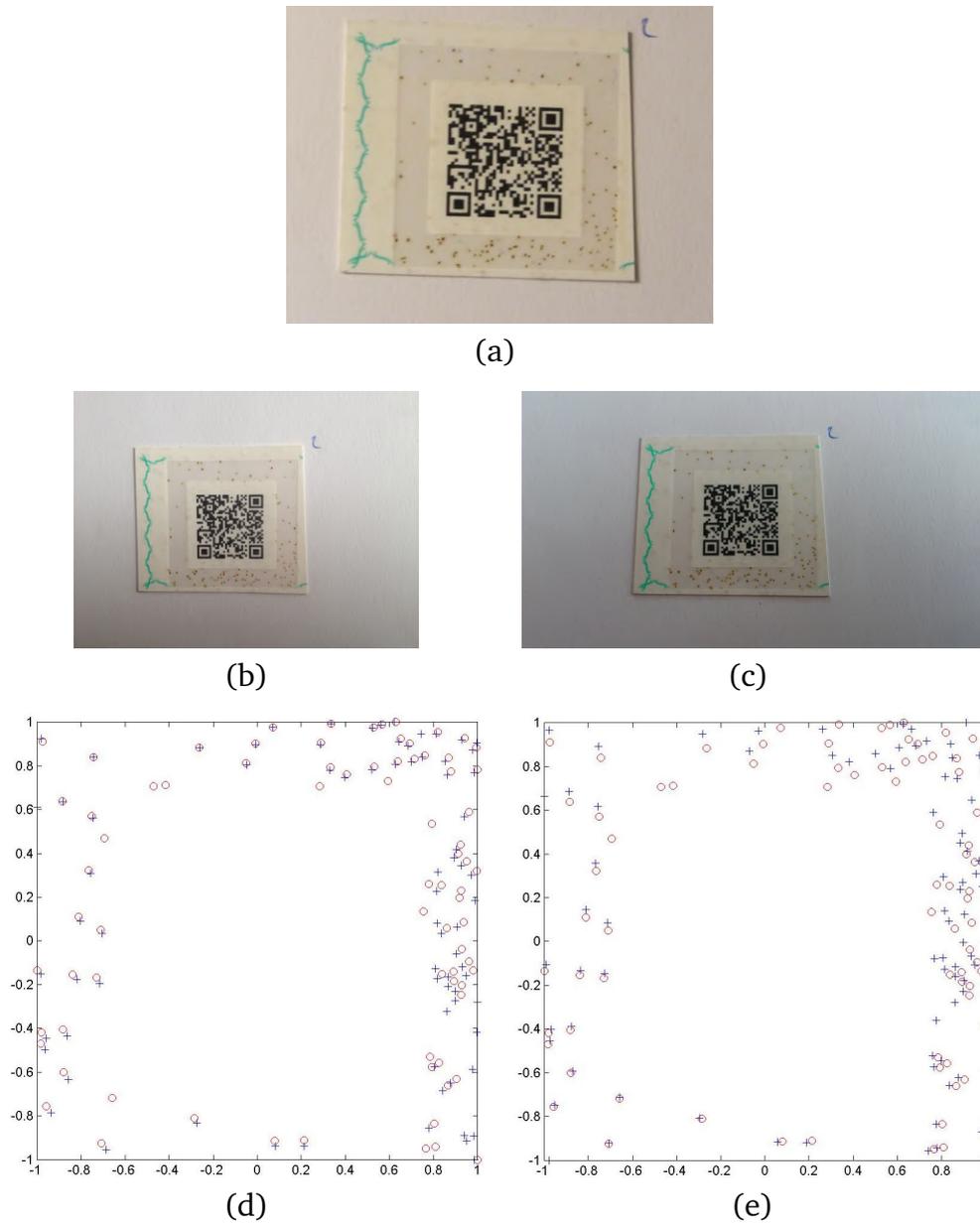


Figure 4.16: Images of the same labels: (a) sample (Canon), (b) test (iPhone), (c) test (Xperia). Detected glitter positions in the normalized coordinate system: (d) Canon & iPhone, (e) Canon & Xperia. Red circle denotes the normalized reference points and blue crosses are the normalized detected coordinates.

Table 4.12: Execution time statistics of the full process (mean, 95% confidence interval)

Operation	Computing time (sec/image)	
	Mean	95% CI
Image processing	1.1129	[1.0915, 1.1343]
Verification	3.0×10^{-4}	$[2.24, 3.76] \times 10^{-4}$
Image processing + verification	1.1133	[1.0918, 1.1347]

4.6 Concluding remarks

In this chapter we introduced an algorithm and its improved version for detection of traditional 1D barcodes and stacked 2D codes. Barcodes have a specific pattern thus we proposed an efficient localization method based on pattern matching for 1D codes and another one approach for QR codes. We also presented an image processing algorithm to implement natural feature identifier recognition. The method is robust against variations in image quality, resolution, lighting, and positioning. It automatically detects the QR code, determines the transformation to normalize the QR-NFI label, detects the NFI glitters within the NFI zone and matches with references using a simple strategy.

Bibliography

- [1] Kovács A., Dégi R., **Katona M.**, Nyúl L. G., Varga L., Dombi J., and Grósz T. Biomarkerek automatikus karakterizálása OCT felvételek időbeli elemzésére. In *Magyar Szemorvostársaság Retina Szekciójának Kongresszusa*, page 1 p., 2017.
- [2] A. M. Abhishek, T. T. J. M. Berendschot, S. V. Rao, and S. Dabir. Segmentation and analysis of retinal layers (ILM & RPE) in Optical Coherence Tomography images with Edema. In *Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on*, pages 204–209, 2014.
- [3] R. Adelman, M. Langheinrich, and C. Flörkemeier. A toolkit for bar-code-recognition and - resolving on camera phones – jump starting the internet of things. In *In Workshop Mobile and Embedded Interactive Systems (MEIS'06) at Informatik*, 2006.
- [4] C. Ahlers, C. Simader, W. Geitzenauer, G. Stock, P. Stetson, S. Dastmalchi, and U. Schmidt-Erfurth. Automatic segmentation in three-dimensional analysis of fibrovascular pigmentepithelial detachment using high-definition optical coherence tomography. *British Journal of Ophthalmology*, 92(2):197–203, 2008.
- [5] J. M. Bailey, B. A. Creamer, and M. A. Hollingsworth. What a Fish Can Learn from a Mouse: Principles and Strategies for Modeling Human Cancer in Mice. *Zebrafish*, 6(4):329–337, 2009.
- [6] W. Barkhoda, F. Akhlaqian, M. D. Amiri, and M. S. Nouroozzadeh. Retina Identification Based on the Pattern of Blood Vessels Using Fuzzy Logic. *EURASIP Journal on Advances in Signal Processing*, pages 1–8, 2011.
- [7] G. T. Bartoo, Y. Kim, R. M. Haralick, D. Nochlin, and S. M. Sumi. Mathematical Morphology Techniques For Image Processing Applications In Biomedical Imaging. In *Medical Imaging II.*, volume 0914, 1988.
- [8] L. F. F. Belussi and N. S. T. Hirata. Fast QR Code Detection in Arbitrarily Acquired Images. In *Graphics, Patterns and Images (Sibgrapi), 2011 24th SIB-GRAPI Conference on*, pages 281–288, 2011.

