

Raneem Ghassan Ismail

Detection, localization, and segmentation of colorectal  
polyps in colonoscopy images by computational  
intelligence methods

Doctoral Dissertation

Supervisors:

Professor Dr. Szilvia Nagy

Department of Telecommunications

Dr. Brigita Sziová

Department of Informatics

Faculty of Mechanical Engineering, Informatics and Electrical Engineering

Széchenyi István University

Multidisciplinary Doctoral School of Engineering Sciences

Széchenyi István University

Győr, Hungary

2024

## **Acknowledgments**

I would like to express my sincere gratitude to everyone who has contributed to the completion of this dissertation.

First and foremost, I am grateful to my supervisor, Prof. Szilvia Nagy, for her invaluable guidance, support, and encouragement throughout the course of my research. Her extensive knowledge, expertise, and critical feedback have been very useful in shaping my ideas and improving the quality of my work. I would also like to mention my gratitude to my second supervisor Brigita Sziová for her help in the undersanding of intelligent decision support systems.

I would also like to thank all my professors and colleagues in the Multidisciplinary Doctoral School of Engineering Sciences (MMTDI) at Széchenyi István University- Győr, for their help and cooperation during my entire research journey.

My deepest gratitude goes to my loving husband "Rashad", and my great family for their unwavering love, support, and encouragement during the challenging times of my research. Their constant belief in me and in my abilities has been an essential source of my motivation and inspiration.

Finally, I hope that obtaining my PhD will be the starting point for a successful and fruitful practical life, you were all a solid foundation for it.

## Summary

Colorectal polyps and the cancerous lesions that can potentially develop from them are very important public health problems today, so their screening is extremely important. Several challenges have appeared in the scientific community of medical image processing, which focus on detecting, localizing, and segmenting these abnormalities. In my work, I aimed to achieve better, more flexible, and more plausible processing of colonoscopy images. I utilized several publicly available databases for conducting tests during my work.

I proposed improvements for a polyp detection procedure that classifies colonoscopy image segments using fuzzy rule interpolation to determine their polyp content. First, I presented alternative solutions for determining the rules based on the statistical properties of the training set with rules that fit the median or the histogram of its distribution instead of its mean, and then I examined the effects of dividing the output classes according to polyp content.

I also studied possibilities for improving a procedure using Hough transformations for localizing roundish polyps. I suggested four modifications, aiming to reduce computational requirements and increase tolerance to deviations from the circular shape. To reduce computational demands, I investigated the impact of the edge detection method used as the first step of the Hough transformation. I also analyzed the gradient distribution of polyp edges and other edges in the images to explore thresholding possibilities. Additionally, I introduced a weighting scheme into the Hough transformation that takes into account the gradients of edge pixels. To increase shape deviation tolerance, I proposed the application of fuzzy Hough transform.

For the precise determination of localized polyp outlines, I tested two variational segmentation methods. I investigated which pre-filtering methods best facilitate segmentation, how the initial mask's shape and size influence results, and which type of method, along with what parameter settings, yield the best outcomes.

## Összefoglalás

A kolorektális polipok, illetve az azokból potenciálisan kialakuló rákos elváltozások napjaink igen fontos népegészségügyi problémái, ezért szűrésük kiemelkedően fontos. Több olyan kihívás jelent meg az orvosi képfeldolgozással foglalkozó tudományos közösségben, mely ezen elváltozások lokalizálására, detektálására, szegmentálására irányul. Munkám során én is a kolonoszkópiai képek jobb, rugalmasabb, plauzibilisebb feldolgozását tűztem ki célul. Több nyilvános adatbázis érhető el, ezeket használtam a tesztek elvégzésére.

Javítási lehetőségeket javasoltam egy olyan polip detektáló eljáráshoz, mely kolonoszkópiai felvételeket fuzzy szabályinterpolációs módszerrel osztályoz azok poliptartalmának megállapítására. Először a szabályoknak a tanítóhalmaz statisztikai tulajdonságai alapján való megállapítására mutattam be alternatív megoldásokat a mintaközéppértéke helyett annak mediánjához, illetve eloszlásának hisztogramjához jobban illeszkedő szabályokkal, majd a kimeneti osztályok poliptartalom szerinti felosztásának hatásait vizsgáltam meg.

Egy gömbölyded polipok lokalizálására alkalmas, Hough-transzformációkat használó eljárás javításának lehetőségeit is megvizsgáltam. Négy módosítást javasoltam egyrészt a számítási igény csökkentésére, másrészt a köralaktól való eltéréssel szembeni tolerancia növelésére. A számítási igény csökkentésére megvizsgáltam, milyen hatással van a Hough-transzformáció első lépéseként végrehajtott éldetektálási eljárás típusa, másrészt megvizsgáltam, milyen gradiens eloszlással rendelkeznek a polipok élei, illetve a képeken előforduló más élek, és ez alapján az élek küszöbölési lehetőségeit vizsgáltam, harmadrészt egy olyan súlyozást vezettem be a Hough-transzformációba, mely az élek pixeleihez tartozó gradienseket is figyelembe veszi.

Variációs szegmentálási módszereket vizsgáltam már lokalizált polipok körvonalainak pontosabb meghatározásának céljából. Megvizsgáltam, milyen előszűrési módok segítik elő legjobban a szegmentálást, azt, hogy a kiindulási maszk alakja, mérete hogyan befolyásolja az eredményeket, illetve hogy milyen típusú módszer, milyen paraméterbeállításokkal adja a legjobb eredményeket.



**Contents:**

1. Introduction .....	1
1.1 Background .....	3
1.2 Research methodology .....	6
2. Colorectal polyps' detection on colonoscopy images using fuzzy inference.....	9
2.1 Basic concepts of fuzzy classification.....	9
2.1.1 Introduction to fuzzy sets .....	9
2.1.2 Fuzzy inference .....	9
2.1.3 Fuzzy rule interpolation .....	11
2.1.4 Evaluation metrics.....	12
2.2 Application of fuzzy rule interpolation in colorectal polyp detection .....	13
2.2.1 Introducing the uniform classification scheme .....	13
2.2.2 Derivation of the antecedents .....	15
2.2.3 Median-centered and histogram-fitted rulebase generation methods.....	17
2.2.4 Experimental study and analysis .....	22
2.3 The influence of introducing more consequent classes on the classification performance.....	24
2.3.1 Experimental results .....	24
2.4 Thesis statements.....	29
2.4.1 Thesis 1 .....	29
2.4.2 Thesis 2 .....	30
3. Colorectal polyps' localization in colonoscopy images using fuzzy circular Hough transform .....	31
3.1 Motivation .....	31
3.2 Theoretical and mathematical background of the implemented techniques ....	35
3.2.1 Classical and fuzzy Hough transforms.....	35
3.2.2 Gradient filtering .....	40
3.2.3 Roberts, Prewitt, and Sobel edge detection algorithms.....	43
3.2.4 Canny edge detection algorithm.....	43
3.3 Practical application of the proposed method .....	44
3.3.1 Edge detection methods—application, evaluation, and selection .....	44
3.3.2 Gradient-based thresholding for Prewitt edge detection results.....	53

3.3.3	Gradient-weighted voting approach for classical and fuzzy Hough transforms.....	57
3.4	Results and discussion.....	60
3.4.1	Number of circles found by the algorithm .....	60
3.4.2	Roundness metrics evaluation .....	68
3.4.3	Time evaluation.....	70
3.5	Thesis statements.....	72
3.5.1	Thesis 3 .....	72
3.5.2	Thesis 4 .....	72
4.	Active contour methods in colonoscopy image segmentation .....	74
4.1	Introduction .....	74
4.2	Variational methods for image segmentation – mathematical background .....	76
4.2.1	Snakes: active contour models .....	76
4.2.2	Mumford-Shah approach.....	78
4.3	Level Set Methods for image segmentation– mathematical background .....	79
4.3.1	Geodesic active contours.....	80
4.3.2	Chan-Vese active contours.....	81
4.4	Practical implementation and results.....	83
4.4.1	Evaluation criteria .....	83
4.4.2	Reference case.....	83
4.4.3	Pre-processing phase .....	85
4.4.4	Post pre-processing case.....	87
4.4.5	Circular initial mask case .....	90
4.4.6	Modified rectangular initial mask and tuned parameters case .....	96
4.4.7	Special cases.....	101
4.5	Thesis statements.....	103
4.5.1	Thesis 5 .....	103
5.	Outlook and future work.....	104
6.	Publications .....	105
	References .....	106
	List of figures .....	117
	List of Tables.....	120
	Appendix A .....	122

Appendix B ..... 142  
Appendix C ..... 162  
Appendix D ..... 165

## 1. Introduction

Colorectal polyps are abnormal growths that develop on the inner lining of the colon or rectum. While most polyps are benign and do not cause any symptoms, some polyps can potentially become cancerous over time. Colorectal cancer (CRC) is the third lethal cancer type and the fourth most common cause of cancer mortality worldwide, (depending on the year and the country where the statistics took place) (Ismail & Nagy, 2021). The survival rate of CRC depends on the stage it is detected in, going from rates higher than 95% in the case of early stages to rates lower than 35% in latter ones (Bernal J. , et al., 2015); hence, the prevention of CRC by finding and removing the so-called preneoplastic lesions (early-stage precancerous polyps or colorectal adenomas) is one of the most important worldwide public health studies and has a high priority in computer-aided diagnostic (CAD) systems research (Bernal, Sanchez, & Vilariño, 2012).

Clinical guidelines based on medical practice as well as scientific evidence recommend that individuals who are over the age of fifty must undergo a regular examination. The conventional, classical, colonoscopy is the most precise visual inspection procedure for early colorectal cancer detection and prevention. A colonoscope, which is an endoscope equipped with a flexible tube, camera, light source, and other tools, can navigate within the bowel, and perform various tasks like inflating the bowel, spraying liquids, taking a biopsy, and removing lesions (Alam & Fattah, 2023), (Krenzer, et al., 2023).

Several reputed studies with sufficiently large number of cases demonstrate that colonoscopy screening reduces the cancer's incidence by 40–90%. During the colonoscopy procedure, its flexible tube with a camera is inserted into the colon and rectum to examine the inner lining, and if polyps are found, they can be removed or biopsied for further analysis, by this way, colonoscopy allows both the identification and the removal of such polyps, which are at their first step in the many years long process of developing “adenoma” to “high grade dysplasia” to “carcinoma,” i.e., the development of colorectal cancer (Bernal J. , et al., 2015), (Bernal, Sanchez, & Vilariño, 2012).

However, colonoscopy has the following drawbacks. Before colonoscopy the patient must perform a thorough bowel cleansing, and the procedure itself is also rather inconvenient, which makes people less willing to sign up for colonoscopy screening. The procedure requires an expensive device and expert medical staff to carry out the examination. It also has a considerable risk of damaging the bowel wall; moreover, it is often carried out in anesthesia,

which has its own risk. Besides the costs and risks of the medical process, the most significant drawback is the reported high miss-rate of polyps, which can be as high as 22%. While colonoscopy is effective in detecting polyps larger than 10 mm, its accuracy decreases significantly for smaller or flat polyps (Bernal J. , et al., 2015), (Bernal, Sanchez, & Vilariño, 2012).

Colorectal polyps can be categorized based on their location to the layers of the bowel wall, as mucosa, muscular mucosa, and submucosa (Sziová, Ismail, Lilik, Kóczy, & Nagy, 2020). Moreover, as polyp morphology might be useful for predicting the presence of invasive growth, a group of endoscopists, pathologists, and surgeons from Western and Japanese countries established an endoscopic classification system for polyp morphology in superficial neoplastic lesions in the colon, esophagus, and stomach. This classification, known as the Paris classification, divides polyps into several groups based on their endoscopic appearance, such as: pedunculated, sessile, slightly elevated, flat, slightly depressed and excavated (Vleugels, Hazewinkel , & Dekker , 2017).

The Kudo classification was created in 1994 to describe the micro-architecture of epithelial pits, also known as pit patterns, seen during chromoendoscopy under high magnification (Kudo , Hirota, Nakajima, Hosobe, & Kusaka, 1994). Its purpose is to differentiate between non-neoplastic, adenomatous, and cancerous lesions. If the lesion surface has roundish pits (class A1) or stellar, asteroid, or papillary shaped pits (class A2), a light-pink color, and a regular pattern, then the polyp is not likely to develop into a malignancy. However, if the lesion has tubular shapes (class B3L) or extremely small roundish pits (class B3S), it has the potential to develop into a carcinoma and may show some discoloration under narrow light band imaging. Lesions with dark, dendritic, gyrus- or sulcus-like pit structures (class B4) are already in a cancerous state, as are irregularly shaped patterns with colors tending towards blackish and/or whitish (class B5) (Vleugels, Hazewinkel , & Dekker , 2017), (Nagy, Lilik, & Kóczy, 2017).

The size of the colonic lesions is another important factor indicating the risk of cancer. Lesions measuring 1 to 5 mm (diminutive) have a very low risk of containing invasive growth, ranging from 0 to 0.1%. Small lesions measuring 6 to 9 mm have a risk ranging between 0 and 0.4%. For lesions larger than 10 mm, the risk of cancer gradually increases. Lesions measuring between 10-20 mm have a risk of 2.4%, while polyps greater than 20 mm in size have the highest risk of 19.4% (Vleugels, Hazewinkel , & Dekker , 2017).

Even for experienced gastroenterologists, detecting polyps with the naked eye can be a challenging task, as some polyps may be very small in size or hidden in folds of the colon lining, in addition to various reasons like the inadequate intestinal preparation and doctor's visual exhaustion (Alam & Fattah, 2023), (Krenzer, et al., 2023), (Yue, et al., 2022).

Developing a computer-aided diagnosis system using computational intelligence methods for colorectal polyps' medical tasks, (i.e., colorectal polyps' detection, localization, and segmentation) is currently one of the most demanding needs in the healthcare sector. Such a system would simplify the diagnostic process and facilitate the assessment of the examined cases severity. Although none of these intelligence methods are intended to replace human diagnosis, at least, not in the near future, but they definitely help the often-overloaded medical staff to collect and organize the available data for decision support. They can draw their attention to certain phenomena, identify and highlight potential polyps, and provide accurate measurements of polyps' size, location, and the degree of their severity (Alam & Fattah, 2023), (Krenzer, et al., 2023), (Yue, et al., 2022).

By employing robust computer-aided diagnosis techniques, gastroenterologists can significantly enhance the accuracy and efficiency of colorectal cancer screening. This, in turn, enables early detection and timely treatment of polyps, thereby leading to a reduction in the incidence and mortality rates associated with colorectal cancer.

### ***1.1 Background***

By utilizing advanced image processing and computer vision techniques, many algorithms have been developed by researchers to automatically assist gastroenterologists in detecting, localizing, and segmenting polyps during colonoscopy. These algorithms vary from polyp shape or texture-based handcrafted methods to fully automated machine learning methods based on highly efficient neural networks, as well as the hybrid methods which combine both. Some of them are rather successful, but the possibility to apply them in real-life applications is still not near.

Mostly, polyp characterization methods are based on the calculation of some feature descriptors over a tile of the image or the complete image. These descriptors support the decision of an intelligent system whether there is a polyp in the image (segment) or not, and often if there is a polyp, the location within the image (segment) is also determined. Two main groups of descriptors exist, the shape-based properties, and the texture-based ones.

The texture analysis is often used in the case of differential diagnosis as based on its surface texture the nature of the polyp can be determined (Kudo , Hirota, Nakajima, Hosobe, & Kusaka, 1994), (Horváth, Spindler, Szalai, & Rácz, 2016), (Rácz, et al., 2015), (Georgieva, Nagy, Kamenova, & Horváth, 2015). However, magnified endoscopy images are needed to perform this type of analysis, thus it is rarely used for the detecting the polyp; it is used after the polyp is already found. Polyps tend to have different patterns than the bowel walls which contain visible blood vessels too. Some works are based on the use of wavelet descriptors to extract this information (Karkanis, Iakodivis, Karras, & Maroulis, 2001), paying attention to the detail and approximation coefficients of the wavelet transform. Other alternatives include the use of local binary patterns (Häfner, et al., 2010) or co-occurrence matrices (Tan, et al., 2020).