- [9] R. Besenczi, J. Tóth, and A. Hajdú. A review on automatic analysis techniques for color fundus photographs. *Computational and structural biotechnology*, 14: 371–384, 2016.
- [10] P. Bodnár and Nyúl L. G. Barcode Detection with Morphological Operations and Clustering. In *Signal Processing, Pattern Recognition, and Applications, Proceedings of the Ninth IASTED International Conference on*, pages 51–57, 2012.
- [11] P. Bodnár and L. G. Nyúl. Improving Barcode Detection with Combination of Simple Detectors. In *8th International Conference on Signal Image Technology and Internet Based Systems, SITIS 2012*, pages 300–306, 2012.
- [12] P. Bodnár and L. G. Nyúl. Barcode detection with uniform partitioning and distance transformation. *IASTED International Conference on Computer Graphics and Imaging*, pages 48–53, 2013.
- [13] P. Bodnár and L. G. Nyúl. QR Code Localization Using Boosted Cascade of Weak Classifiers. In *Image Analysis and Recognition*, pages 338–345. 2014.
- [14] P. Burt and E. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.
- [15] W.-S. Chen, K.-H. Chih, S.-W. Shih, and C.-M. Hsieh. Personal Identification Technique Based on Human Iris Recognition with Wavelet Transform. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, pages 949–952, 2005.
- [16] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu. Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. *Optics Express*, 18(18):19413–19428, 2010.
- [17] C. Creusot and A. Munawar. Real-Time Barcode Detection in the Wild. In *2015 IEEE Winter Conference on Applications of Computer Vision*, pages 239–245, 2015.
- [18] P.-E. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14(3):227–248, 1980.
- [19] R. Dégi, A. Kovács, **M. Katona**, L. G. Nyúl, L. G. Varga, J. Dombi, and T. Grósz. Quantitative assessment of the manifestation of age-related macular degeneration via automated image analysis of spectral-domain optical coherence tomography. In *European Society of Retina Specialists 15th EURETINA Congress*, page 1 p., 2015.

- [20] W. Ding, M. Young, S. Bourgault, S. Lee, D. A. Albiani, A. W. Kirker, F. Forooghian, M. V. Sarunic, A. B. Merkur, and M. F. Beg. Automatic detection of subretinal fluid and sub-retinal pigment epithelium fluid in optical coherence tomography images. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7388–7391, 2013.
- [21] B. I. Dodo, Y. Li, K. Eltayef, and X. Liu. Graph-Cut Segmentation of Retinal Layers from OCT Images. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: BIOIMAGING*,, pages 35–42, 2018.
- [22] J. Dombi. *Modalities*, pages 53–65. Springer Berlin Heidelberg, 2012.
- [23] M. Fausto, N. Nassir, and A. Seyed-Ahmad. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571, 2016.
- [24] A. F. Fercher, C. K. Hitzenberger, W. Drexler, G. Kamp, and H. Sattmann. In Vivo Optical Coherence Tomography. *American Journal of Ophthalmology*, 116(1):113–114, 1993.
- [25] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI’98*, pages 130–137, 1998.
- [26] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman. Blood vessel segmentation methodologies in retinal images – A survey . *Computer Methods and Programs in Biomedicine*, 108(1): 407–433, 2012.
- [27] D. Fu, H. Tong, L. Luo, and F. F. Gao. Retinal automatic segmentation method based on prior information and optimized boundary tracking algorithm. In *Proc. SPIE*, volume 10033, pages 100331C:1–100331C:6, 2016.
- [28] O. Gallo and R. Manduchi. Reading 1D Barcodes with Mobile Phones Using Deformable Templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(9):1834–1843, 2011.
- [29] M. K. Garvin, M. D. Abramoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka. Automated 3-D Intraretinal Layer Segmentation of Macular Spectral-Domain Optical Coherence Tomography Images. *IEEE Transactions on Medical Imaging*, 28(9):1436–1447, 2009.

- [30] G. A. Geoffrey, W. Fu, and K. D. Gary. Temozolomide-Mediated Radiosensitization of Human Glioma Cells in a Zebrafish Embryonic System. *Cancer Research*, 68(9):3396–3404, 2008.
- [31] T. Goldstein, X. Bresson, and S. Osher. Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction. *Journal of Scientific Computing*, 45(1):272–293, 2010.
- [32] R. C. Gonzalez and R. E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2006.
- [33] R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital Image Processing Using MATLAB*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2003. ISBN 0130085197.
- [34] D. K. Hansen, K. Nasrollahi, C. B. Rasmussen, and T. B. Moeslund. Real-Time Barcode Detection and Classification using Deep Learning. In *IJCCI*, pages 321–327, 2017.
- [35] B. Harangi, J. Tóth, and A. Hajdú. Fusion of Deep Convolutional Neural Networks for Microaneurysm Detection in Color Fundus Images. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3705–3708, 2018.
- [36] B. Hassan, G. Raja, T. Hassan, and M. U. Akram. Structure tensor based automated detection of macular edema and central serous retinopathy using optical coherence tomography images. *Journal of the Optical Society of America*, 33(4):455–463, 2016.
- [37] M. R. Hee, C. R. Baumal, C. A. Puliafito, J. S. Duker, E. Reichel, J. R. Wilkins, J. G. Coker, J. S. Schuman, E. A. Swanson, and J. G. Fujimoto. Optical Coherence Tomography of Age-related Macular Degeneration and Choroidal Neovascularization. *Ophthalmology*, 103(8):1260–1270, 1996.
- [38] D. Huang, E.A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and et al. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.
- [39] O. Ishaq, J. Negri, M.-A. Bray, A. Pacureanu, R. T. Peterson, and C. Wählby. Automated quantification of Zebrafish tail deformation for high-throughput drug screening. *2013 IEEE 10th International Symposium on Biomedical Imaging*, pages 902–905, 2013.

- [40] R.B. Jarvis and J.F. Knowles. DNA damage in zebrafish larvae induced by exposure to low-dose rate γ -radiation: detection by the alkaline comet assay. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 541(1): 63–69, 2003.
- [41] J. Juett and X. Qi. Barcode localization using bottom-hat filter. *NSF Research Experience for Undergraduates*, 2005.
- [42] C. Kalifa and J. Grill. The Therapy of Infantile Malignant Brain Tumors: Current Status? *Journal of Neuro-Oncology*, 75(3):279–285, 2005.
- [43] M. K. Khan, J. Zhang, and X. Wang. Chaotic Hash-Based Fingerprint Biometric Remote User Authentication Scheme on Mobile Devices . *Chaos, Solitons & Fractals*, 35:519–524, 2008.
- [44] S. Kong. QR Code Image Correction based on Corner Detection and Convex Hull Algorithm. *Journal of Multimedia*, 8:662–668, 2013.
- [45] Varga L., **Katona M.**, Grósz T., Dombi J., Kovács A., Dégi R., and Nyúl L. G. Időskori makula degeneráció kvantitatív jellemzése SD-OCT képek automatikus elemzésével. In *Új alapokon az egészségügyi informatika : A XXVIII. Neumann Kollokvium konferencia-kiadványa*, pages 43–48, 2015.
- [46] L. Lam, S. Lee, and C. Y. Suen. Thinning methodologies-a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(9): 869–885, 1992.
- [47] A. Lang, A. Carass, M. Hauser, E. S. Sotirchos, P. A. Calabresi, H. S. Ying, and J. L. Prince. Retinal layer segmentation of macular OCT images using boundary classification. *Biomedical Optics Express*, 4(7):1133–1152, 2016.
- [48] V. Larouche, A. Huang, U. Bartels, and E. Bouffet. Tumors of the central nervous system in the first year of life. *Pediatric Blood & Cancer*, 49(7):1074–1082, 2007.
- [49] C. S. Lee, A. J. Tying, N. P. Deruyter, Y. Wue, A. Rokem, and A. Y. Lee. Deep-learning based, automated segmentation of macular edema in optical coherence tomography. *Biomedical Optics Express*, 8(7):3440–3448, 2017.
- [50] H. Lee, K. E. Kang, H. Chung, and H. C. Kim. Automated Segmentation of Lesions Including Subretinal Hyperreflective Material in Neovascular Age-related Macular Degeneration. *American Journal of Ophthalmology*, 191:64–75, 2018.

- [51] L. S. Lim, P. M., J. M. Seddon, F. G. Holz, and T. Y. Wong. Age-related macular degeneration. *The Lancet*, 379(9827):1728–1738, 2012.
- [52] D-T. Lin and C-L. Lin. Automatic location for multi-symbology and multiple 1d and 2d barcodes. *Journal of Marine Science and Technology*, 21:663–668, 2013.
- [53] S. Lu, C. Y. I. Cheung, J. Liu, J. H. Lim, C. K. s. Leung, and T. Y. Wong. Automated Layer Segmentation of Optical Coherence Tomography Images. *IEEE Transactions on Biomedical Engineering*, 57(10):2605–2608, 2010.
- [54] T. Mahashwari and A. Asthana. Image Enhancement Using Fuzzy Technique. *International Journal of Research Review in Engineering Science and Technology*, 2(2):1–4, 2013.
- [55] M. F. McAleer, C. Davidson, W. R. Davidson, B. Yentzer, Steven A. Farber, U. Rodeck, and A. P. Dicker. Novel use of zebrafish as a vertebrate model to screen radiation protectors and sensitizers. *International Journal of Radiation Oncology*Biography*Physics*, 61(1):10–13, 2005.
- [56] I. Megyeri, **M. Katona**, and L. G. Nyúl. A Novel Approach to Detect Outer Retinal Tubulation using U-Net in SD-OCT images. In *The 15th International Conference on Signal Image Technology & Internet Based Systems*, 2019. Submitted for publication.
- [57] A. Mittal, R. Soundararajan, and A. C. Bovik. Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2013.
- [58] H. Narasimha-Iyer, A. Can, B. Roysam, V. Stewart, H. L. Tanenbaum, A. Majerovics, and H. Singh. Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy. *IEEE Transactions on Biomedical Engineering*, 53(6):1084–1098, 2006.
- [59] L. Ngo, J. Cha, and J. Han. Deep Neural Network Regression for Automated Retinal Layer Segmentation in Optical Coherence Tomography Images. *IEEE Transactions on Image Processing*, pages 1–1, 2019.
- [60] J. Novosel, Z. Wang, H. de Jong, M. van Velthoven, K. A. Vermeer, and L. J. van Vliet. Locally-adaptive loosely-coupled level sets for retinal layer and fluid segmentation in subjects with central serous retinopathy. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 702–705, 2016.
- [61] E. Ohbuchi, H. Hanaizumi, and L. A. Hock. Barcode Readers Using the Camera Device in Mobile Phones. In *Proceedings of the 2004 International Conference on Cyberworlds, CW '04*, pages 260–265, 2004.

- [62] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.
- [63] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7): 629–639, 1990.
- [64] C. Pylatiuk, D. Sanchez, R. Mikut, R. Alshut, M. Reischl, S. Hirth, W. Rottbauer, and S. Just. Automatic Zebrafish Heartbeat Detection and Analysis for Zebrafish Embryos. *Zebrafish*, 11(4):379–383, 2014.
- [65] O. Ronneberger and et al. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241. Springer International Publishing, 2015.
- [66] A. G. Roy, S. Conjeti, S. P. K. Karri, D. Sheet, A. Katouzian, C. Wachinger, and N. Navab. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed Opt Express*, 8(8):3627–3642, 2017.
- [67] J. Sahlsten, J. Jaskari, J. Kivinen, L. Turunen, E. Jaanio, K. Hietala, and K. Kaski. Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading. *Scientific Reports*, 9:10750, 2019.
- [68] M. B. Salah, A. Mitiche, and I. B. Ayed. Multiregion Image Segmentation by Parametric Kernel Graph Cuts. *IEEE Transactions on Image Processing*, 20(2): 545–557, 2011.
- [69] R. W. Schafer. What Is a Savitzky-Golay Filter? *IEEE Signal Processing Magazine*, 28(4):111–117, 2011.
- [70] F. Shi, L. Wang, W. Zhu, D. Xiang, C. Yan, L. Li, and X. Chen. An automated framework for 3d serous pigment epithelium detachment segmentation in sd-oct images. *Scientific reports*, 6:21739, 2016.
- [71] G. Sörös and C. Flörkemeier. Blur-resistant Joint 1D and 2D Barcode Localization for Smartphones. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*, pages 11:1–11:8, 2013.
- [72] G. Sörös, S. Semmler, L. Humair, and O. Hilliges. Fast Blur Removal for Wearable QR Code Scanners. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ISWC ’15, pages 117–124, 2015.