On the other hand, shape-based approaches (Rácz, Jánoki, & Saleh, 2010), (Bernal, Sanchez, & Vilariño, 2012), (Bernal J. , et al., 2015), (Yuji, et al., 2015), (Bernal J. ,et al., 2017) aim to search in colonoscopy frames for those specific shapes that polyps commonly have, both in the intensity distribution and in the boundary shape. The first subgroup of such methods (i.e., the intensity distribution-based one) clusters a series of methods mostly consisting of low-level image processing operations such as gradient filters, valley information, or edge detectors. Whereas the second subgroup (i.e., the boundary shape-based one) assumes that polyps tend to have elliptical shape, or often close to the circular shape.

Various machine learning and deep learning architectures have been currently developed as prominent solutions to automatize polyp detection and localization tasks and enhance their accuracy (Ahmad, Brandao, Sami, & et al., 2019), (Sornapudi, Meng, & Yi, 2019), (Wittenberg, Zobel, Rathke, & Mühldorfer, 2019). A method for real-time detection, classification, and localization of gastrointestinal tract disorders from colonoscopy images was presented in (Aliyi, Dese, & Raj, 2023). Both online available (hyper-Kvasir dataset) and private locally collected samples were utilized. The method was developed using the pre-trained transfer learning SSD, YOLOv4, and YOLOv5 object detection models, with minimal fine-tuning of the hyperparameters, and their final performances were compared. The utilization of the YOLOv5 object detection algorithm and the artificial bee colony (ABC) optimization algorithm was proposed in (Karaman, et al., 2023). The YOLOv5 algorithm was employed for polyp detection, while the ABC algorithm was used to enhance the performance of the model by finding the optimal activation functions and hyperparameters for the YOLOv5 algorithm. The proposed

method was executed on the SUN and PICCOLO datasets and achieved good performance in real-time polyp detection.

Different convolutional neural network (CNN) architectures were proposed and evaluated using different currently available public colonoscopy databases and a wide variety of convenient metrics in (Sánchez-Peralta, Bote-Curiel, Picón, Sánchez-Margallo, & Pagador, 2020). The effectiveness of the proposed encoder-decoder ColonSegNet architecture, in conjunction with several deep learning methods was evaluated using the Kvasir-SEG database in (Jha, et al., 2021). Researchers in (Mo, Tao, Wang, & Wang, 2018) employed the faster region-based convolutional neural network (Faster RCNN) method with VGG16 as the backbone for polyp detection. An automated system based on a fusion of color wavelet features and convolutional neural network features along with the use of a support vector machine for polyp detection and localization was proposed in (Billah, Waheed, & Rahman, 2017). A combination of the existing deep neural network, DeepLab v3, with Long Short-Term Memory networks (LSTMs) was also presented in (Xiao, Chang, & Liu, 2018) for polyps' detection in colonoscopy images.

One of the fundamental challenges in medical image processing is the limited number of training and testing database samples. Unlike the usual image classification tasks, the medical images cannot be processed by unskilled personnel, the preparation of the labeled ground truth masks need medical experts, that makes the creation of the medical databases very costly and time-consuming process. In addition legal, social, and ethical challenges come with managing personal data corresponding to the medical images. The above-mentioned facts clarify why the large training set based learning algorithms are not the only possibilities in medical image processing, but other approaches, such as expert knowledge based fuzzy inference are also preferable.

In the past few years, a research team has been formed at Széchenyi István University that focuses on detecting and classifying colorectal polyps within colonoscopy image segments using a fuzzy inference-based strategy. These image segments either contain a complete polyp or at least a portion of it or they are completely empty regarding the polyps (Nagy, Lilik, & Kóczy, 2017), (Nagy, Sziova, & Koczy, 2018). The purpose of the fuzzy inference method developed in our group is to determine, whether an image segment that is not included in the training set contains areas belonging to polyp or not.

The proposed strategy was based on comprehensive consultations with expert gastroenterologists and on the utilization of plausible statistical parameters and entropies of



image segments as antecedents, and finally, a fuzzy decision value was assigned to each segment. Subsequently, additional preprocessing methods and several antecedents were incorporated into the original strategy, and their effect was also explored (Nagy, Sziova, & Koczy, 2017), (Sziová, Nagy, & Kóczy, 2021).

Even though fuzzy inference is not used by other groups in this field of medical image analysis, the previous results are somewhat promising, but leave place for further development.

For the research presented in my dissertation, the continuous improvement and exploring of the fuzzy inference-based method was essential. My contribution was introducing new rulebase generation methods and more fitting consequents for colorectal polyp detection purposes using the stabilized Kóczy-Hirota rulebase interpolation method.

As the classical circular Hough transform was already found to be successful in colorectal polyp localization task, to deal with the not perfectly circular shapes of the polyps and to improve the Hough transform's tolerance for shape deviation, I studied the applicability of fuzzy Hough transform, together with special preprocessing methods and a gradient-based weighting approach in its voting methodology.

As a last step, I also researched the colorectal polyp segmentation possibilities. One of the most effective, non artificial intelligent based segmentation methods is the active contours method. I tested the efficiency of two active contours methods and the necessary preprocessing steps to optimize their performance.

## **1.2 Research methodology**

A comprehensive literature review has been carried out to establish a robust background on the research topic and to achieve the above mentioned goals, with an intense focus on the international contributions and recently developed algorithms together with the evaluation employed metrics in this interesting field.

In this dissertation, three publicly available still colonoscopy image databases CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and Etis-Larib (Silva, Histace, Romain, Dray, & Granado, 2014) were used. Each of these databases contains a specific number of images, namely 612, 380, and 195, respectively. The colonoscopy images have different sizes, the dimensions of the images are  $384 \times 288$ ,  $574 \times 500$ , and  $1225 \times 966$  pixels, and they have a resolution of 96, 72, and 72 dpi, respectively. Furthermore, the colour depth of the RGB color pictures is 24 bit per pixel. In addition to the frames, a ground truth

mask corresponding to the region covered by the polyp/ polyps in the image is also included in each database.

The research methodology can be well organized by briefly listing the practical steps that were implemented and investigated in this dissertation as follows.

- ❖ Utilizing computational intelligence methods through simple measured data that are easy to interpret by medical experts is very advantageous for computer-aided diagnosis systems. Based on the concept of fuzzy reasoning, if the rulebase is derived from plausible statistical parameters of the image being analyzed, it can be easily comprehended and accepted by society. Thus, a further development for a fuzzy inference-based strategy was carried out. Simple statistical parameters and entropies of image segments were used as rulebase antecedents.
  - Investigating the impact of determining the rulebase parameters on the efficiency of the detecting of the polyp segment using stabilized Kóczy-Hirota fuzzy rule interpolation. The performance of the proposed method was evaluated using the true positive rate (TPR), the false positive rate (FPR), the true negative rate (TNR) and the false negative rate (FNR) metrics.
  - Introducing more consequent classes in the colorectal polyp's detection approach to better align with real-world scenarios, and examining the effect of the refinement of the consequent categories in stabilized Kóczy-Hirota fuzzy rule interpolation. The number of the consequents was extended from two classes, (i.e., “with-polyp”, or “without-polyp”), to a more refined set distribution based on the polyp content of the image segment. Other metrics, namely accuracy (ACC), Matthews correlation coefficient (MCC) and  $F_1$  score were introduced for a more detailed evaluation.
- ❖ Applying classical circular Hough transform was proven to be effective for some types of colorectal polyps. However, the polyps are very rarely perfectly circular, so more tolerance to shape uncertainty is needed. That was the reason why the switching from the classical circular Hough transform to the fuzzy circular Hough transform for the colorectal polyp's localization purposes was proposed.
  - Depending on statistical data from the three colonoscopy databases, the performance of four edge detection algorithms (Canny, Prewitt, Roberts, and Sobel) was compared, and the most ideal one that gave the most polyp contour related and least unnecessary edge points, i.e., the lowest number of points to be transformed, was

selected. Two metrics based on normalized gradients of contour and non-contour edges were developed to establish this selection.

- According to a statistical study that has been performed during this research, in the colonoscopy images the polyp contours usually belong to gradient domain of neither too large, nor too small gradients, though they can also have stronger or weaker segments. To prioritize the gradient domain typical for the polyps, a relative gradient-based thresholding as well as a gradient-weighted voting approach was introduced.
- Characterizing the roundness of the objects to be detected was applied for evaluating the improvement of the shape deviation tolerance of the classical and fuzzy Hough transforms using the maximum radial displacement and the average radius.
- ❖ Considering the importance of segmenting detected polyps for achieving precise diagnosis, it was worth searching systematically, how colonoscopy databases are responding to two of the most influential variational segmentation methods, the geodesic and Chan–Vese active contour methods.
  - As the quality of colonoscopy databases varies, preprocessing steps were taken before evaluating 14 filtered images as inputs for the active contour methods.
  - The performance of the segmentation techniques was measured using the Sørensen-Dice Similarity Coefficient. The effects of the initial mask shape and its size together with the number of iterations, contraction bias and smoothness factor were studied.

Finally, it is worth mentioning that all steps formulated in this dissertation have been implemented by computer simulation in a MATLAB environment. All image processing, analyzing, and visualizing were performed using the powerful image processing toolbox as well as codes developed for the given purpose by me and our research group.

## 2. Colorectal polyps' detection on colonoscopy images using fuzzy inference

### 2.1 Basic concepts of fuzzy classification

#### 2.1.1 Introduction to fuzzy sets

Fuzzy set theory is a novel mathematical framework for dealing with uncertainty and imprecision. It was introduced by Lotfi A. Zadeh in 1965 (Zadeh, 1965) as an extension of classical set theory, with the aim of providing a formalism for representing and manipulating vague concepts. Fuzzy sets have been successfully applied in a wide range of fields, including artificial intelligence, control theory, decision analysis, pattern recognition, and many others.

Fuzzy sets concept reflects the nature of the human thinking logic, where there are numerous or even infinite transitions or possibilities between the concepts such as the human characteristics like tall and short man, or the weather conditions like cold and hot weather.

In contrast to classical sets, where an element either belongs or does not belong to a set, Zadeh's fuzzy sets theory emphasizes that the degree of an element's membership in a set is more important than simply determining whether it belongs to the set or not. Fuzzy sets were developed to allow partial membership degrees to be assigned to the set elements. These membership degrees are represented by real values ranging from fully belonging "1" to not belonging at all "0" (Zadeh, 1965). Figure 2-1 illustrates the difference between crisp and fuzzy membership functions.

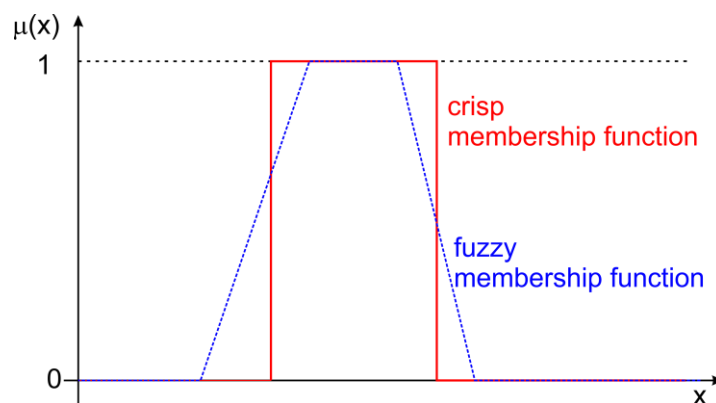


Figure 2-1 Examples of crisp and fuzzy membership functions

#### 2.1.2 Fuzzy inference

Fuzzy inference systems as computing frameworks rely on fuzzy set theory to reach the conclusion and connect the system antecedents and consequents. A rulebase and a corresponding

inference engine make up a typical fuzzy inference system. Many types of inference systems have been developed since Lotfi A. Zadeh first came out with the idea of non-crisp sets (Zadeh, 1965), (Zadeh, 1968); the most commonly utilized are the Mamdani inference in 1977 (Mamdani, 1977), and the Takagi-Sugeno inference in 1985 (Takagi & Sugeno, 1985).

The Mamdani fuzzy model is more intuitive and well-suited for linguistic input, as it consists of rules, like “IF... THEN ....” and the connection between the rules can be expressed with “AND” as the lowest of the membership values (called t-norm), and “OR” as the highest (called s-norm or t-conorm).

The aggregated consequent fuzzy set arises as the s-norm, i.e., the maximum of the results for the whole model consequents, whereas the result for each individual consequent is the t-norm, i.e., the minimum of the antecedents' rules belonging to that consequent. However, the outputs are often fuzzy sets, which makes a defuzzification step at the end necessary. An example of an inference system with 2 antecedents and 2 consequents can be seen in Figure 2 - 2.

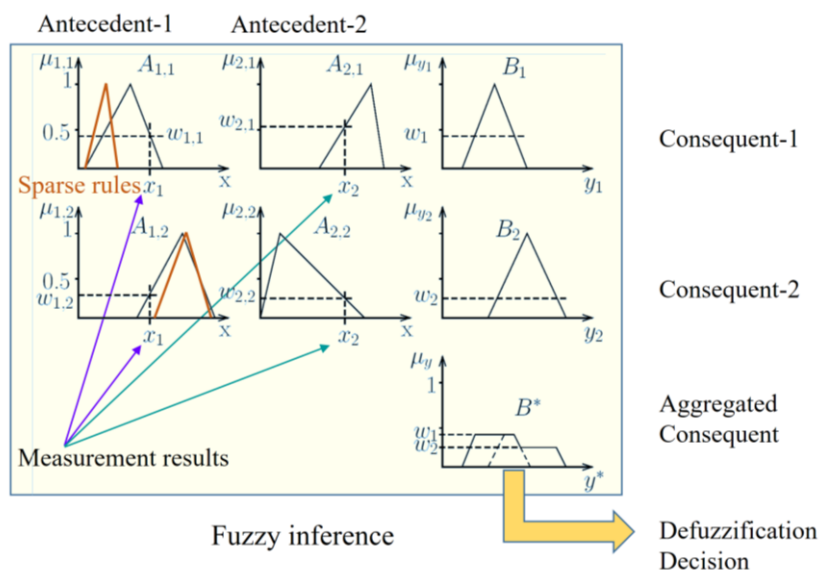


Figure 2-2 Fuzzy inference system based on 2 antecedents. The dark membership functions demonstrate such antecedents, where the overlap between the supports of the rules is sufficient, while the orange rules give an example of sparse rulebase

Unlike Mamdani-type fuzzy systems, which use fuzzy sets as outputs, the Takagi-Sugeno-Kang (TSK) model is a type of fuzzy inference system that uses a set of fuzzy if-then rules and produces crisp outputs by computing weighted averages of the inputs for each rule. The rules in a Sugeno model have the form “if  $x$  is  $A$  and  $y$  is  $B$ ”, then  $z = px + qy + r$ , where  $A$  and  $B$  are fuzzy sets defined on the input variables  $x$  and  $y$ , respectively, and  $p$ ,  $q$ , and  $r$  are

constants representing the linear coefficients and the offset of the linear function for the output variable  $z$ . The output of the Sugeno model is obtained by taking the weighted average of the outputs of all the rules, with the weights determined by the degree of match between the input values and the fuzzy sets defined in the if-part of the rules.

A Mamdani or Takagi-Sugeno-Kang (TSK) rulebase styles can be derived from data or obtained based on professional experience. Both the Mamdani and the TSK inference models can only be used with dense rulebases that cover the whole domain of the relevant antecedents. However, the most commonly used triangular, gaussian, or trapezoidal rulebases, often fail to form a dense set of rules, which makes the above-mentioned inference systems not applicable for antecedents, where the measured value in at least one of the dimensions has zero membership degree in all the rules, as it can be seen with the orange-colored rules in the first antecedent of Figure 2-2.

### **2.1.3 Fuzzy rule interpolation**

The need for dense rulebases in classical fuzzy inference systems can be substituted by the application of fuzzy interpolation. Interpolation can be used efficiently to produce appropriate outputs when a given observation does not overlap with any rule. Fuzzy interpolation can also contribute to the minimization of the system complexity; it can eliminate rules that can be estimated by the neighboring rules (Li, Yang, Qu, & Sexton, 2018). There are many interpolation approaches in fuzzy systems, and the stabilized Kóczy–Hirota inference method (SKH) method, which was developed in 1993, is one of the most important milestones (Koczy & Hirota, 1993), (Tikk, et al., 2002).

This approach relies on the natural concept that when an observed value (even if it is fuzzy) is more similar to one of the antecedent rules, its output will also be more similar to the corresponding consequent. To apply this approach, it is essential to compute the distance between the observed value and the fuzzy sets. If the membership functions are simple triangular ones, this distance can be calculated by the characteristic points of the functions, such as the vertices of the corresponding triangle.

Considering the nearest antecedent sets, SKH creates temporary rules for the domains not covered by any antecedent set for each observation. These new rules can be used only for the cases they were created for. The method calculates the infima and suprema of the characteristic

$\alpha$ -cuts of the new fuzzy conclusion  $B^*$  according to the Euclidean distances between the observation and the bounds of the characteristic  $\alpha$ -cuts of the original antecedent sets and conclusion. The method itself is given by the following formulas

$$\inf B_{\alpha}^* = \frac{\sum_i \inf B_{i\alpha} \left( \frac{1}{d_{\alpha L}(A^*, A_i)} \right)^k}{\sum_i \left( \frac{1}{d_{\alpha L}(A^*, A_i)} \right)^k}, \quad (2.1)$$

$$\sup B_{\alpha}^* = \frac{\sum_i \sup B_{i\alpha} \left( \frac{1}{d_{\alpha U}(A^*, A_i)} \right)^k}{\sum_i \left( \frac{1}{d_{\alpha U}(A^*, A_i)} \right)^k}, \quad (2.2)$$

with  $A_i$  being the  $i$ th antecedent,  $A^*$  the measured value,  $k$  is the number of antecedents, and  $d_{\alpha L/U}$  represent the euclidean distances between the lower and upper bounds of the distance between the  $\alpha$ -cuts of observation and the antecedents, and  $B^*$  stands for the corresponding fuzzy conclusion.

In the following sections, I show in detail my main contributions to the development of the fuzzy inference strategy. First, I begin by explaining why a uniform classification scheme is used and clarifying the antecedents derivation process. Then, I propose two rulebase generation methods, and I compare their results with the original method. Finally, I study the influence of introducing more consequent classes on the classification performance, and also, I discuss the practical results.

#### 2.1.4 Evaluation metrics

For the evaluation purposes, we assessed the tiles in the test set and compared their inferred polyp content with the corresponding image ground truth masks, which had been manually drawn by the professional personnel. The true positive rate (TPR), the false positive rate (FPR), the true negative rate (TNR) and the false negative rate (FNR) were basically utilized in the process of evaluating the classification results of the three rulebase generation methods. These parameters are mathematically expressed below.

1. True Positive Rate (TPR): TPR, also known as sensitivity or recall, is the proportion of true positives (correctly classified positive cases) among all actual positive cases. It is calculated as the ratio of true positives to the sum of true positives and false negatives.

$$TPR = \frac{\text{True Positive}}{\text{Actual Positive}} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})}. \quad (2.3)$$

2. False Positive Rate (FPR): FPR is the proportion of false positives (incorrectly classified positive cases) among all actual negative cases. It is calculated as the ratio of false positives to the sum of false positives and true negatives.

$$FPR = \frac{\text{False Positive}}{\text{Actual Negative}} = \frac{\text{False Positives}}{(\text{False Positives} + \text{True Negatives})}. \quad (2.4)$$

3. True Negative Rate (TNR): TNR, also known as specificity, is the proportion of true negatives (correctly classified negative cases) among all actual negative cases. It is calculated as the ratio of true negatives to the sum of true negatives and false positives.

$$TNR = \frac{\text{True Negative}}{\text{Actual Negative}} = \frac{\text{True Negatives}}{(\text{True Negatives} + \text{False Positives})}. \quad (2.5)$$

4. False Negative Rate (FNR): FNR is the proportion of false negatives (incorrectly classified negative cases) among all actual positive cases. It is calculated as the ratio of false negatives to the sum of false negatives and true positives.

$$FNR = \frac{\text{False Negative}}{\text{Actual Positive}} = \frac{\text{False Negatives}}{(\text{False Negatives} + \text{True Positives})}. \quad (2.6)$$

Also, besides TPR, TNR, FPR and FNR, more evaluation metrics were introduced to make the evaluation process more comprehensive in the advanced stages of this chapter. The Accuracy (ACC), Matthews Correlation Coefficient (MCC), and the F<sub>1</sub> score were used, they are defined the following way:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN}. \quad (2.7)$$

$$MCC = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}}. \quad (2.8)$$

$$F_1 \text{ score} = \frac{2*TP}{2*TP+FP+FN}. \quad (2.9)$$

## 2.2 Application of fuzzy rule interpolation in colorectal polyp detection

### 2.2.1 Introducing the uniform classification scheme

Continuing the development of the fuzzy inference-based approach that was proposed by the research team at our university is one of the primary goals of this chapter. The approach objec-



tive is not to achieve perfect polyp segmentation within colonoscopy images, but rather to highlight image segments that potentially contain polyps and draw the attention of medical personnel to detect and classify colorectal polyps more efficiently.

The colonoscopy databases have various types of polyps, and these polyps have been captured from different angles and viewpoints. Moreover, there are difficulties arising from variations in the bowel cleansing, these are other challenges to consider. To address these issues, a grouping process to organize the images within the databases into distinct image groups was employed. Each group consisted of subsequent images taken from the same examination with relatively similar angles of the polyps and similar properties, i.e., color, contrast, focus, and cleanliness state.

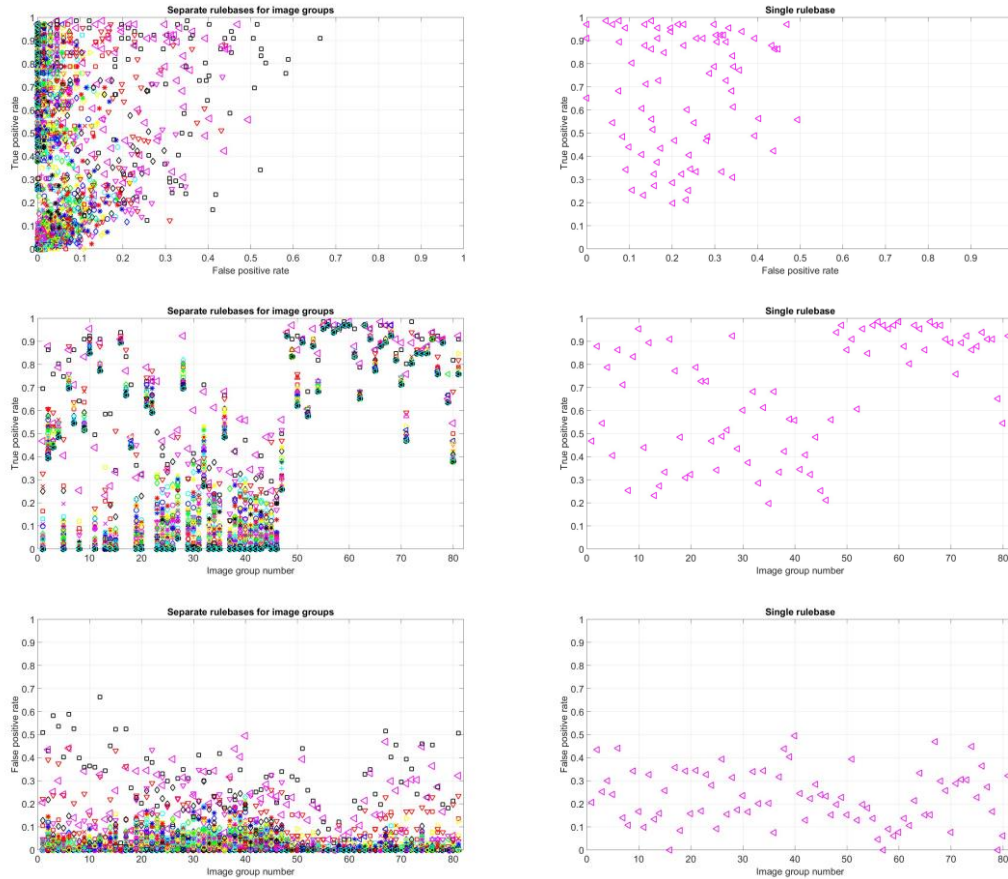
The three databases CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and Etis-Larib (Silva, Histace, Romain, Dray, & Granado, 2014) were organized in 81 image groups. The images were divided into smaller tiles of different sizes based on their original size:  $30 \times 30$ ,  $40 \times 40$ ,  $50 \times 50$ ,  $60 \times 60$  , and  $70 \times 70$  for small images;  $60 \times 60$ ,  $80 \times 80$ ,  $100 \times 100$ ,  $120 \times 120$ , and  $140 \times 140$  for medium images; and  $120 \times 120$ ,  $160 \times 160$ ,  $200 \times 200$ ,  $240 \times 240$ , and  $280 \times 280$  for large images. As the size of the tiles did not influence the results in a consequent way, and the middle one was the best for almost all cases, the size series  $50 \times 50$ ,  $100 \times 100$ ,  $200 \times 200$  were used in most of the studies, and I followed this approach as well.

A training set was constructed from every second tile, while the remaining half of the tiles was used to test the derived rulebases.

A very simple statistical approach was used in the first attempts of generating the training set fuzzy rules of the two outputs “contains polyp”, and “does not contain polyp”. The mean of the measured data was selected as the  $\alpha = 1$  cut point of the triangular rule, while the support was generated from the minimum and maximum measured values of the empirical measured data from the training set’s part corresponding to the respective consequent.

The results for the 81 image groups with the 81 different rulebases are presented in Figure 2-3. Also, the results of a joint, single rulebase based on the statistical data of the training set consisting of the training sets of all the image groups can be seen. It is visible that in most of the cases the fuzzy inferences with rulebases of separate image groups outperform that of the joint rulebase, not only for the group the rules were generated from, but also for most of the other groups. However, in this research, as it is more realistic to apply the uniform rulebase, only the performance of classification of the joint rulebase was studied. Developing a uniform

classification scheme is motivated by the need for very rapid and simple algorithms in the live video image classification applications.



*Figure 2-3 True positive rates and false positive rates of the stabilized Kóczy-Hirota fuzzy rule interpolation-based colonoscopy image classification. First column: the results for all the 81 image groups, of which the first 18 belongs to CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), the next 36 belongs to CVC-Clinic (Bernal J. , et al., 2015), and the last 36 belongs to ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014). The different colors mean rulebases based on training sets from different image groups. Second column: the results for a joint, common rulebase. The magenta larger triangles of the second column are also visible in the graphs of the first column, for comparison*

## 2.2.2 Derivation of the antecedents

The selection of the antecedents for the fuzzy classification system must be done very carefully, considering the complex spatial, biological, and optical conditions present during colonoscopy. Moreover, for the algorithm to be suitable for live video investigation, it is necessary to have computationally simple antecedents. Also, considering that medical experts and some patients

prefer to comprehend the functionality of a program on their device, it is recommended to use easily understandable input parameters.

The endoscopic camera operates in a dark, wet, and mucous environment, illuminated by multiple point-like light sources, resulting in pictures with significant reflections. Thus, as the polyps are protruding into the bowel, even if their native color is the same as the bowel wall, their color representation is different due to lighting. Moreover, the natural pink color of the bowel wall restricts the color range of the images, except in cases where the bowel is inadequately cleansed, or a malignant lesion is present.

The colonoscopy images were sorted into three categories based on the polyp angle of view, namely zenithal, lateral, and semilateral (Bernal J. , et al., 2015). The zenithal view presents the polyps from a top perspective, where they appear not significantly brighter than the surrounding area. However, their circumference and surface texture are usually distinguishable. Whereas the polyps viewed from a lateral angle have a much darker background with a low - light conditions bowel wall. And lastly, there is an in-between view, the semilateral view, where the background is somewhat darker than the polyp, and the contour, particularly at the polyp's bottom, is barely detectable.

Considering the characteristics of polyps mentioned above, the simplest statistical parameters: mean and standard deviation were proposed to be studied, together with the mean and standard deviation of the gradient-filtered image, since the gradients surrounding the polyp-blobs can be relatively large.

When the edges in an image are prominent and clearly defined, they can be valuable in determining the location of a polyp (Maini & Aggarwal, 2009). This was the reason for calculating the edge-to-pixel ratio using the Canny edge detection filter as another simple antecedent as well. Edges in a colonoscopy image segment may arise due to the polyp contours, polyps' texture or the micro-architecture of epithelial pits, and the presence of veins or folds in the bowel wall. In cases where a segment contains a lot of edges, the density of edges within that segment will be high. Conversely, smoother image segments tend to have lower edge density values.

The other set of antecedents is relatively unique, it is based on the Rényi entropies (Rényi, 1960) of the image

$$S_n = \frac{1}{1-n} \ln(\sum_{i=1}^N I_i^n), \quad (2.10)$$

which is a generalization of the well-known Shannon entropy  $S_1$  (Shannon, 1948). Here  $I_i$  denotes the  $i$ th pixel intensity normalized to be a probability distribution component. The structural entropy  $S_{str}$  and the spatial filling factor  $\ln q$  of the normalized pixel intensities were also calculated, and they are given by the following formulas

$$S_{str} = S_1 - S_2, \quad (2.11)$$

$$\ln q = S_0 - S_2. \quad (2.12)$$

These quantities are proven to give information about the shape of the distribution (Pipek & Varga, 1992), (Varga & Pipek, 2003), and they are effective in characterizing surfaces (Molnár, Nagy, & Mojzes, 2009), (Bonyár, Molnár, & Harsányi, 2012), (Bonyár, 2016). It is also shown that structural entropy is different along a polyp-like object than in the case of the probable background patterns in bowels (Nagy, Sziova, & Pipek, 2019).

The powerful widely used tool for filtering noise, capturing fine details, analyzing rough scale behavior, and detecting various patterns is Wavelet transform (Nagy, Sziova, & Koczy, 2017). Wavelet transform results in a set of coefficients that represent the image at different scales and orientations. The low-frequency coefficients correspond to the coarse details of the image, while the high-frequency coefficients capture the fine details and edges. Since there are rapid changes around the polyps, the fine-scale components were expected to have relatively greater significance in the regions surrounding the polyps. Due to this observation, the same parameters were calculated for the wavelet transform four output types (the low-pass–low-pass, the two mixed and the high-pass–high-pass) as for the original images themselves.

Using the RGB color space a total of 99 input parameters which include the measurable quantities mentioned earlier were calculated. Table 2-1 contains an ordered list of these input parameters, which were used as an initial set of antecedents for the fuzzy rulebase.

### 2.2.3 Median-centered and histogram-fitted rulebase generation methods

In the first attempts of generating the fuzzy rules of the two outputs for the training set using the mean of the measured data, the support of the resulting rules of the tiles with polyp was very often included into the support of the non-polyp rules, and the  $\alpha = 1$  points were often also near, thus plotting the histograms of the measured antecedents was introduced to analyze the shape of the probability distributions in the complete training set. Lot of the histograms

*Table 2-1 The numbering of the antecedent parameters. R,G, and B means the color channel,  $S_{str}$  is the structural entropy,  $\ln q$  is the logarithm of the spatial filling factor from (Pipek & Varga, 1992). In the case of the wavelet transforms, LL stands for the Low-pass Low-pass filter branch output, i.e., the rough details, HH for the output of the High-pass High-pass branch, i.e., the fine details, while the LH and HL combinations denote the mixed outputs, i.e., if one of the directions is filtered with Low-pass, the other with High-pass filter*

<i>Antecedent Index</i>	<i>Antecedent Type</i>
1–2	mean and standard deviation, R
3–4	mean and standard deviation, G
5–6	mean and standard deviation, B
7	edge density, R
8	edge density, G
9	edge density, B
10–11	$S_{str}, \ln q, R$
12–13	$S_{str}, \ln q, G$
14–15	$S_{str}, \ln q, B$
16–30	similar to 1–15, wavelet transform LL
31–45	similar to 1–15, wavelet transform LH
46–60	similar to 1–15, wavelet transform HL
61–75	similar to 1–15, wavelet transform HH
76–77	gradient magnitude's mean and standard deviation, R
78–79	gradient magnitude's mean and standard deviation, G
80–81	gradient magnitude's mean and standard deviation, B
82–87	similar to 76–81, gradient direction
88–93	similar to 76–81, gradient x component
94–99	similar to 76–81, gradient y component

seem very different from the shape of the applied mean fuzzy rules as it can be seen in Figure 2-4. In the next three sections, I give a brief characterization of the histograms, based on Figure 2-4.