- [73] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.
- [74] E. R. Szabó, I. Plangár, T. Tóké, I. Mán, R. Polanek, R. Kovács, G. Fekete, Z. Szabó, Zs. Csenki, F. Baska, and K. Hideghéty. l-Alpha Glycerylphosphorylcholine as a Potential Radioprotective Agent in Zebrafish Embryo Model. *Zebrafish*, 13(6):481–488, 2016.
- [75] I. Szentandrás, A. Herout, and M. Dubská. Fast Detection and Recognition of QR Codes in High-resolution Images. In *Proceedings of the 28th Spring Conference on Computer Graphics, SCCG '12*, pages 129–136, 2013.
- [76] E. Tekin and J. Coughlan. BLaDE: Barcode localization and decoding engine. Technical report, Technical Report 2012-RERC, 2012.
- [77] E. Tekin and J. M. Coughlan. An algorithm enabling blind users to find and read barcodes. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–8, 2009.
- [78] **Katona M.** and Nyúl L. G. Vonalkódok és természetes vizuális azonosítók felismerése valós időben. In *A Képfeldolgozók és Alakfelismerők Társaságának 10. országos konferenciája - KÉPAF 2015*, pages 562–577, 2015.
- [79] **Katona M.** and Nyúl L. G. Szubretinális folyadékter és ciszta automatikus detektálása retinafelvételeken. In *A Képfeldolgozók és Alakfelismerők Társaságának 11. országos konferenciája - KÉPAF 2015*, page Paper 51. 11 p., 2017.
- [80] **Katona M.**, T. Tóké, E. R. Szabó, Sz. Brunner, R. Polanek, Hideghéty K., and Nyúl L. G. Sugárkezelésen átesett zebrahal embriók automatikus morfológiai elemzése. In *A Képfeldolgozók és Alakfelismerők Társaságának 12. országos konferenciája - KÉPAF 2019*, page Paper 21. 12 p., 2019.
- [81] **M. Katona** and L. G. Nyúl. A Novel Method for Accurate and Efficient Barcode Detection with Morphological Operations. In *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, pages 307–314, 2012.
- [82] **M. Katona** and L. G. Nyúl. Efficient 1D and 2D Barcode Detection Using Mathematical Morphology. In *Mathematical Morphology and Its Applications to Signal and Image Processing*, pages 464–475. Springer Berlin Heidelberg, 2013.

- [83] **M. Katona** and L. G. Nyúl. Improved 1D and 2D barcode detection with morphological operations. In *A Képfeldolgozók és Alakfelismerők Társaságának 9. országos konferenciája - KÉPAF 2013*, pages 309–324, 2013.
- [84] **M. Katona** and L. G. Nyúl. Fast recognition of natural feature identifiers by a mobile phone. In *The 9th Conference of PhD Students in Computer Science (CSCS 2014) : Volume of Extended Abstracts*, page 26, 2014.
- [85] **M. Katona** and L. G. Nyúl. Fast Recognition of Natural Feature Identifiers by a Mobile Phone. *Acta Cybernetica*, 22:101–116, 2015.
- [86] **M. Katona** and L. G. Nyúl. Quantitative Assessment of Retinal Layer Distortions and Subretinal Fluid in SD-OCT Images. In *The 10th Conference of PhD Students in Computer Science (CSCS 2016) : Volume of Extended Abstracts*, page 37, 2016.
- [87] **M. Katona** and L. G. Nyúl. An Approach to the Quantitative Assessment of Retinal Layer Distortions and Subretinal Fluid in SD-OCT Images. *Acta Cybernetica*, 23:615–628, 2017.
- [88] **M. Katona**, P. Bodnár, and L. G. Nyúl. Distance Transform and Template Matching Based Methods for Localization of Barcodes and QR Codes. *Computer Science and Information Systems*. doi: 10.2298/CSIS181011020K.
- [89] **M. Katona**, A. Kovács, R. Dégi, and L. G. Nyúl. Automatic Detection of Subretinal Fluid and Cyst in Retinal Images. In *Image Analysis and Processing – ICIAP 2017*, pages 606–616, 2017.
- [90] **M. Katona**, A. Kovács, L. Varga, T. Grósz, J. Dombi, R. Dégi, and L. G. Nyúl. Automatic detection and characterization of biomarkers in oct images. In Aurélio Campilho, Fakhri Karray, and Bart ter Haar Romeny, editors, *Image Analysis and Recognition – ICIAR 2018*, pages 706–714, Cham, 2018. Springer International Publishing. doi: 10.1007/978-3-319-93000-8_80.
- [91] **M. Katona**, T. Tókécs, E. R. Szabó, Sz. Brunner, I. Z. Szabó, R. Polanek, K. Hideghéty, and L. G. Nyúl. Automatic Segmentation and Quantitative Analysis of Irradiated Zebrafish Embryos. In *Computational Modeling of Objects Presented in Images. Fundamentals, Methods, and Applications*, pages 95–107, 2019.
- [92] **M. Katona**, A. Kovács, R. Dégi, and L. G. Nyúl. Segmentation of Subretinal Hyperreflective Material and Pigment Epithelial Detachment Using Kernel Graph Cut. In *Progress in Computer Recognition Systems*, pages 98–105. Springer International Publishing, 2020.

- [93] Y. Tian, Z. Che, G. Zhai, and Z. Gao. BAN, A Barcode Accurate Detection Network. In *2018 IEEE Visual Communications and Image Processing (VCIP)*, pages 1–5, 2018.
- [94] T. R. Tuinstra. *Reading Barcodes from Digital Imagery*. PhD thesis, Cedarville University, 2006.
- [95] V. Varjas and A. Tanács. Car recognition from frontal images in mobile environment. In *Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on*, pages 819–823, 2013.
- [96] R. Velez-Montoya and et al. Current knowledge and trends in age-related macular degeneration: today’s and future treatments. *Retina*, 334:1487–1502, 2013.
- [97] N. N. Ventsov and L. A. Podkolzina. Localization of Barcodes Using Artificial Neural Network. *2018 IEEE East-West Design & Test Symposium (EWDTS)*, pages 1–6, 2018.
- [98] J. Walther, M. Gaertner, P. Cimalla, A. Burkhardt, L. Kirsten, S. Meissner, and E. Koch. Optical coherence tomography in biomedical research. *Analytical and Bioanalytical Chemistry*, 400(9):2721–2743, 2011.
- [99] J. Wang, M. Zhang, A. D. Pechauer, L. Liu, T. S. Hwang, D.J. Wilson, D. Li, and Y. Jia. Automated volumetric segmentation of retinal fluid on optical coherence tomography. *Biomedical Optics Express*, 7(4):1577–1589, 2016.
- [100] X. Wang, E. Cheng, I. S. Burnett, Y. Huang, and D. Wlodkowic. Automatic multiple zebrafish larvae tracking in unconstrained microscopic video conditions. In *Scientific Reports*, volume 7, 2017.
- [101] W. Wieclawek. Automatic cysts detection in optical coherence tomography images. In *Mixed Design of Integrated Circuits Systems (MIXDES), 2015 22nd International Conference*, pages 79–82, 2015.
- [102] G. R. Wilkins, O. M. Houghton, and A. L. Oldenburg. Automated Segmentation of Intraretinal Cystoid Fluid in Optical Coherence Tomography. *IEEE Transactions on Biomedical Engineering*, 59(4):1109–1114, 2012.
- [103] T. Wu, J. Lu, Y. Lu, T. Liu, and J. Yang. Embryo zebrafish segmentation using an improved hybrid method. *Journal of Microscopy*, 250(1):68–75, 2013.
- [104] X. Xu, K. Lee, L. Zhang, M. Sonka, and M. D. Abràmoff. Stratified Sampling Voxel Classification for Segmentation of Intraretinal and Subretinal Fluid in