The histograms of the mean values of the original tile have two wide peaks apart from each other for the two consequents (see No. 1, 3 and 5), while the standard deviations have almost the same peak location (see No. 2), with one of the peaks being visibly broader. The edge densities of the original form two almost overlapping wide heaps (see No. 8), while the

structural entropy and the filling factor have only slightly different thin peaks at the sides of the interval, sometimes with more smaller peaks (see No. 10, 11 and 12), (this again suggests building a hierarchical classification scheme where these entropy-based quantities play role only in some cases, even though a hierarchical classification scheme would make the calculations more complicated and time demanding).

The wavelet analysis changes the histograms of the mean values to thinner and more distinguishable peaks (see No. 1, compared to 31 and 61), the standard deviations to slightly thinner, too (see No. 2, compared to 32 or 21 and 51) and the edge densities to more distinguishable (see No. 22, 23, 38, and 52). The structural entropy and the filling factor also become even thinner (see No. 40 and 44).

The gradients' mean values usually differ only slightly in either width or peak location (see No. 90), while their standard deviations have peaks at the lower values which can be located at almost the same, or slightly different positions (see No. 87 and 91).

As the shape of the probability distributions in the training sets differ from the applied fuzzy rules very much, the goal became to determine whether selecting fuzzy membership functions fitting more to the measured data improves the classification properties. Two approaches were used. First, instead of having the  $\alpha = 1$  point to the mean of the measured values, it was moved to the median, which improves the fitting only slightly, as it can be seen with the dotted lines in Figure 2-4. This approach will be referred to as "median-centered" one in the following discussions. Second, to achieve better fitting the  $\alpha = 1$  peak of the triangular membership function was set to the maximum of the histogram, while the infimum and supremum of the support were set to those points, where the histogram values first cross a given percentage of its peak value, starting from both sides of the interval  $[0; 1]$ . In some cases, this also resulted in bad fitting (see antecedent No. 52 in Figure 2-4) but mostly this fits the histograms rather well. Later this method will be called "histogram-fitted" rule generation. This is the method which produces highly sparse rules (i.e., containing gaps), especially when the support threshold is set to higher levels, causing the support of the histogram-based rules to become narrower, like the 10% case shown in Figure 2-4. This was the reason why the need for rule interpolation arose, and the application of one of the interpolation methods, such as the SKH approach, was necessary.

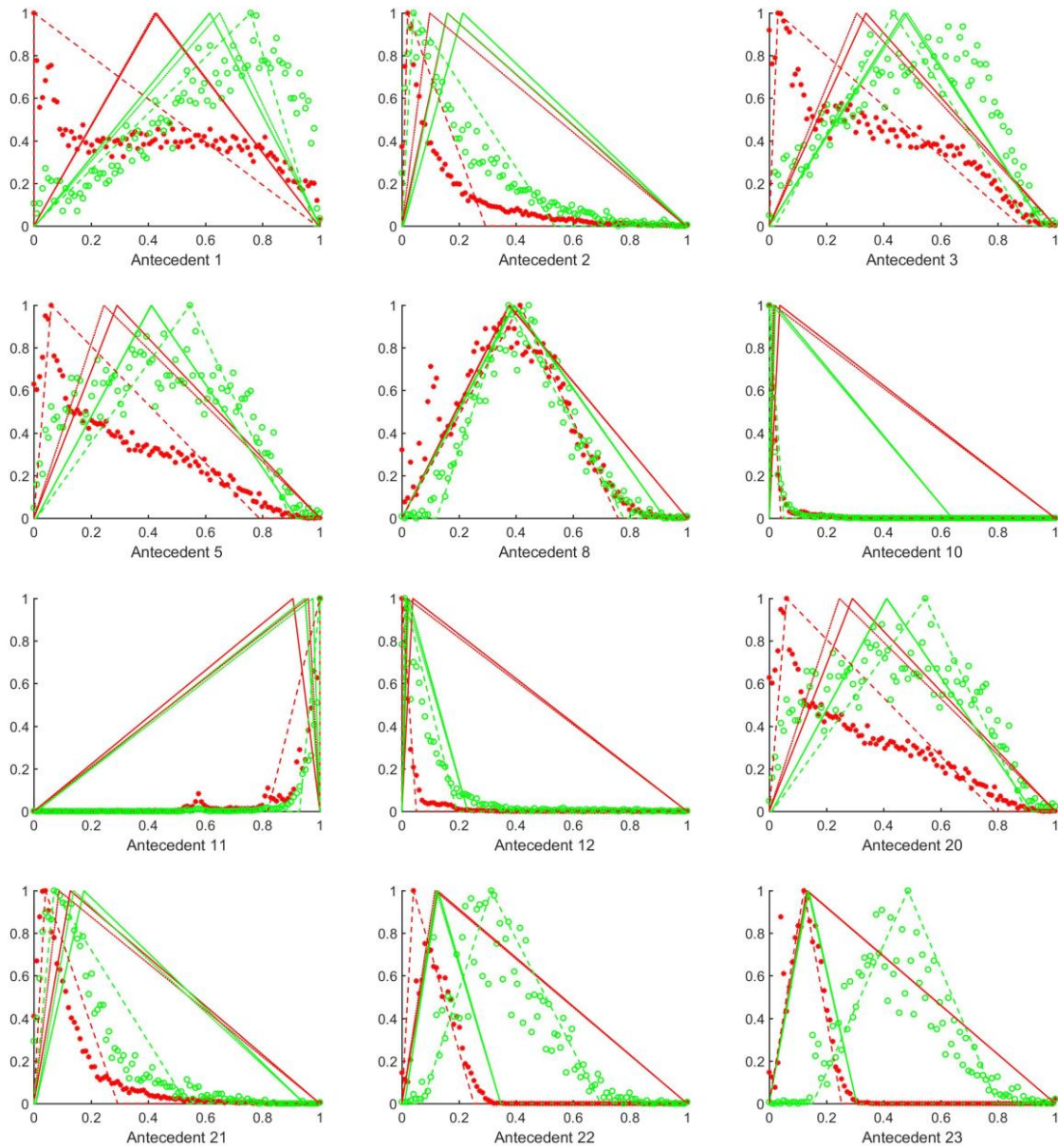


Figure 2-4 Examples of the RGB color space histograms together with the triangular rules resulting from the three rulebase generation methods of the measured data in the training sets for the three databases. Red color means the consequent “with polyp”, green color means the consequent “no polyp” cases. The histogram points are denoted by red stars for the “with polyp” cases, and green circles for the “no polyp” cases. Continuous line denotes the original rules with the mean value being the  $\alpha = 1$  peak of the triangular rule, dotted line the median-centered ones, and dashed line the histogram-based rules with the center being the peak of the histogram and the support set to the location where the histogram first meets the threshold level 10%. All the measured data are normalized so that the training set would be located into the closed interval  $[0; 1]$ , while the histogram peaks were scaled to be 1, for better visibility



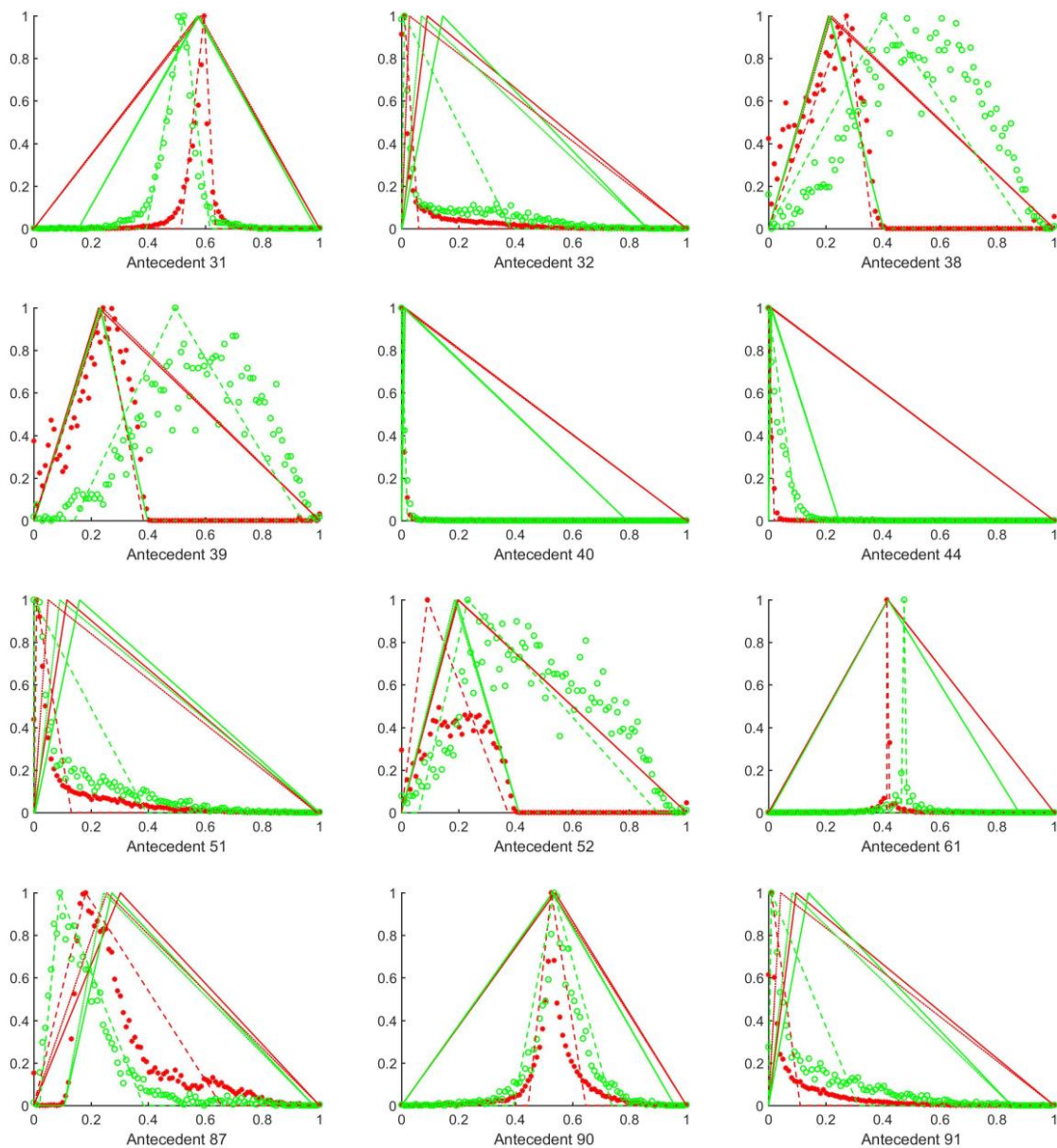


Figure 2-4 (continued): Examples of the RGB color space histograms together with the triangular rules resulting from the three rulebase generation methods of the measured data in the training sets for the three databases. Red color means the consequent “with polyp”, green color means the consequent “no polyp” cases. The histogram points are denoted by red stars for the “with polyp” cases, and green circles for the “no polyp” cases. Continuous line denotes the original rules with the mean value being the  $\alpha = 1$  peak of the triangular rule, dotted line the median-centered ones, and dashed line the histogram-based rules with the center being the peak of the histogram and the support set to the location where the histogram first meets the threshold level 10%. All the measured data are normalized so that the training set would be located into the closed interval  $[0; 1]$ , while the histogram peaks were scaled to be 1, for better visibility



#### 2.2.4 Experimental study and analysis

In the first reference case, all the 99 antecedents from Table 2-1 for RGB colour space were used. The performance of the “median-centered” and “histogram-fitted” rule generation methods was compared to the original “mean-centered” method by the following two directions, how the selection of the threshold for the histogram fitting, and the selection of the antecedents influence the classification results?

Later, however, during the experiments three representative threshold levels {1%, 5%, and 10%}, and three representative number of antecedents {66, 50, and 33} were selected for further study.

The antecedents' selection was made in a way where the histograms, and thus the resulting rules had the maximum dissimilarity. The reason for decreasing the number of the antecedents was to give the priority to the antecedents that had more different membership functions for the two consequents. The antecedents were sorted according to the total sum of difference in the magnitudes of the three characteristic points of the triangular membership functions for the two consequents. The larger the sum of magnitude differences was, the better the rule for classification was considered. For each type of the three rulebase generation methods, the 66, 50 and 33 antecedents with the largest differences were used in the classification, respectively.

The classification results of the three rulebase generation methods using the basic evaluation metrics are numerically quantified in Table 2-2.

In the 99 antecedents case, the positioning of the  $\alpha = 1$  point of the triangular membership function to the median of the measured data improves the performance slightly for the “yes polyp” case (TPR increased) but makes more tiles without polyp segment be detected as polyp containing one (FPR increased). However, moving the membership functions to histogram based central points and supports does not improve the results, even though the true positive rate improves significantly, the true negative rate drops to below 50%. (For these histogram-fitted results, if all the antecedents were taken into consideration, then such rules were also included, where the support of the membership functions totally overlapped, thus only the core points of the rules determined the results in the SKH rule interpolation. This is the reason why the last 3 lines are the same. This statement holds for all the cases if the used antecedents are the same, the outputs will be the same for all histogram percentages.)

If we study the effect of leaving out the antecedents with the most overlapping membership functions, then, it is clear, that the histogram-fitted rulebases are not much influenced. For

Table 2-2 Classification performance for the different methods with different number of selected antecedents

<i>Method</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>
<b>Number of antecedents: 99</b>				
<b>Mean-centered</b>	0.6001	0.6698	0.3302	0.3999
<b>Median-centered</b>	0.6894	0.593	0.407	0.3106
<b>Histogram-fitted 1%</b>	0.8273	0.4371	0.5629	0.1727
<b>Histogram-fitted 5%</b>	0.8273	0.4371	0.5629	0.1727
<b>Histogram-fitted 10%</b>	0.8273	0.4371	0.5629	0.1727
<b>Number of antecedents: 66</b>				
<b>Mean-centered</b>	0.6584	0.6257	0.3743	0.3416
<b>Median-centered</b>	0.683	0.6068	0.3932	0.317
<b>Histogram-fitted 1%</b>	0.8286	0.4325	0.5675	0.1714
<b>Histogram-fitted 5%</b>	0.8281	0.4352	0.5648	0.1719
<b>Histogram-fitted 10%</b>	0.8271	0.4366	0.5634	0.1729
<b>Number of antecedents: 50</b>				
<b>Mean-centered</b>	0.6712	0.6193	0.3807	0.3288
<b>Median-centered</b>	0.6736	0.6093	0.3907	0.3264
<b>Histogram-fitted 1%</b>	0.8091	0.4559	0.5441	0.1909
<b>Histogram-fitted 5%</b>	0.7895	0.482	0.518	0.2105
<b>Histogram-fitted 10%</b>	0.7814	0.4912	0.5088	0.2186
<b>Number of antecedents: 33</b>				
<b>Mean-centered</b>	0.7651	0.3989	0.6011	0.2349
<b>Median-centered</b>	0.5607	0.6771	0.3229	0.4393
<b>Histogram-fitted 1%</b>	0.7828	0.4905	0.5095	0.2172
<b>Histogram-fitted 5%</b>	0.7895	0.48	0.52	0.2105
<b>Histogram-fitted 10%</b>	0.7779	0.4922	0.5078	0.2221

the mean-centered rules, the true positive rates improve, however, the true negative rates deteriorate. For the median-centered rules the tendency is the opposite, but none of these tendencies are very expressed. It is interesting to mention, that in the case of 33-antecedents, for the mean-centered case the true negative rate became extremely low, while for the median-centered case it increased, but there the true positive rate drops significantly compared to the other numbers of antecedents.