- Longitudinal Clinical OCT Data. *IEEE Transactions on Medical Imaging*, 34(7): 1616–1623, 2015.
- [105] Y. Xu, K. Yan, J. Kim, X. Wang, C. Li, L. Su, S. Yu, X. Xu, and D. D. Feng. Dual-stage deep learning framework for pigment epithelium detachment segmentation in polypoidal choroidal vasculopathy. *Biomed Opt Express*, 8(9): 4061–4076, 2017.
- [106] A. Yazdanpanah, G. Hamarneh, B. Smith, and M. Sarunic. *Intra-retinal Layer Segmentation in Optical Coherence Tomography Using an Active Contour Approach*, pages 649–656. Springer Berlin Heidelberg, 2009.
- [107] I. Yun and J. Kim. Vision-based 1D barcode localization method for scale and rotation invariant. In *TENCON 2017 - 2017 IEEE Region 10 Conference*, pages 2204–2208, 2017.
- [108] A. Zamberletti, I. Gallo, M. Carullo, and E. Binaghi. Neural Image Restoration for Decoding 1-D Barcodes using Common Camera Phones. In *Computer Vision, Imaging and Computer Graphics. Theory and Applications*, pages 5–11, 2010.
- [109] H. Zhao, J. Zhou, A. Robles-Kelly, J. Lu, and J. Y. Yang. Automatic Detection of Defective Zebrafish Embryos via Shape Analysis. In *Digital Image Computing: Techniques and Applications*, pages 431–438, 2009.
- [110] Y. Zheng, J. Sahni, A. Raj C. Campa, A. N. Stangos, and S. P. Harding. Computerized Assessment of Intraretinal and Subretinal Fluid Regions in Spectral-Domain Optical Coherence Tomography Images of the Retina. *American Journal of Ophthalmology*, 155:277–286, 2012.
- [111] S. A. Zweifel, M. Engelbert, K. Laud, R. Margolis, R. F. Spaide, and K. B. Freund. Outer Retinal Tubulation: A Novel Optical Coherence Tomography Finding. *JAMA Ophthalmology*, 127(12):1596–1602, 2009.

Summary

In this dissertation, we deal with the various problems of image analysis in areas of medical imaging based diagnostics, microscopy in biology and localization of visual codes using classical image processing approaches, but we also applied a deep learning approach. The classical methods are well applicable for various detection and segmentation tasks, do not require a large set of image data which was not available during our work.

The dissertation consist of three main parts: analysis of retinal images in Chapter 2, detection of morphological changes in zebrafish embryos in Chapter 3, and localization of visual codes in Chapter 4.

Retinal Image Analysis

In Chapter 2 we dealt with age-related macular degeneration (AMD) which is one of the leading causes of age-related visual impairment. Optical coherence tomography (OCT) is a crucial and essential tool to diagnose the disease and to evaluate the treatment efficiency. It can also render the 3D structure of the retina layer by layer. OCT shows many features of eye diseases, including AMD. In recent years, many methods have been developed to detect these features, most of them trying to solve the current problem using neural networks.

To help segment the specific biomarkers, we proposed two methods for delimiting the retina area. The first one is based on the analysis of vertical profiles to highlight boundary layers, because there is an intensity difference between the layers of the retina. This approach provided many inadequate results in RPE layer detection when AMD caused highly distorted layers, hence we developed a new approach that relies on graph cut. Retinal layer detection algorithms were evaluated in two different ways. Our algorithms were compared against manual delineations from experts and re-implemented literature methods. We calculated the mean, maximum and standard deviation of boundary errors for each surface. The mean error of our graph cut based algorithm is less than 3 pixels in all cases that shows our approach is more stable than the re-implemented algorithms.

Usually the images are affected by different distortions during the OCT examination. Noise and artifacts due to shadows of blood vessels may appear in cross-

sectional layers. We introduced a method that detects blood vessels in scanning laser ophthalmoscope (SLO) images. Then the positions of these blood vessels can be determined in the B-scan image for exclusions from further investigations. The extent of distortion due to the disease is also important for medical diagnosis and a method has been suggested to determine this.

Many approaches in literature focus on automatic detection of subretinal (SRF) and intraretinal (IRF) fluid in OCT images. We proposed a method based on the analysis of vertical profiles to delineate the boundaries of retina layers. Since the hyperreflectivity of SRF and IRF are similar to that of the vitreous body, we developed a method for detecting and separating them simultaneously. The annotations by ophthalmologists were not yet available at this stage of the research so we only did visual comparison of our proposed and re-implemented methods.

Due to retinal diseases pigment epithelial detachment (PED) and subretinal hyperreflective material (SHRM) may occur, thus a method was presented to detect and separate these biomarkers. We evaluated our method on annotated dataset using Dice coefficient and Sensitivity metrics. In many cases SHRM is not clearly detectable, but the average Dice coefficients are above 0.75 and the sensitivity is 0.93 for PED and 0.77 for SHRM.

Only a few publications deal with the localization of subretinal material, and none of them is focusing on segmenting outer retinal tubulation, but publications in the medical field indicate that cardinality of these tubulations is an important descriptor of the disease for doctors. The problem was investigated in two ways: using a classical image processing procedure and applying a nowadays popular deep neural network approach. In order to achieve widespread validation, we calculated various metrics both for automatic and for manual segmentations. We measured the global and object level Dice coefficient and Recall. Our method achieves 0.847 object level Recall and 0.583 Dice coefficient, which is comparable to human professionals.

Quantitative Analysis of Irradiated Zebrafish Embryos

An automated method was presented for the analysis of irradiation treated zebrafish embryos in Chapter 3. Due to irradiation, dose-dependent morphological changes are observable in the embryos. The advantage of using the zebrafish vertebrate model is the biological changes assessment on large number of embryos, enabling reliable statistical analysis. However, the more sophisticated morphological measurements — beyond the simple survival detection — are extremely time- and labor-intensive. We introduced a framework to reduce human effort for faster experimentation. We assumed that a protocol was followed during the experiments, including the placement and handling of embryos and taking microscopic photos. The automated analysis consists of several steps. It is expected that at least some part of the circular

wall of the well containing the fish is observable in the images thus a method was proposed to delineate the well region. We presented an approach to localize fish and distinguish head and tail parts from each other. The endpoints of the previously segmented fish object, the total length of the body, and the location and diameter of the eyes were also determined. During the evaluation of 393 images, no significant difference was found between the results of manual measurement and the automatic method.

Localization of Visual Codes

In Chapter 4, we discussed our proposed methods for detecting different visual codes, including traditional 1D barcodes and 2D QR, and 2D stacked types. We introduced two methods based on morphological operations for efficient localization of classical 1D codes. The results of the algorithms were compared with other literature methods. From the evaluation outcome we can conclude that the proposed procedures had second best running efficiency for all test images and in many cases it also showed best accuracy. The weakness of the method appears when the image is very noisy or when there are such image areas which are similar to a barcode.

The 1D barcodes consists of parallel bars and the bars are located at varying distance from each other. This structural feature was used in our detection method based on template matching. We found that the stacked 2D codes have some similarities in pattern with the conventional ones, so that some 2D codes, such as PDF417, Codablock, can also be localized by the proposed method. Choosing another template, such a method can also be adapted to determine the location of QR codes. Our approaches were compared against other algorithms in the literature using the Muenster database as a benchmark. We investigated and found that our approaches are fast and not sensitive to rotation in the case of 1D codes. Efficiency of our QR template matching approach was compared with methods on the Dubská datasets and the proposed algorithm performed well.

Codes are designed using commonly used geometric patterns. However, patterns that are unique in nature and can therefore be used to verify originality or authenticity can be produced. We introduced a possible hybrid visual code construct in which identification and authentication can be achieved using a combination of artificial and natural patterns, so that the task can be divided into three main steps: imaging, image processing and identification. We proposed a pipeline to determine the position and orientation of the QR code, delineate the zone containing the glitters on the label and detect the glitters. A simple pairing strategy were proposed for authentication. Image analysis proved to be fast on mobile devices.

Contributions of the thesis

Contributions in the first group are related to the investigation of various biomarkers of age-related macular degeneration on Optical Coherence Tomography (OCT) images. (Detailed discussion in Chapter 2.)

- I/1. I proposed two new approaches to define internal limiting membrane (ILM) and retinal pigment epithelial (RPE).
- I/2. I developed a new method to determine the position and extent of the distortion which is an important feature in the diagnosis of AMD.
- I/3. I proposed two new algorithms to detect subretinal (SRF) and intraretinal fluid (IRF).
- I/4. I developed an approach to determine pigment epithelial detachment (PED) and subretinal hyperreflective material (SHRM).
- I/5. I proposed a new method based on classical operations and another one on relies on convolutional neural network to localize outer retinal tubulation (ORT).

Contributions in the second thesis point are related to automatic detection of morphological deteriorations of irradiated zebrafish embryos in microscopic images used in radiobiological research. (Detailed discussion in Chapter 3.)

- II. I developed a new algorithm to delimitate the zone of the well and obtain various morphological features of zebrafish: I determined the end points and length of the embryos, the position and diameter of the eyes, and the orientation of the fish.

Contributions in the third group are related to localization of visual codes in digital images. (Detailed discussion in Chapter 4.)

- III/1. I proposed two new methods based on morphological operations to localize traditional 1D and stacked 2D barcodes.
- III/2. I developed a new approach to detect classical 1D barcodes and another one to segment QR codes using pattern matching.
- III/3. I proposed a method to localize and recognize a hybrid visual code (NFI).
- III/4. I performed an extensive comparative analysis of 12 code localization methods on publicly available and on our databases of tens of thousands of real and synthetic images depicting 1D and 2D codes.

Összefoglalás

Az értekezésben a képanalízis különféle problémáival foglalkoztam az orvosi kép alapú diagnosztika, biológiai mikroszkópia és a vizuális kódok lokalizálása területén. A klasszikus képfeldolgozási megközelítések jól alkalmazhatók különféle detektálási és szegmentálási feladatokra és ezek nem követelik meg a nagy mennyiségű képi adatot, amely a munkánk során eleve nem volt elérhető.

Az értekezés három főbb témát dolgoz fel: retina felvételek elemzése a 2. fejezetben, zebraalak embriók morfológiai elváltozásainak detektálása a 3. fejezetben, valamint vizuális kódok lokalizálása a 4. fejezetben.

Retina felvételek elemzése

A 2. fejezetben az időskori látásromlás egyik fő vezető okával, az időskori makuladegenerációval (AMD) foglalkoztunk. Az optikai koherencia tomográfiának (OCT) a betegség felismerésében és a kezelés hatékonyságának megítélésében meghatározó és elengedhetetlen szerepe van. A retina egyes rétegeinek 3D felépítése is leképezhető. Az OCT felvételeken jól láthatóak az egyes betegségek, így az AMD biomarkerei is. Az elmúlt években számos eljárás született ezeknek a jellemzőknek a detektálására, melyek közül a legtöbb már valamilyen neurális hálózat segítségével igyekszik megoldani az aktuális problémát.

A konkrét biomarkerek szegmentálása előtt a retina behatárolására javasoltunk kétféle módszert. Az egyik esetben arra támaszkodtunk, hogy a retina egyes rétegei között intenzitásbeli különbség van és ezáltal a vertikális profilok elemzésével a határoló rétegek kiemelhetőek. Azokban az esetekben, amikor a betegség következtében a rétegek nagyon torzultak, kevésbé hatékonyan volt meghatározható az eljárással a retinális pigment epithel réteg (RPE), ezért kidolgoztunk egy gráf vágáson alapuló megközelítést. Az eljárásokat kétféleképpen értékeltük ki. Az algoritmusok által elért eredményeket összehasonlítottuk a szakértők annotációjával és más, újraimplementált szakirodalmi módszerekkel. Kiszámítottuk az átlagos és maximális hibát, valamint a szórását is az egyes szekvenciákra. A javasolt módszerek átlagos hibája minden esetben kevesebb, mint 3 pixel, ami azt mutatja, hogy a megközelítéseink stabilabbak, mint az újraimplementált algoritmusok.

Az OCT vizsgálat során készült felvételek különféle torzulásokkal lehetnek terhelvek. A keresztmetszeti képeken a zaj mellett megjelennek a vérerek árnyékából származó artifaktumok. Bemutattunk egy módszert, mely a pásztázó lézer oftalmoszkópiával (SLO) készült felvételen detektálja a vérereket, melyek pontos pozíciója a B-scan képen megadható. Ezáltal ismertté válnak azok a területek, melyek fals információval bírnak, mert a kép ezen a területen torzult. Az orvosi diagnózis szempontjából fontos a betegség miatt bekövetkező rétegfelgyűrődés mértéke is, ezért javasoltunk egy módszert ennek meghatározására.

A szakirodalomban a biomarkerek közül a legtöbb publikáció a szubretinális (SRF) és intraretinális folyadék (IRF) detektálásával foglalkozik. Ismertetünk egy eljárást, mely a határoló rétegek esetében használt vertikális profilok további elemzésén alapul. Mivel az SRF és IRF hiperreflektivitása is hasonló az üvegtestéhez, így kidolgoztunk egy módszert ezek egyidejű detektálására, valamint szétválasztására. A kutatás ezen fázisában a szemorvosok által annotált képek nem voltak elérhetőek, ezért csak vizuális összehasonlítást végeztünk más, általunk újrainplementált szakirodalmi eljárásokkal.