If the tendency between the histogram percentages that determine the support of the rules is studied, I can conclude the following. The 1% case gives the best results, and as the number of highly overlapping rules that are excluded increases, this tendency becomes clearer.

### ***2.3 The influence of introducing more consequent classes on the classification performance***

The influence of selecting the rulebase parameters on improving the effectiveness of the classification system based on stabilized Kóczy-Hirota rule interpolation method was investigated in the previous study (Sziová, Ismail, Lilik, Kóczy, & Nagy, 2020). This work was extended by authors in (Sziova, Nagy, & Fazekas, 2021) to a more thorough antecedents selection approach.

Moreover, as it was found in the first histograms analysis (Sziová, Ismail, Lilik, Kóczy, & Nagy, 2020) that many of the antecedents had histograms with multiple local maxima, this research was extended to check if these multiple peaks belong to image segments with different polyp percentages: the image segments that contain only a small portion of a polyp might behave differently from those that are almost entirely covered with a polyp part.

The point, which was thought to be interesting to study, is the possibility of including more consequents between “containing polyp” and “not containing polyp” classes to better match real-life problems. The number of consequents was increased from 2 - contains polyp or does not contain polyp - to more output classes based on the polyp percentages of the image segments. The basic idea is to allow the tiles with different polyp coverage to form different, more structured classes, instead of the bulk “no polyp” and “yes polyp” ones, as the polyps behave differently along their perimeters and inside, even if only protruding, roundish polyps were studied.

Regarding the newly tested output classes, first the consequent classes as “low polyp percentage” and “high polyp percentage” were introduced instead of the “yes polyp” class, with a borderline at 50% polyp coverage. Then, an even more refined, 4-step system was studied, with thresholds at 20% and 50% polyp percentages. The no polyp class was kept intact (Ismail, Sziova, Taha, & Nagy, 2022).

#### **2.3.1 Experimental results**

Analyzing the colonoscopy images in the HSV color space, which separates color components from intensity, may offer additional valuable information, and the switching from RGB to HSV color space has also been proven to have a positive effect on the results in (Sziova, Nagy, & Fazekas, 2021). That was the reason why in this study, the same 99 antecedents were also used

as an initial set of antecedents for the fuzzy rulebase, but they were also calculated in the HSV color space as well.

Together with the initial 99-antecedents case and using the antecedent selection process similar to the one followed with the study of the rulebase parameters, the influence of introducing more consequent classes on the classification performance was studied using the 66, 50, and 33 most distinguishable antecedents.

Besides TPR, TNR, FPR and FNR, the Accuracy (ACC), Matthews Correlation Coefficient (MCC), and the  $F_1$  score were used to determine whether the introduction of the new output classes and the new color space improve the results. As these metrics are valid for binary classification tasks, and our requirement is also a binary classification, I merged the 3 and 4 output classes into binary decision the following way. Both for 3 and for 4 classes, the class with 0% polyp content remained the non-polyp output, and all the other categories were built into the “with polyp” case.

The resulting histograms and the mean-, median-centered, and histogram-fitted 1% triangular rules for the 2, 3, and 4 consequent classes of all 99 antecedents for RGB and HSV color spaces are plotted in Appendix A and Appendix B, respectively. Tables 2-3, 2-4, and 2-5 list the numerical results for all evaluation metrics.

The true positive rates improved by introducing the 3 and 4 consequent classes, however, the true negative rates became worse, both for the RGB and the HSV colour spaces. In some cases, especially for the histogram-based rulebases, the improvement of the true positive rate stops at the 3-class case.

For accuracy, the original, 2-class case is almost always the best. However, based on MCC and  $F_1$  score, the tendency is not clear, sometimes the 3-class, sometimes the 2-class case is the best, and there are some cases with the 4-outputs giving the best results. Generally, the 4-class case did not improve the results, while the 3-class and 2-class cases were comparable.

The  $F_1$  score emphasizes the ability to find the polyp cases, but it can be degraded by a very large false positive rate, if the number of the negative samples is much higher than the number of positive samples (i.e., the dataset is imbalanced), like in this case. This is the reason why  $F_1$  is below 0.5, even if the TPR is over 0.9.

Table 2-3 Classification performance for different number of selected antecedents with different consequent classes using mean-centered rulebase generation method for RGB and HSV color spaces

<i>Consequent classes</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>	<i>ACC</i>	<i>MCC</i>	<i>F<sub>1</sub></i>
<b>Number of antecedents: 99 RGB</b>							
<b>2 output classes</b>	0.6001	0.6698	0.3302	0.3999	0.65391	0.23216	0.44203
<b>3 output classes</b>	0.6561	0.623	0.377	0.3439	0.63052	0.23594	0.44786
<b>4 output classes</b>	0.7075	0.5741	0.4259	0.2925	0.60456	0.23647	0.44973
<b>Number of antecedents: 66 RGB</b>							
<b>2 output classes</b>	0.6584	0.6257	0.3743	0.3416	0.63318	0.24034	0.45056
<b>3 output classes</b>	0.6819	0.6085	0.3915	0.3181	0.62529	0.24474	0.45394
<b>4 output classes</b>	0.6726	0.5883	0.4117	0.3274	0.60754	0.21938	0.43909
<b>Number of antecedents: 50 RGB</b>							
<b>2 output classes</b>	0.6712	0.6193	0.3807	0.3288	0.63116	0.24529	0.45396
<b>3 output classes</b>	0.5154	0.6926	0.3074	0.4846	0.6521	0.18243	0.40358
<b>4 output classes</b>	0.8576	0.2976	0.7024	0.1424	0.4255	0.14811	0.40542
<b>Number of antecedents: 33 RGB</b>							
<b>2 output classes</b>	0.7651	0.3989	0.6011	0.2349	0.48256	0.14333	0.40319
<b>3 output classes</b>	0.7791	0.3809	0.6191	0.2209	0.47185	0.14136	0.40257
<b>4 output classes</b>	0.8888	0.2128	0.7872	0.1112	0.36723	0.10886	0.39085

<i>Consequent classes</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>	<i>ACC</i>	<i>MCC</i>	<i>F<sub>1</sub></i>
<b>Number of antecedents: 99 HSV</b>							
<b>2 output classes</b>	0.7218	0.563	0.437	0.2782	0.59925	0.23913	0.45143
<b>3 output classes</b>	0.7577	0.5374	0.4626	0.2423	0.58771	0.2482	0.45636
<b>4 output classes</b>	0.7776	0.5085	0.4915	0.2224	0.56995	0.24178	0.45235
<b>Number of antecedents: 66 HSV</b>							
<b>2 output classes</b>	0.7541	0.5618	0.4382	0.2459	0.60571	0.26526	0.46631
<b>3 output classes</b>	0.7897	0.5316	0.4684	0.2103	0.59052	0.27066	0.46834
<b>4 output classes</b>	0.8244	0.487	0.513	0.1756	0.56405	0.26519	0.46345
<b>Number of antecedents: 50 HSV</b>							
<b>2 output classes</b>	0.7437	0.5638	0.4362	0.2563	0.60492	0.25826	0.46239
<b>3 output classes</b>	0.7845	0.5292	0.4708	0.2155	0.58754	0.26437	0.4649
<b>4 output classes</b>	0.8199	0.4869	0.5131	0.1801	0.56298	0.26127	0.46151
<b>Number of antecedents: 33 HSV</b>							
<b>2 output classes</b>	0.7659	0.4671	0.5329	0.2341	0.53537	0.19859	0.42958
<b>3 output classes</b>	0.8403	0.4268	0.5732	0.1597	0.52124	0.23283	0.44499
<b>4 output classes</b>	0.8797	0.3268	0.6732	0.1203	0.45308	0.19316	0.42355

Table 2-4 Classification performance for different number of selected antecedents with different consequent classes using median-centered rulebase generation method for RGB and HSV color spaces

<i>Consequent classes</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>	<i>ACC</i>	<i>MCC</i>	<i>F<sub>1</sub></i>
<b>Number of antecedents: 99 RGB</b>							
<b>2 output classes</b>	0.6894	0.593	0.407	0.3106	0.61498	0.23744	0.44996
<b>3 output classes</b>	0.723	0.559	0.441	0.277	0.59619	0.23649	0.4499
<b>4 output classes</b>	0.762	0.512	0.488	0.238	0.56939	0.23161	0.44704
<b>Number of antecedents: 66 RGB</b>							
<b>2 output classes</b>	0.683	0.6068	0.3932	0.317	0.62419	0.24411	0.45365
<b>3 output classes</b>	0.709	0.571	0.429	0.291	0.60254	0.23527	0.44907
<b>4 output classes</b>	0.725	0.538	0.462	0.275	0.58063	0.22077	0.44115
<b>Number of antecedents: 50 RGB</b>							
<b>2 output classes</b>	0.6736	0.6093	0.3907	0.3264	0.62402	0.23855	0.45012
<b>3 output classes</b>	0.698	0.556	0.444	0.302	0.58832	0.21309	0.43643
<b>4 output classes</b>	0.787	0.312	0.688	0.213	0.42044	0.0914	0.38274
<b>Number of antecedents: 33 RGB</b>							
<b>2 output classes</b>	0.5607	0.6771	0.3229	0.4393	0.65054	0.20604	0.42301
<b>3 output classes</b>	0.732	0.328	0.672	0.268	0.42044	0.05435	0.36582
<b>4 output classes</b>	0.97	0.059	0.941	0.03	0.267	0.05448	0.37674

<i>Consequent classes</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>	<i>ACC</i>	<i>MCC</i>	<i>F<sub>1</sub></i>
<b>Number of antecedents: 99 HSV</b>							
<b>2 output classes</b>	0.5824	0.6547	0.3453	0.4176	0.63818	0.20323	0.42377
<b>3 output classes</b>	0.6325	0.6283	0.3717	0.3675	0.62928	0.22108	0.43799
<b>4 output classes</b>	0.6426	0.6005	0.3995	0.3574	0.61012	0.20492	0.4295
<b>Number of antecedents: 66 HSV</b>							
<b>2 output classes</b>	0.5794	0.6617	0.3383	0.4206	0.6429	0.20723	0.42573
<b>3 output classes</b>	0.6502	0.6282	0.3718	0.3498	0.63322	0.2357	0.44744
<b>4 output classes</b>	0.6792	0.5764	0.4236	0.3208	0.5999	0.21477	0.43676
<b>Number of antecedents: 50 HSV</b>							
<b>2 output classes</b>	0.5912	0.6672	0.3328	0.4088	0.64987	0.22229	0.43551
<b>3 output classes</b>	0.6566	0.6209	0.3791	0.3434	0.62906	0.23454	0.44707
<b>4 output classes</b>	0.7508	0.5319	0.4681	0.2492	0.58186	0.23783	0.45061
<b>Number of antecedents: 33 HSV</b>							
<b>2 output classes</b>	0.6429	0.6376	0.3624	0.3571	0.6388	0.23809	0.44849
<b>3 output classes</b>	0.6999	0.5774	0.4226	0.3001	0.60535	0.23287	0.44754
<b>4 output classes</b>	0.9734	0.0529	0.9471	0.0266	0.26318	0.05235	0.37636

Table 2-5 Classification performance for different number of selected antecedents with different consequent classes using histogram-fitted 1% rulebase generation method for RGB and HSV color spaces

<i>Consequent classes</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>	<i>ACC</i>	<i>MCC</i>	<i>F<sub>1</sub></i>
<b>Number of antecedents: 99 RGB</b>							
<b>2 output classes</b>	0.8273	0.4371	0.5629	0.1727	0.52627	0.22914	0.4438
<b>3 output classes</b>	0.8895	0.3307	0.6693	0.1105	0.45837	0.20586	0.42864
<b>4 output classes</b>	0.8512	0.37	0.63	0.1488	0.47989	0.19911	0.42777
<b>Number of antecedents: 66 RGB</b>							
<b>2 output classes</b>	0.8286	0.4325	0.5675	0.1714	0.52295	0.22664	0.44245
<b>3 output classes</b>	0.88	0.3517	0.6483	0.12	0.47236	0.21246	0.4324
<b>4 output classes</b>	0.8637	0.3523	0.6477	0.1363	0.4691	0.19733	0.42633
<b>Number of antecedents: 50 RGB</b>							
<b>2 output classes</b>	0.8091	0.4559	0.5441	0.1909	0.5366	0.2276	0.44376
<b>3 output classes</b>	0.9018	0.3211	0.6789	0.0982	0.4537	0.21073	0.4299
<b>4 output classes</b>	0.8701	0.3309	0.6691	0.1299	0.45404	0.18692	0.4213
<b>Number of antecedents: 33 RGB</b>							
<b>2 output classes</b>	0.7828	0.4905	0.5095	0.2172	0.55728	0.23192	0.44686
<b>3 output classes</b>	0.9151	0.2911	0.7089	0.0849	0.43359	0.20154	0.42463
<b>4 output classes</b>	0.9109	0.2931	0.7069	0.0891	0.43421	0.19875	0.42378

<i>Consequent classes</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>	<i>FNR</i>	<i>ACC</i>	<i>MCC</i>	<i>F<sub>1</sub></i>
<b>Number of antecedents: 99 HSV</b>							
<b>2 output classes</b>	0.9567	0.1987	0.8013	0.0433	0.37187	0.17658	0.41034
<b>3 output classes</b>	0.9678	0.1812	0.8188	0.0322	0.36088	0.17656	0.40888
<b>4 output classes</b>	0.9648	0.1664	0.8336	0.0352	0.34875	0.16048	0.4036
<b>Number of antecedents: 66 HSV</b>							
<b>2 output classes</b>	0.9538	0.2024	0.7976	0.0462	0.37407	0.17592	0.41044
<b>3 output classes</b>	0.9663	0.1848	0.8152	0.0337	0.3633	0.17751	0.40942
<b>4 output classes</b>	0.9651	0.1632	0.8368	0.0349	0.34633	0.1581	0.40277
<b>Number of antecedents: 50 HSV</b>							
<b>2 output classes</b>	0.9535	0.2025	0.7975	0.0465	0.37407	0.17568	0.41037
<b>3 output classes</b>	0.9665	0.185	0.815	0.0335	0.36347	0.17794	0.40955
<b>4 output classes</b>	0.9636	0.1717	0.8283	0.0364	0.35257	0.16329	0.40471
<b>Number of antecedents: 33 HSV</b>							
<b>2 output classes</b>	0.9575	0.1976	0.8024	0.0425	0.3712	0.17661	0.41026
<b>3 output classes</b>	0.9641	0.1931	0.8069	0.0359	0.3692	0.18129	0.41112
<b>4 output classes</b>	0.9658	0.1629	0.8371	0.0342	0.34627	0.15885	0.40294

Both accuracy and the MCC metrics reflect the method's general ability to evaluate the results. The neutral point, where the random guessing is located in the scales of these metrics is 0.5 for accuracy (with the scale of  $[0,1]$ ), and 0 for MCC (with a scale  $[-1,1]$ ). The fuzzy classification method is clearly better than this neutral point in the case of mean- and median-centered rulebases, except for the histogram-based method for HSV images. For mean and median most of the accuracies are above 0.6, and the best accuracies are 0.65: they are for the RGB color space, for the original, mean-centered, 99-antecedent, 2-class, and 50-antecedents, 3-class cases results and the median-centered, 33-antecedent, 2-class case. However, for these cases, the true positive rates are some of the lowest.

For MCC the best results can be achieved in the HSV color space, mean-centered case, for 66-antecedents, but the 50-antecedent case closely follows. In these cases, the number of output classes do not make great impact on MCC or  $F_1$ , however, the 3-class cases are the best.

From the rulebase generation methods point of view, clearly, the histogram-based one seems to be the least successful for these number of antecedents. The other two methods do not differ significantly.

Regarding the number of antecedents, there is no clear tendency. The TPR for the mean-centered rulebase generation method and 2-outputs improves as the number of the deleted antecedents increases (and also the TNR decreases), while for the median-centered method, the tendency is the opposite, for the RGB case. The other metrics have a kind of optimum for the 66- and 50- antecedent cases, especially for 3 output classes.

This implies that there is a tradeoff, not only for the true positive and true negative rates, but the other metrics as well, there is no clear optimum within the studied range of parameters and further research is needed.

## ***2.4 Thesis statements***

### **2.4.1 Thesis 1**

I suggested two novel rulebase generation methods for the application of fuzzy inference with stabilized Kóczy-Hirota rule interpolation in colorectal polyp detection, a median-centered and a histogram-fitted method. Both methods depend on the statistics of the training set. The median-centered method connects the characteristic points of a triangular membership function, i.e., the edges of its support and the core point to the minimum, maximum and median of the distribution of the training set, respectively. The histogram-fitted method generates triangular



rules where the core of the fuzzy membership function is the highest point of the 100-bin, equal-width histograms of the normalized measured data of each antecedent and each consequent in the training set. The support edges extend to the domain where the histogram values first cross a given percentage (1%, 5%, or 10%) of its peak value, starting from both sides of the interval [0;1].

- A. I compared the results of the two proposed rulebase generation methods to the original, mean-centered method. For three publicly available databases, I calculated the true positive rate, the false positive rate, the true negative rate and the false negative rate for antecedent numbers being  $2/3$ ,  $1/2$ , and  $1/3$  of the original 99 antecedents. I concluded, that the results of the three methods are comparable to each other for the studied numbers of antecedents.
- B. Among the histogram-fitted rulebases, the lowest percentage one, with the widest support gave the best results.

#### 2.4.2 Thesis 2

Based on the multiple peaks visible on the 100-bin, equal-width histograms of the distributions of the normalized measured data for many antecedents, I proposed to increase the number of consequents during the fuzzy decision process to improve the binary “with polyp”-“no polyp” classification of tiles of colonoscopy images. Instead of polyp area percentage of 0% or above 0% classes, I studied a 3-output class case with borderlines at 0% and 50%, and a 4-output class case with borderlines at 0%, 20%, and 50%.

- A. The true positive rates are the best for the histogram-fitted rules with HSV version of the images. However, in these cases, the true negative rates became very low. These results are not much influenced by the number of antecedents used and the number of consequent classes.
- B. The accuracy is the best in the case of the original 99-antecedents, 2-consequents, mean-centered case for RGB versions of the images.
- C. The Matthews Correlation Coefficient (MCC), as well as the  $F_1$  score are the best in the case of HSV version of the images, 3-consequents, 66-antecedents and mean-centered rules.

### **3. Colorectal polyps' localization in colonoscopy images using fuzzy circular Hough transform**

#### **3.1 Motivation**

In pattern recognition literature, Hough transform is considered as a great mathematical tool for object detection since the first appearance of its classical version for machine analysis of bubble chamber pictures by Paul V C Hough in 1959 (Hough, 1959). In 1981, it was extended by D. H. Ballard into the Generalized HT (GHT), which is a two-phase learning-detection process to detect arbitrary complex non-parametric shapes (Ballard, 1981). Continuously, different variations of Hough transform have been proposed by many researchers, like probabilistic Hough transform (Nahum, Eldar, & Bruckstein, 1991), randomized Hough transform (Lei & Erkki, 1993), and Vector–Gradient Hough Transform (VGHT) (Cucchiara & Filicori, 1998).

In 1994, Han, Kóczy, and Poston introduced the fuzzy Hough transform (Han, Kóczy, & Poston, 1994) to detect fuzzy lines and circles in the noisy environments by roughly fitting the data points to given parametric forms.

Hough transform and all its successive versions have proven to be typical powerful techniques with promising outcomes. They were used in numerous fields of applications like object detection (Zhao, Han, Zhang, Xu, & Cheng, 2022) and (Lin, 2020), lanes and roads detection (Liu, Zhang, Li, & Tao, 2017) and (Mathavan, et al., 2017), industrial automation (Pugin, Zhiznyakov, & Zakharov, 2018), mechanical engineering (Nagy, Solecki, Sziová, Sarkadi-Nagy, & Kóczy, 2020), robot navigation (Chen, Qiang, Wu, Xu, & Wang, 2021), and medical image processing fields (Nagy, Kovács, Sziová, & Kóczy, 2019), (Djekoune, Messaoudi, & Amara, 2017), (Hapsari, Utoyo, Rulaningtyas, & Suprajitno, 2020), (Vijayarajeswari, Parthasarathy, Vivekanandan, & Basha, 2019), and (Shaaf, Jamil, & Ambar, 2022).

Derived from prior experiences and a comprehensive analysis of previous studies employing the Hough transform, the practical strengths and weaknesses of this method are briefly outlined in Table 3-1.

The research community has investigated the limitations of Hough transform and suggested different approaches to make it a more plausible tool. To solve large calculation demand and to ensure integration with the Wireless Capsule Endoscopy (WCE) system, authors

*Table 3-1 The practical strengths and weaknesses of Hough transform*

<b>Strengths</b>	<b>Weaknesses</b>
✓ Not limited to specific shapes	☒ With more parameters, the dimension of the accumulator space increases
✓ Any parametric curve can be used as basis	☒ The calculation demand grows exponentially with the number of parameters
✓ Easy to parallelize (the transformation itself is independently executed for each edge point)	☒ Can be computationally expensive <ul style="list-style-type: none"> <li>- for high-resolution images</li> <li>- for complex shapes</li> <li>- for dense line images</li> </ul>
✓ Can identify partial and obstructed curves	☒ Shorter curve segments with fewer points result in weaker peaks in the accumulator space
	☒ Sensitive to noise. Noise may lead to additional peaks in the accumulator space
✓ It can detect multiple objects	☒ Overlapping objects may lead to unphysical, ghost peaks
	☒ Partially covered, or smaller objects can be suppressed by the larger objects with more points

in (Chuquimia, Pinna, Dray, & Granado, 2020) improved the real-time computation of the Hough transform. The design of the new approach took into consideration specific constraints of WCE such as limited space and limited energy resources. Within the same limitation direction (i.e., minimizing the Hough transform's computational cost), an Edge Orientation-based Fuzzy Hough Transform (EOFHT) was proposed in (Montseny, Sobrevilla, & Marès Martí, 2003). Instead of using all the edge-detected image points in the voting process, just those specific points whose representation is consistent with the selected gradient orientation range were eligible to vote.

Moreover, since shorter curves give fewer votes, thus, circles with smaller radii give weaker peaks in the accumulator space, a weighted vote which is inversely proportional to the radius in the parameter space was given to each entitled point in the image space by the modified voting method in (Barbosa & Vieira, 2019).

The efficiency of applying Hough transform in the field of automatic colorectal polyp detection and localization in colonoscopy images was also studied. This is a promising research

area and a challenging problem because of the high variability of the colorectal polyps' characteristics in both shape and texture.

Classical Hough transform was applied in (Silva, Histace, Romain, Dray, & Granado, 2014) to identify potential regions of interest (ROIs) in 300 video endoscopy pictures. A good detection of the ROIs containing a polyp was possible using classical Hough transform based on Canny edge detection approach. However, in several samples the method generated many alternative weaker circles and raised the classification system's False Positive Rate (FPR). To enhance the effectiveness of the proposed method, after Hough transform step, the textural characteristics from co-occurrence matrices were computed, and then used within a boosting-based technique for the final classification purpose. Hough transform was also used in other colorectal polyp localizing methods as a ready-made preprocessing step to find ROIs (Ruano, Barrera, Bravo, Gomez, & Romero, 2019), (Ruiz, Guayacán, & Martínez, 2019). However, they concentrated on the steps determining whether a ROI contains polyp or not, and not on the Hough transform step itself. Linear Hough transform was also used in colonoscopy for detecting folds within the bowel (Yao, Stidham, Soroushmehr, Gryak, & Najarian, 2019).

In my research, rather than improving the steps after selecting the ROIs with the Hough transform (as done in previously mentioned studies), an effort was made to overcome certain weaknesses of the Hough transform itself.

As a first step the fuzzy Hough transform was introduced, as it provides more tolerance to the deviations from the ideal curve's points, and colorectal polyps in real colonoscopy images are not precisely circular (Nagy, Ismail, Sziová, & Kóczy, 2021).

As a second step, the large computational demand of the Hough transform was the target, as the low computational load is one of the most essential requirements for algorithms used in computer-aided diagnosis (CAD) systems (Ismail, Prukner, & Nagy, 2023).

Hough transform starts with an edge detection (mostly Canny edge detection), and all the edge pixels have to be transformed. However, the edge pixels can be decimated, as the Canny edge detection tends to have a dense edge map because of the necessarily high connectivity of the edge points. Colonoscopy images contain edges other than colorectal polyps' borderlines, however, these edges often either have smaller or larger intensity steps compared to the edges of colorectal polyp contours. Based on these considerations, the possibility of removing Canny edge detected points with too small or too high gradient magnitude values was investigated in (Ismail, Prukner, & Nagy, 2023).

The research was performed on all the images of the three colonoscopy image datasets CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014). The study's goal was to eliminate edge points that do not belong to polyp contours, and at the same time, to keep as many polyp contour edge points, as it is necessary for the colorectal polyp contour to be detectable.

Setting a global gradient magnitude threshold domain that could achieve both a low total number of Canny edge pixels and at the same time a sufficiently accurate matching with the colorectal polyp contour was not possible. However, if the continuity of the Canny edges is given up, other edge detecting methods could provide a better basis for Hough transform, as there the continuity of the lines is not needed, only that the edge points are on the contours of the polyp. This is one of the ideas that were studied further in the following research which is presented in detail in this chapter (Ismail & Nagy, 2023).

The main contributions of this chapter are as follows. First, the performance of four edge detection algorithms (Canny, Prewitt, Roberts, and Sobel) was compared, and the most ideal one that gave the most polyp-contour-related and least unnecessary edge points was selected. Two metrics, based on the normalized gradients of contour and non-contour edges were used to determine which algorithm is the most appropriate. Second, to further reduce the number of edges that do not belong to the polyp contours, a gradient magnitude thresholding process was applied for the results of the selected edge detection method. Finally, to make the circle detection more tolerant to shape uncertainty, the fuzzy version of the Hough transform was also tested together with the classical one using a gradient-weighted voting approach. To evaluate the results, the radial displacement and the average radius were introduced to characterize the roundness of the objects to be detected. These contributions are summarized in the following points,

- selecting the edge detection method that is the most suitable for colorectal polyp localization purposes. Developing a metric to base this selection,
- determining gradient limits for removing the unnecessary edges,
- applying fuzzy Hough transform on colonoscopy images and comparing its results with the classical Hough transform,
- introducing a gradient-weighted voting to both classical and fuzzy Hough transforms and study its effects,
- characterizing the roundness of the objects to be detected.

In the following sections, first, I give a theoretical and mathematical summary of the classical and fuzzy Hough transforms, the gradient filtering and edge detection algorithms. Then, I introduce the practical application of the proposed method steps in detail. Finally, I apply the proposed method and evaluate and discuss the results.

### 3.2 *Theoretical and mathematical background of the implemented techniques*

#### 3.2.1 **Classical and fuzzy Hough transforms**

Hough transform has been in use for detecting straight lines and circles since a rather long time. The Hough transform is meant to find a parametric curve fitting to some measured points of the curve (Hough, 1959), if the shape, i.e., the general parametric formula of the curve is known. One of the simplest Hough transforms uses the two-parameter equation of a straight line, i.e.,

$$y = a_0 \cdot x + b_0 \quad (3.1)$$

with  $x$  and  $y$  being the coordinates of points in the Cartesian space, and  $a_0$  and  $b_0$  being the parameters from the parameter space  $(a, b)$ . Hough transform basically generates the curves belonging to each line point  $(x_0, y_0)$  in the space of the parameters; in the case of straight lines, each point  $(x_0, y_0)$  will form a straight line

$$b = (-x_0) \cdot a + y_0 \quad (3.2)$$

in the parameter space, too. If another point is on line (3.1) in the real space, it will have another line in the transformed space, similar to (3.2), but with different slope and offset. However, these lines, formed by the points on line (3.1) will have an intersection at  $(a_0, b_0)$ , i.e., at the parameter pair belonging to (3.1). This means, that the point  $(a_0, b_0)$  will be arising from all the points  $(x, y)$  of (3.1), thus, if we add one vote to each point of (3.1), then  $(a_0, b_0)$  will have a high number of votes, whereas other points, that belong only to one of the lines of type (3.2) will have only one. This consideration was used by Hough and later by many others to develop the following method.

1. Divide the space  $(x, y)$  by a finite grid (if not already executed).
2. Divide the transformed space  $(a, b)$  by a finite grid, it gives the resolution of the result.
3. For each point  $(a, b)$  in the transformed space add a vote for each point  $(x, y)$  in the original space, that is contained by the line.
4. Search for the maximum of votes
  - a) If there is just one line, the global maximum  $(a_{max,0}, b_{max,0})$  will be the approximation of parameter  $(a_0, b_0)$  of the line we were looking for.

b) If there are multiple lines, longer and shorter segments, the local maxima  $(a_{max,k}, b_{max,k})$  will also approximate parameter pairs of lines.

The longer the line in the original space, the more votes it obtains. By setting up a threshold, the length of the detected line segment can be controlled.

5. The lines  $y = a_{max,k} \cdot x + b_{max,k}$  with the detected approximate parameters  $(a_{max,k}, b_{max,k})$  can be drawn in the original space  $(x, y)$ .

The algorithm is given for the case of straight lines like (3.1), but it can be generated easily to any kind of parametric curve. In the case of colorectal polyps, the circular or elliptic Hough transform is the most plausible, with parametric equations

$$r_0^2 = (x - a_0)^2 + (y - b_0)^2, \quad (3.3)$$

$$1 = \frac{(x-h_0)^2}{a_0^2} + \frac{(y-k_0)^2}{b_0^2} \quad (3.4)$$

It can be seen that the circle has three parameters, the radius  $r_0$ , and the centre coordinates  $a_0$  and  $b_0$ , while the ellipse has four parameters, the half axes  $a_0$  and  $b_0$ , and the centre coordinates  $h_0$  and  $k_0$ .

Hough transforms are applied to images after edge detection, thus, the original space already has an innate grid, i.e., step 1 is mostly not necessary (Ballard, 1981), (Nahum, Eldar, & Bruckstein, 1991), (Lei & Erkki, 1993). Very often this edge detection step is considered as the first step of the Hough transform.

As the focus of our work is on circular Hough transform, here the pseudocode of the classical circular Hough transform is given in Algorithm 1.

In real life, the images have noise, thus, even if the original objects had straight lines, circles or ellipses as their edges, the images will probably have distorted edges. If this distortion is not too large, then the classical Hough transform is still effective, though sometimes more lines arise instead of one. However, there is a method that can handle not only slightly distorted edges, but larger deviations from the circles. This method is the fuzzy Hough transform, introduced by Han, Kóczy and Poston (Han, Kóczy, & Poston, 1994).