A betegség esetében bekövetkezhet pigment epithel leválás (PED) szubretinális hiperreflektív anyag (SHRM) kíséretében, melyek detektálására és megkülönböztetésére prezentáltunk egy módszert. A módszereinket az annotált adatbázison értékeltük ki a Dice együttható és Sensitivity mutatók segítségével. Sok esetben az SHRM nem egyértelműen kimutatható, de mindkét biomarker esetében az átlagos Dice értékek meghaladják 0.75-öt, és a Sensitivity értékek 0.93 és 0.77 a PED és az SHRM esetén.

Csak néhány publikáció foglalkozik a szubretinális anyag lokalizációjával, de egy sem a külső retinális tubuláció szegmentálásával, ám az orvosi szakterületen megjelenő publikációk azt mutatják, hogy ezeknek a tubulációknak a számossága az orvosok számára a betegség egy fontos leírója. A problémát kétféle módon vizsgáltuk: egy klasszikus képfeldolgozási eljárás alkalmazásával és egy mély neurális hálózaton alapuló megközelítéssel. A széleskörű kiértékelés érdekében különféle mutatókat számítottunk ki mind az automatikus, mind a kézi szegmentáláshoz. Megmértük a globális és az objektum szintű Dice együtthatót és a Recall-t. Módszerünk objektum szintű Recall értéke 0.847 és Dice együtthatója 0.583, amely összehasonlítható az emberi szakemberek eredményeivel.

Zebrahal embriók morfológiai elváltozásainak detektálása

A 3. fejezetben sugárzás által károsodott zebrahal embriók morfológiai elváltozásainak meghatározására kidolgozott automatizált eljárást mutattunk be. Sugárkezelés után dóziszfüggő módon morfológiai elváltozások észlelhetők az embriókban. A gerinces zebrahal-modell alkalmazásának előnye, hogy nagy számú embrió biológiai változásainak mérésére alkalmas, amely lehetővé teszi a megbízható statisztikai elem-

zést. A kifinomultabb morfológiai mérések – az egyszerű túlélés észlelésén túl – rendkívül idő- és munkaigényesek. Bemutattunk egy keretrendszert az emberi munkaigény csökkentésére a gyorsabb kísérletezés elérése érdekében. Feltételezzük, hogy a kísérletek során ugyanazt a protokollt követték, beleértve az embriók elhelyezését és kezelését, valamint a mikroszkópos képek készítését. Az automatizált elemzés több lépésből áll. Elvárható, hogy a halakat tartalmazó lyuk kör alakú falának legalább egy része megfigyelhető a képeken, ezért javasoltunk egy módszert a lyuk régiójának körülhatárolására. Bemutattunk egy eljárást a halak lokalizálására, valamint a fej és a farok részeinek megkülönböztetésére. Meghatároztuk a korábban már szegmentált hal végpontjait, a teljes testhosszát, valamint a szem helyét és átmérőjét is. A 393 kép kiértékelése során nem találtunk szignifikáns különbséget a kézi mérés eredményei és az automatikus módszer között.

Vizuális kódok lokalizálása

A 4. fejezetben különböző vizuális kódok – beleértve a hagyományos 1D, valamint 2D QR és 2D halmozott kódokat – detektálására kidolgozott módszereink tárgyalására kerül sor. Bemutattunk két morfológiai műveleten alapuló eljárást a klasszikus 1D-s kódok hatékony lokalizálására. Az algoritmusok eredményeit összehasonlítottuk más szakirodalmi módszerekkel. Az eredményekből azt a következtetést vonhatjuk le, hogy a javasolt eljárások a második legjobb futási idővel rendelkeznek, és sok esetben a legnagyobb pontosságot mutatják. A módszer gyengesége akkor jelenik meg, amikor a kép nagyon zajos, vagy ha vannak olyan képterületek, amelyek hasonlóak a vonalkódhoz.

Az 1D vonalkódok párhuzamos vonalakkal állnak, melyek különböző távolságra vannak egymástól. Ezt a szerkezeti jellemzőt használjuk a mintaillesztésen alapuló detektálási módszerünkben. Megállapítottuk, hogy a halmozott 2D kódok elegendő hasonlóságot mutatnak a mintákban a hagyományosakkal, ezért néhány, például a PDF417 és a Codablock szintén lokalizálható a javasolt módszerrel. Másik minta használatával a módszer adaptálható a QR kódok helyének meghatározására is. A megközelítéseinket összehasonlítottuk a szakirodalom más algoritmusaihoz a Muenster adatbázison. Megállapítottuk, hogy az eljárásaink gyorsak és nem érzékenyek a forgatásra az 1D-s kódok esetében. A QR mintaillesztéses megközelítés hatékonyságát kiértékeltek a Dubská adathalmazon más szakirodalmi módszerekkel együttesen, és a javasolt algoritmusunk teljesített a legjobban.

A szokásosan használt kódok geometriailag jól definiált minták szerint tervezettek. Előállíthatók azonban olyan mintázatok, melyek természetüknél fogva egyediek és így eredetiség vagy hitelesség ellenőrzésére is alkalmazhatók. Bemutattunk egy olyan lehetséges hibrid vizuális kód konstrukciót, melyben mesterséges és természetes mintázatok együttes alkalmazásával elérhető az azonosítás és a hitelesítés is, így a fela-

dat szétbontható három fő lépésre: képkalkotás, képfeldolgozás és azonosítás. Javasoltunk egy módszert a QR kód helyzetének és orientációjának meghatározására, a glitterzónában található szemcsék detektálására. Egy egyszerű párosítási stratégiát is bemutattunk a hitelesítéshez. A képanalízis mobil eszközökön is gyorsnak bizonyult.

A disszertáció tézisei

Az **első téziscsoport** eredményei az időskori makuladegeneráció (AMD) különböző biomarkereinek optikai koherencia tomográf (OCT) képek elemzésén alapuló vizsgálatához kapcsolódnak, beleértve új módszerek kidolgozását és orvosi referencia szegmentálásokkal való összehasonlítását. (Részletes tárgyalás a 2. fejezetben.)

- I/1. Javasoltam két új megközelítést a retina belső határhártyája (ILM) és a retina pigment epithel (RPE) réteg lokalizálására.
- I/2. Kidolgoztam egy eljárást a rétegfelgyűrődés pozíciójának és mértékének meghatározására, mely fontos jellemző az AMD megítélése szempontjából.
- I/3. Javasoltam két új algoritmust a szubretinális (SRF) és intraretinális (IRF) folyadék detektálására.
- I/4. Kidolgoztam egy eljárást a pigment epithel leválás (PED) és a szubretinális hiperreflektív anyag (SHRM) régiók meghatározására.
- I/5. Javasoltam egy új, klasszikus képfeldolgozó műveleteket alkalmazó módszert, valamint egy konvolúciós neurális hálózatot a külső retinális tubuláció (ORT) lokalizálására.