The fuzzy Hough transform considers the points  $(x, y)$  as fuzzy points. Fuzzy points are based on Zadeh's original idea of fuzzy sets (Zadeh, 1965), which generalized the classical, Boolean sets to have fuzzy perimeters by introducing a so-called membership function  $\mu$ . In the classical, Boolean algebra, only the membership values of  $\mu = 0$  and  $\mu = 1$  are possible, i.e., something can either be a member of a set or not. In Zadeh's approach, the objects can not only

**Algorithm 1: Classical Hough transform for a circle with parameters a, b and r****Requirements:**


---

an edge image  $I[i_x, i_y]$  with size  $L_x, L_y$ ,  
a finite parameter space  $V[j_a, j_b, j_r]$  with size  $L_a, L_b, L_r$  with initial values of 0  
a threshold for peak percentage  $P_p$ ,  
a result image  $R[i_x, i_y]$ , with size  $L_x, L_y$ , with initial values of 0

- 1: **for** each image row  $i_x$  from 1 to  $L_x$
- 2:     **for** each image column  $i_y$  from 1 to  $L_y$
- 3:         **for** each parameter space row  $j_a$  from 1 to  $L_a$
- 4:             **for** each parameter space column  $j_b$  from 1 to  $L_b$
- 5:                 **for** each parameter space 3rd dimension  $j_r$  from 1 to  $L_r$
- 6:                     **if**  $j_r^2 = (i_x - j_a)^2 + (i_y - j_b)^2$
- 7:                          $V[j_a, j_b, j_r] = V[j_a, j_b, j_r] + 1$
- 8:                     **end if**
- 9:                 **end for**
- 10:             **end for**
- 11:         **end for**
- 12:     **end for**
- 13: **end for**
- 14: compute the global maximum  $M_G$  in  $V[j_a, j_b, j_r]$
- 15: compute local maxima  $M(k) = V[j_{a,k}, j_{b,k}, j_{r,k}]$
- 16: select local maxima with  $M(k) > P_p \cdot M_G$
- 17: calculate the number  $N_M$  of the local maxima from line 16
- 18: **for** each local maximum  $k$  from 1 to  $N_M$
- 19:     **for** each result image row  $i_x$  from 1 to  $L_x$
- 20:         **for** each result image column  $i_y$  from 1 to  $L_y$
- 21:             **if**  $j_{r,k}^2 = (i_x - j_{a,k})^2 + (i_y - j_{b,k})^2$
- 22:                  $R[i_x, i_y] = 1$
- 23:             **end if**
- 24:         **end for**
- 25:     **end for**
- 26: **end for**

---

be elements or not of a set, but there is a strength of their membership. In the case of a geometrical point  $(x, y)$ , it can also be considered as a fuzzy set, it has a membership value  $\mu = 1$  at the coordinate point  $(x, y)$ , and the membership value decreases to zero monotonously as we get further from the coordinate  $(x, y)$ , leaving its environment still be partially belonging to the fuzzy point  $(x, y)$ . Using this approach in the fuzzy transform, we can consider not only the coordinate point  $(x, y)$  to give a vote 1 to the corresponding parameter space points, but also its environment can give votes proportional to their membership value.



This approach modifies the previously described classical Hough transform in the following way. (For the sake of generality,  $(a, b, \dots)$  were used, instead of  $(a, b)$  to express the applicability to any type of parametric curves, not only straight lines).

1. Divide the space of fuzzy points  $(x, y)$  by a finite grid (if not already executed).
2. Divide the transformed space  $(a, b, \dots)$  by a finite grid, it gives the resolution of the result.
3. For each point  $(a, b, \dots)$  in the transformed space, and its environment add a vote proportional to the membership function for each fuzzy point  $(x, y)$  in the original space, that is part of the fuzzy curve.
4. Search for the maximum of votes
  - a) If there is just one curve segment, the global maximum  $(a_{max,0}, b_{max,0}, \dots)$  will be the approximation of parameter  $(a_0, b_0, \dots)$  of the curve we were looking for.
  - b) If there are multiple curves, with longer and shorter segments, the local maxima  $(a_{max,k}, b_{max,k}, \dots)$  will also approximate parameters of the curves.

The longer the curve in the original space, the more votes it obtains. By setting up a threshold the length of the detected segment can be controlled.

5. The curves with the detected approximate parameters  $(a_{max,k}, b_{max,k}, \dots)$  can be drawn in the original space  $(x, y)$ .

Practically, if all fuzzy points  $(x, y)$  have the same membership distribution around them, then the 3rd step of the voting manifests in adding a vote  $\mu(|\alpha - a|, |\beta - b|, \dots)$  to the neighboring points  $(\alpha, \beta, \dots)$  of the studied parameter space point  $(a, b, \dots)$ .

Using this approach made the Hough transform more tolerant to distortions from the original parametric curves. The utilization of circular Hough transform for searching for contours of polyps was made possible by this, i.e., the three-parameter Equation (3.3) could be used instead of the four-parameter Equation (3.4).

Using these considerations, the pseudocode of the circular fuzzy Hough transform is as follows in Algorithm 2.

The applied voting membership function was a 3D-Gaussian, as it had a rather wide region around the center that is still close to 1. For a given parameter point  $(a_0, b_0, r_0)$ , the votes were given to the neighboring points  $([a_0 - \sigma, a_0 + \sigma], [b_0 - \sigma, b_0 + \sigma], [r_0 - \sigma, r_0 + \sigma])$  according to the membership function

**Algorithm 2: Fuzzy Hough transform for a circle with parameters a, b, and r****Requirements:**

- an edge image  $I[i_x, i_y]$  with size  $L_x, L_y$ ,
  - a finite parameter space  $V[j_a, j_b, j_r]$  with size  $L_a, L_b, L_r$ , with initial values of 0
  - a threshold for peak percentage  $P_p$ ,
  - a voting membership matrix  $\mu[j_a, j_b, j_r]$  with size  $2d_a + 1, 2d_b + 1, 2d_r + 1$
  - a result image  $R[i_x, i_y]$ , with size  $L_x, L_y$ , with initial values of 0
- 1: **for** each image row  $i_x$  from 1 to  $L_x$
  - 2:     **for** each image column  $i_y$  from 1 to  $L_y$
  - 3:         **for** each parameter space row  $j_a$  from 1 to  $L_a$
  - 4:             **for** each parameter space column  $j_b$  from 1 to  $L_b$
  - 5:                 **for** each parameter space 3rd dimension  $j_r$  from 1 to  $L_r$
  - 6:                     **if**  $j_r^2 = (i_x - j_a)^2 + (i_y - j_b)^2$
  - 7:                          $V[\{j_a - d_a, \dots, j_a + d_a\}, \{j_b - d_b, \dots, j_b + d_b\}, \{j_r - d_r, \dots, j_r + d_r\}] = \dots$
  - 8:                          $V[\{j_a - d_a, \dots, j_a + d_a\}, \{j_b - d_b, \dots, j_b + d_b\}, \{j_r - d_r, \dots, j_r + d_r\}] + \mu[j_a, j_b, j_r]$
  - 9:                     **end if**
  - 10:             **end for**
  - 11:         **end for**
  - 12:     **end for**
  - 13: **end for**
  - 14: compute the global maximum  $M_G$  in  $V[j_a, j_b, j_r]$
  - 15: compute local maxima  $M(k) = V[j_{a,k}, j_{b,k}, j_{r,k}]$
  - 16: select local maxima with  $M(k) > P_p \cdot M_G$
  - 17: calculate the number  $N_M$  of the local maxima from line 16
  - 18: **for** each local maximum  $k$  from 1 to  $N_M$
  - 19:     **for** each result image row  $i_x$  from 1 to  $L_x$
  - 20:         **for** each result image column  $i_y$  from 1 to  $L_y$
  - 21:             **if**  $j_{r,k}^2 = (i_x - j_{a,k})^2 + (i_y - j_{b,k})^2$
  - 22:                  $R[i_x, i_y] = 1$
  - 23:             **end if**
  - 24:         **end for**
  - 25:     **end for**
  - 26: **end for**

$$\mu(a, b, r) = \exp\left(-\frac{(a-a_0)^2 + (b-b_0)^2 + (r-r_0)^2}{2\sigma}\right). \quad (3.5)$$

As an example, an image (No. 220 from database CVC-Colon), its preprocessing steps, and transformed images are shown in the following figures. Figure 3-1 shows the image after reflection removal, its ground truth mask and its Prewitt edge detected version. Figure 3-2 shows the resulting votes for classical and fuzzy Hough transforms (for one with  $\sigma = 5$  slight

fuzziness, and another with  $\sigma = 15$ , wide fuzziness). Each column has four images, two at the radius values belonging to the two main detected circles, and there are 2 pictures with radii slightly smaller and larger than the circle at the polyp. Figure 3-3 shows the detected circles for different thresholds compared to the global maximum of the votes.

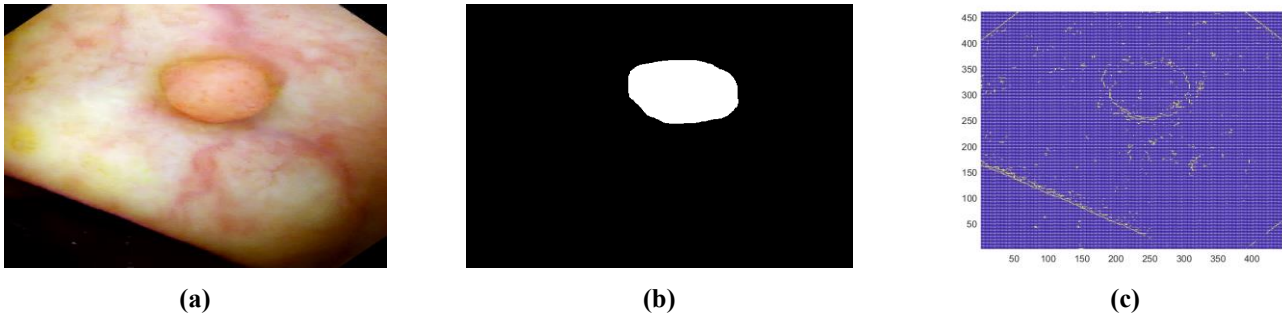


Figure 3-1 A sample image (220) from database CVC-Colon (Bernal, Sanchez, & Vilariño, 2012) for demonstrating the Hough transform steps, is shown in subplot (a), its ground truth mask in subplot (b), and the Prewitt edge detected version of the picture in subplot (c)

### 3.2.2 Gradient filtering

In mathematics, for a 2D continuous function, we use the partial derivatives to measure the degree of variation along each dimension. The edges in an image are segments that can be formed from the point locations where there is a rapid change in the image gray-level intensity in a small region. The connection between the previous two concepts made it possible to apply gradient filtering techniques in the field of image processing to detect edges.

The gradient of an image intensity function is a 2D vector with two components defined by the horizontal and vertical derivatives at each image point, and using these two values, we can identify the strength of the edge's magnitude and its orientation at each pixel.

The common mathematical formulation of the gradient for 2D image is the following vector:

$$\mathbf{G}[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}, \quad (3.6)$$

where  $f$  is the image intensity function, and  $x$  and  $y$  are the spatial coordinates of the image. The magnitude and direction of the gradient are consecutively given by the two equations below:

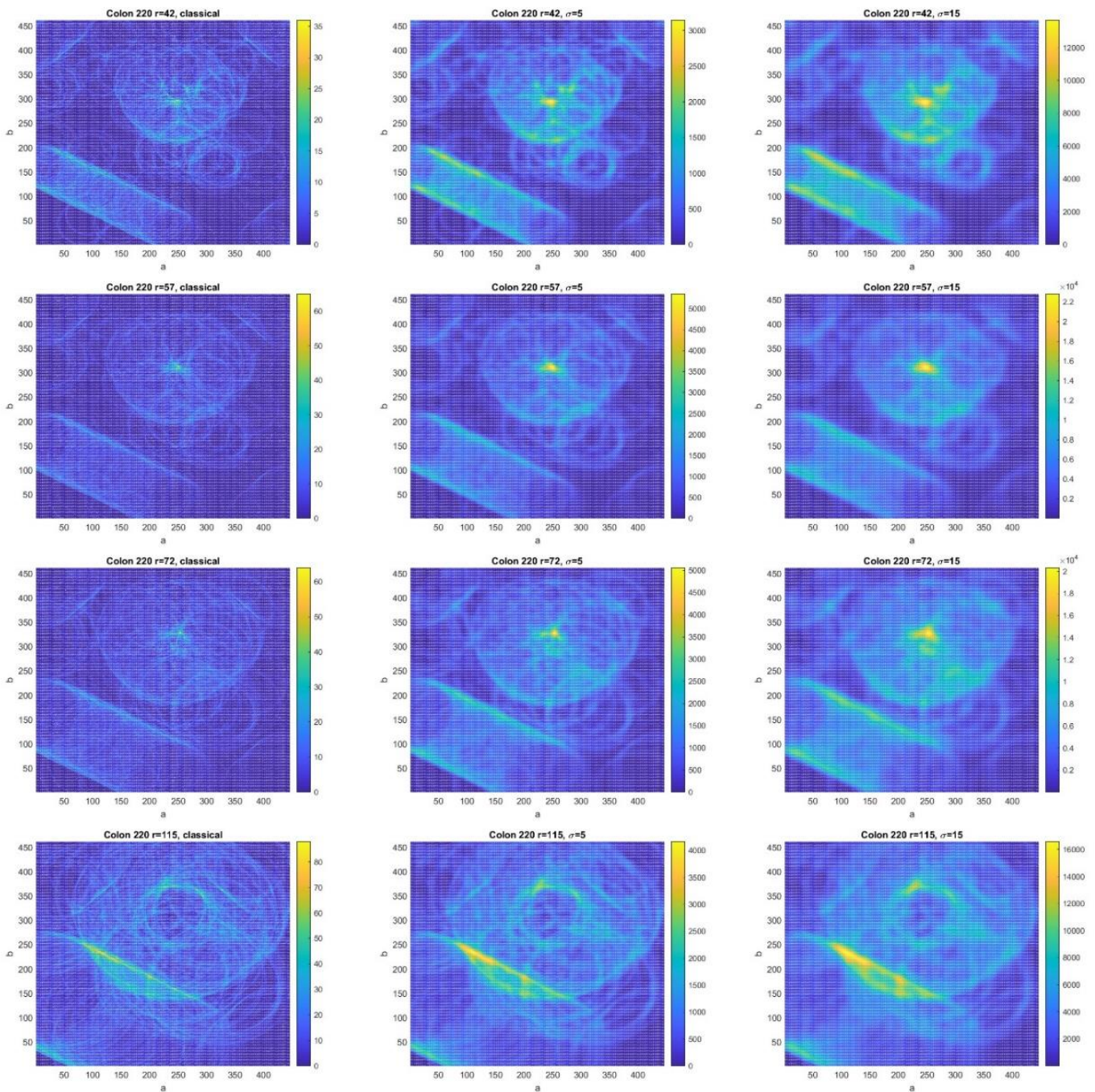
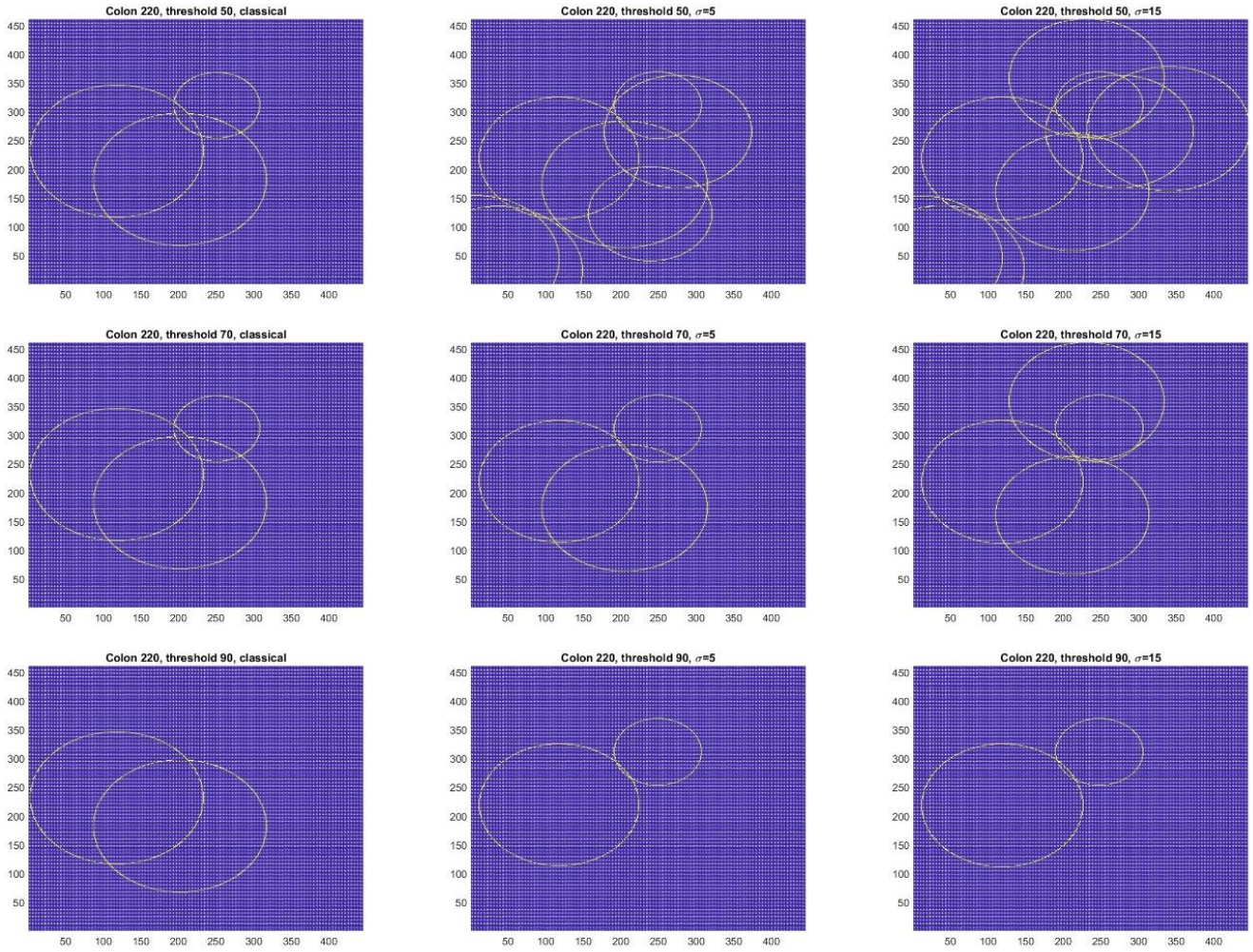