A **második tézis** eredménye a radiobiológiai kutatásokban használt zebrahal embriók sugárzás hatására bekövetkező morfológiai változásainak automatikus detektálása mikroszkópos képeken. (Részletes tárgyalás a 3. fejezetben.)

- II. Kidolgoztam egy új eljárást a lyuk zónájának behatárolására és a zebrahalak különböző alaktani jellemzőinek kinyerésére: meghatároztam az embriók végpontjait és hosszát, az embrió szemének pozícióját és átmérőjét, a hal irányultságát.

A **harmadik téziscsoport** eredményei vizuális kódok lokalizációjához kapcsolódnak digitális képeken. (Részletes tárgyalás a 4. fejezetben.)

- III/1. Javasoltam két új, morfológiai műveleteken alapuló algoritmust az 1D-s klaszterikus és 2D-s halmazos vonalkódok lokalizálására.

-
- III/2. Kidolgoztam egy-egy új, mintaillesztésen alapuló eljárást a hagyományos kódok és a QR kódok helyének meghatározására.
 - III/3. Javasoltam egy algoritmust egy hibrid vizuális kód (NFI) detektálására és felismerésére.
 - III/4. Elvégeztem 12 kódlokálizációs módszer kiterjedt összehasonlító elemzését 1D és 2D kódokat ábrázoló, több tízezer valós és szintetikus képet tartalmazó publikusan elérhető illetve saját összeállítású adatbázison.

Publications

Journal publications

M. Katona and L. G. Nyúl. Fast Recognition of Natural Feature Identifiers by a Mobile Phone. *Acta Cybernetica*, 22:101–116, 2015.

M. Katona and L. G. Nyúl. An Approach to the Quantitative Assessment of Retinal Layer Distortions and Subretinal Fluid in SD-OCT Images. *Acta Cybernetica*, 23:615–628, 2017.

M. Katona, P. Bodnár, and L. G. Nyúl. Distance Transform and Template Matching Based Methods for Localization of Barcodes and QR Codes. *Computer Science and Information Systems*. doi: 10.2298/CSIS181011020K.

Full papers in conference proceedings

M. Katona and L. G. Nyúl. A Novel Method for Accurate and Efficient Barcode Detection with Morphological Operations. In *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, pages 307–314, 2012.

M. Katona and L. G. Nyúl. Efficient 1D and 2D Barcode Detection Using Mathematical Morphology. In *Mathematical Morphology and Its Applications to Signal and Image Processing*, pages 464–475. Springer Berlin Heidelberg, 2013.

M. Katona, A. Kovács, R. Dégi, and L. G. Nyúl. Automatic Detection of Subretinal Fluid and Cyst in Retinal Images. In *Image Analysis and Processing – ICIAP 2017*, pages 606–616, 2017.

M. Katona, A. Kovács, L. Varga, T. Grósz, J. Dombi, R. Dégi, and L. G. Nyúl. Automatic detection and characterization of biomarkers in oct images. In Aurélio Campilho, Fakhri Karray, and Bart ter Haar Romeny, editors, *Image Analysis and Recognition – ICIAR 2018*, pages 706–714, Cham, 2018. Springer International Publishing. doi: 10.1007/978-3-319-93000-8_80.

M. Katona, A. Kovács, R. Dégi, and L. G. Nyúl. Segmentation of Subretinal Hyperreflective Material and Pigment Epithelial Detachment Using Kernel Graph Cut.

In *Progress in Computer Recognition Systems*, pages 98–105. Springer International Publishing, 2020.

M. Katona, T. Tóké, E. R. Szabó, Sz. Brunner, I. Z. Szabó, R. Polanek, K. Hideghéty, and L. G. Nyúl. Automatic Segmentation and Quantitative Analysis of Irradiated Zebrafish Embryos. In *Computational Modeling of Objects Presented in Images. Fundamentals, Methods, and Applications*, pages 95–107, 2019.

I. Megyeri, **M. Katona**, and L. G. Nyúl. A Novel Approach to Detect Outer Retinal Tubulation using U-Net in SD-OCT images. In *The 15th International Conference on Signal Image Technology & Internet Based Systems*, 2019. Submitted for publication.

Other publications

M. Katona and L. G. Nyúl. Improved 1D and 2D barcode detection with morphological operations. In *A Képfeldolgozók és Alakfelismerők Társaságának 9. országos konferenciája - KÉPAF 2013*, pages 309–324, 2013.

M. Katona and L. G. Nyúl. Fast recognition of natural feature identifiers by a mobile phone. In *The 9th Conference of PhD Students in Computer Science (CSCS 2014) : Volume of Extended Abstracts*, page 26, 2014.

Katona M. and Nyúl L. G. Vonalkódok és természetes vizuális azonosítók felismerése valós időben. In *A Képfeldolgozók és Alakfelismerők Társaságának 10. országos konferenciája - KÉPAF 2015*, pages 562–577, 2015.

R. Dégi, A. Kovács, **M. Katona**, L. G. Nyúl, L. G. Varga, J. Dombi, and T. Grósz. Quantitative assessment of the manifestation of age-related macular degeneration via automated image analysis of spectral-domain optical coherence tomography. In *European Society of Retina Specialists 15th EURETINA Congress*, page 1 p., 2015.

Varga L., **Katona M.**, Grósz T., Dombi J., Kovács A., Dégi R., and Nyúl L. G. Időskori makula degeneráció kvantitatív jellemzése SD-OCT képek automatikus elemzésével. In *Új alapokon az egészségügyi informatika : A XXVIII. Neumann Kollokvium konferenciakiadványa*, pages 43–48, 2015.

Katona M. and Nyúl L. G. Szubretinális folyadékter és ciszta automatikus detektálása retinafelvételeken. In *A Képfeldolgozók és Alakfelismerők Társaságának 11. országos konferenciája - KÉPAF 2015*, page Paper 51. 11 p., 2017.

M. Katona and L. G. Nyúl. Quantitative Assessment of Retinal Layer Distortions and Subretinal Fluid in SD-OCT Images. In *The 10th Conference of PhD Students in Computer Science (CSCS 2016) : Volume of Extended Abstracts*, page 37, 2016.

Kovács A., Dégi R., **Katona M.**, Nyúl L. G., Varga L., Dombi J., and Grósz T. Biomarkerek automatikus karakterizálása OCT felvételek időbeli elemzésére. In *Magyar Sze-morvostársaság Retina Szekciójának Kongresszusa*, page 1 p., 2017.

Katona M., T. Tőkés, E. R. Szabó, Sz. Brunner, R. Polanek, Hideghéty K., and Nyúl L. G. Sugárkezelésen átesett zebrahal embriók automatikus morfológiai elemzése. In *A Képfeldolgozók és Alakfelismerők Társaságának 12. országos konferenciája - KÉPAF 2019*, page Paper 21. 12 p., 2019.