Figure 3-2 The Hough transformed images of the edge detected picture in subplot (c) of Figure 3-1. The center coordinates  $a$  and  $b$  are shown in one plane for 4 different values of the radius  $r$ . The number of votes is indicated by colors. It can be seen in the colorbar beside each image. The rows mean the different radii, the 1st row is  $r = 42$ , the 2nd row is  $r = 57$  (the radius of the final circle around the polyp), the 3rd row is  $r = 72$ , and the 4th, last row is  $r = 115$  (the radius of the other final circle). The columns mean the following: the 1st column: the classical Hough transform result, the 2nd column:  $\sigma = 5$  fuzzy Hough transform, and the 3rd column:  $\sigma = 15$  fuzzy Hough transform





*Figure 3-3 The resulting final circles from the inverse Hough transform of the picture in Figure 3-1. Similar to Figure 3-2, the columns indicate the fuzziness of the transform, i.e., the 1st column: classical Hough transform, the 2nd and 3rd columns: fuzzy Hough transform for  $\sigma = 5$  and 15, respectively. The rows in this case give the local maximum threshold compared to the global maximum. The 1st row: 50%, the 2nd row: 70%, and the 3rd row: 90% of the global maximum*

$$G[f(x, y)] = \sqrt{G_x^2 + G_y^2}, \quad (3.7)$$

$$\alpha(x, y) = \tan^{-1} \left( \frac{G_y}{G_x} \right). \quad (3.8)$$

All of the aforementioned considerations are carried out in the continuous domain. In the case of a digital image, where the intensity function is sampled at image discrete points, we replace the gradient operator by a discrete operation, i.e., by a convolution between the image and a kernel, which is a matrix of smaller size. For partial differentiation the discrete counterpart is taking the difference of neighboring pixels. The following gradient kernels are often used in practice.

$$Roberts_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad Roberts_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (3.9)$$

$$Prewitt_x = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad Prewitt_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}. \quad (3.10)$$

$$Sobel_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \quad Sobel_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (3.11)$$

### 3.2.3 Roberts, Prewitt, and Sobel edge detection algorithms

Edge detection methods can be used as mathematical techniques to identify particular locations in an image where the gray level intensities show discontinuities. The resulting edge maps serve as the basis for subsequent processing steps in numerous significant computer vision applications.

The well-known Roberts, Prewitt, and Sobel edge detection algorithms are widely used because of their simplicity and easiness of implementation. All of these algorithms have the same work mechanism, but with different kernels. Each kernel has the effect of calculating the gradient in the specified direction. However, the choice of algorithm to be used depends on the desired application and the characteristics of the image being processed.

Roberts edge detection method uses convolutional filters to detect the variations in the image gray-level intensity in the diagonal directions (Roberts, 1965), whereas Prewitt (Prewitt, 1970) and Sobel (Sobel, 1978) methods use convolutional matrices to detect the changes in both  $x$  and  $y$  directions.

The previously mentioned kernels, (3.9), (3.10), and (3.11), are used by the Roberts, Prewitt, and Sobel edge detection algorithms, consecutively.

### 3.2.4 Canny edge detection algorithm

John Canny first presented Canny edge detection in 1986 (Canny, 1986) as a multistep algorithm. Canny algorithm is looking for the connectivity of the edge points as well as the high gradient image points which makes it the most popular edge detection technique in many computer vision and image processing applications. This technique produces very reliable and highly accurate edge maps that are close to the human perception of edges.

The process of Canny algorithm consists of four main steps. First, the original image is refined using a Gaussian filter to remove unwanted noise. The applied Gaussian filter is defined as follows:

$$g(x, y) = G_{\sigma}(x, y) * f(x, y), \quad (3.12)$$

where

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \quad (3.13)$$

The convolution operator is represented by the symbol  $*$ , and the indices  $x$  and  $y$  are used to identify a pixel's location within an image. The two-dimensional function  $G_{\sigma}(x, y)$  is a Gaussian function with the variance of  $\sigma^2$ .

The smoothed image's gradient magnitude and direction are then calculated using a certain gradient operator, i.e., Roberts, Prewitt, or Sobel. The third step is implementing the Non-maximum Suppression (NMS) approach to check if the pixels are part of the local maxima, and if not, they are put down to zero. Two hysteresis high and low thresholds are computed in the final step. Every edge point with a gradient value greater than the higher threshold is identified as a strong edge, whereas the edge points whose gradient values fall below the lower threshold are eliminated. The connectivity of the residual edge points which have gradient values between the low and high thresholds is tested: the examined point is considered as an edge pixel only if at least one of the neighboring pixels is a strong edge pixel (Kalbasi & Nikmehr, 2020).

### 3.3 *Practical application of the proposed method*

#### 3.3.1 **Edge detection methods—application, evaluation, and selection**

For these tasks, several processing steps were carried out. They are sequentially summarized below together with the illustrated figures and plots. The publicly available colonoscopy image databases, CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014) were used.

1. Cutting off the black frame surrounding all original images to reduce unnecessary information.
2. Removing the colonoscope's light reflections: The colonoscope light's reflections (and consequently their contours) were removed from all the databases' images as a step towards reducing the number of redundant edge pixels (Figure 3-4, subplot (b)). The histogram of the image pixel intensities was used as the basis for the reflection removal step. Briefly, the

histogram's highest and lowest intensity peaks were cut out, and then the pixel intensities were re-normalized to the original [0 to 255] domain. A "white mask" was created using the pixels that made up the histogram's highest peak. Like the procedure described in (Csimadia & Nagy, 2014), the "white mask" was extended and smoothed into the neighboring pixels.

3. Extracting polyp contour: For each of the ground truth masks (Figure 3-4, subplot (c)), the contour was defined (Figure 3-4, subplot (d)). The number of pixels that make up the polyp mask contour was calculated for later use in the evaluation process.
4. Generating the "ring mask" for the colorectal polyp contour: In many cases of the manually drawn masks, the edges of the polyp's contour are not completely visible, either because the polyp is located in the area of bowel folds, or it is covered with impurities. Moreover, reasons related to human fatigue or error can also affect the drawing accuracy of the colorectal polyp mask. These are the main reasons why it was necessary to extend the contour of the manually drawn database mask to a finite width ring which is proportional to the size of the examined image.

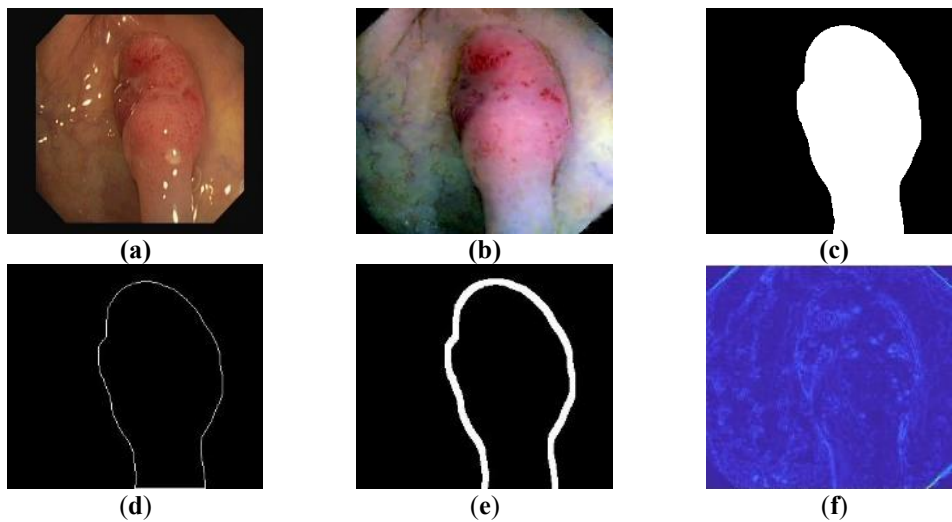
The previously extracted contour was extended into a ring (Figure 3-4, subplot (e)). To do that, the first  $x$  nearest pixels of the entire contour were selected as a width of the ring mask. As the databases images are with different sizes, different ring mask widths based on the database images size were used (they were  $x = 3$  for database CVC-Clinic (Bernal J. , et al., 2015),  $x = 5$  for database CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and  $x = 10$  for database ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014)).

5. Calculating the gradient magnitude for each of the studied samples, like in (Figure 3-4, subplot (f)).
6. Detecting polyp edges: Canny, Prewitt, Roberts, and Sobel techniques were applied as four different edge detection methods (Figure 3-5, subplots (a), (c), (e), (g) respectively). By employing this edge detection operation, it becomes possible to decrease the time required for the following pre-processing steps and offer a comparatively consistent data source that tolerates geometric and environmental variations while performing the Hough transform calculations. The total number of edge pixels resulting from each filtering technique for all the images of the three databases was calculated and plotted in Figure 3-6.
7. Finding the gradient-weighted edges: The edge filtered images (Figure 3-5, subplots (a), (c), (e), (g)) were multiplied by the gradient magnitude output (Figure 3-4, subplot (f)). The reason for performing this multiplication step is to determine the gradient magnitude domain, where the polyp edges are most likely to be present within the whole gradient



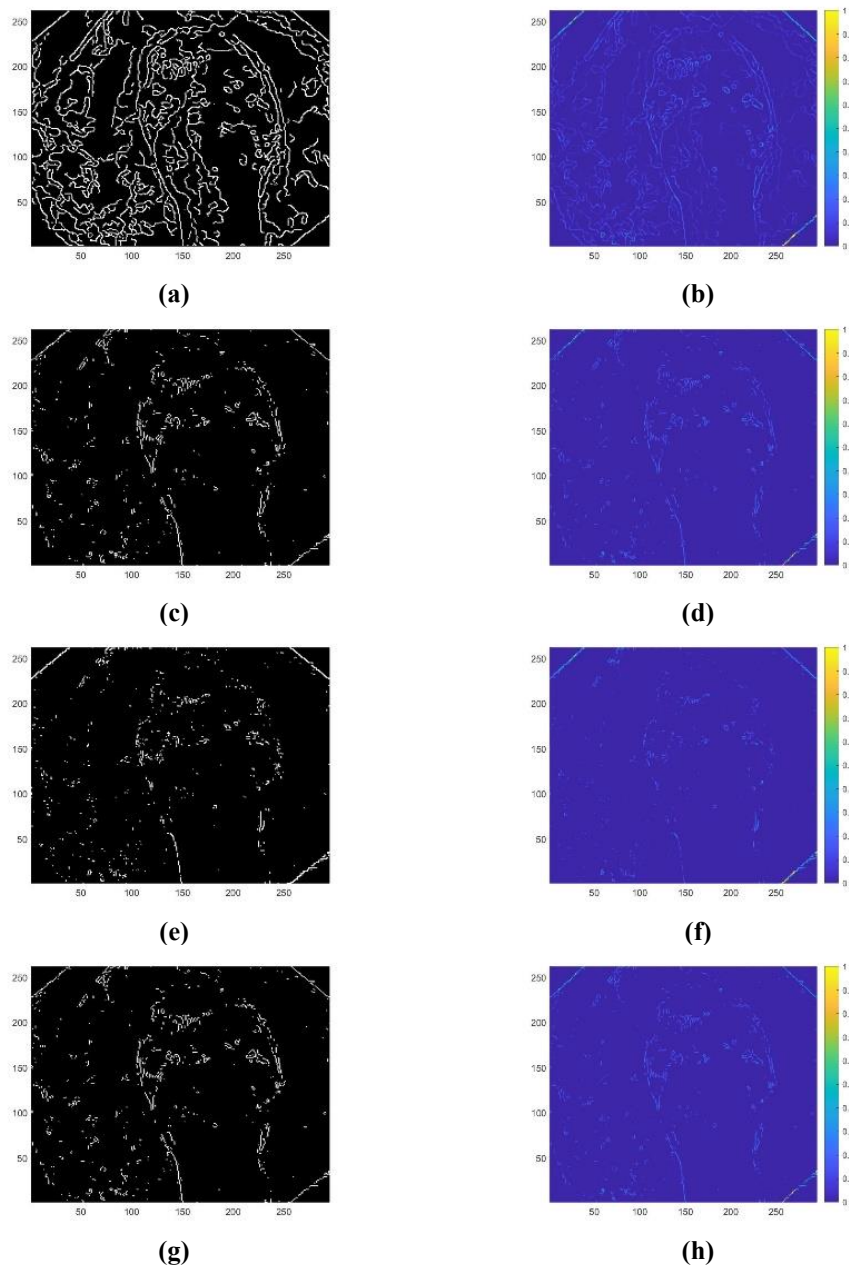
domain.

As an example, images like (Figure 3-5, subplots (b), (d), (f), (h)) for Canny, Prewitt, Roberts, and Sobel techniques were received. (It is visible that in contrast with the full gradient subplot (Figure 3-4, subplot (f)), subplots (b), (d), (f), (h) in Figure 3-5 contain the gradient values only where the edge mask value is 1, i.e., where the white pixels are located in subplots (a), (c), (e), (g) in Figure 3-5.



*Figure 3-4 First 5 steps of the edge detection algorithm selection process. 1st row: (a) The original image of the sample (111) from database CVC-Clinic (Bernal J. , et al., 2015), (b) the preprocessed image version, and (c) its ground truth polyp mask. 2nd row: (d) The mask contour, (e) the extended contour ring mask, i.e., the 3 nearest neighbors in all directions for all the contour pixels, and (f) the gradient magnitude filtered image*

8. Normalizing: To make the proposed approach universally applicable, for all the pictures in all the databases, the gradient-weighted edges pixels were normalized into the interval  $[0, 1]$  for each image separately.
9. Counting the number of the edge pixels located inside the ring mask, (this number serves as the reference: it is the number of the useful edge pixels).



*Figure 3-5 Steps 6 and 7 of the edge detection algorithm selection process. 1st row: (a) The Canny edge detected version of Figure 3-4 image, and (b) the Canny edge masked gradient magnitude image. The 2nd, 3rd, and 4th rows are exactly like the 1st row, but for Prewitt (c), (d), Roberts (e), (f), and Sobel (g), (h) edge detection methods consecutively*

10. Calculating the final evaluation metrics: Considering the application requirements, two quantities were introduced to evaluate each of the four implemented edge detection methods. For each image of the three studied databases, the statistics calculated in the previous steps (the total number of pixels in the polyp mask contour, the total number of edge pixels resulted from each edge detection method, and the total number of edge pixels

inside the ring mask) were used in composition the following two metrics.

- a. The first evaluation parameter, i.e., the one referring to the calculation efficiency, is defined by the ratio between the number of edge pixels in the ring mask around the polyp contour and the total number of edge pixels in the entire image,

$$R_{calc} = \frac{(No.of\ edge\ pixels\ in\ the\ ring\ mask)}{(Total\ No.of\ edge\ pixels)}. \quad (3.14)$$

This ratio represents the quality of the edge detection method regarding the Hough transforms and polyp detection.  $R_{calc}$  values' range is between 0 and 1. The higher this ratio is, the fewer non-mask contour edges are identified, and the less unneeded calculation is required in the classical or fuzzy Hough transform. Figure 3-7 shows the values of this metric resulting from Canny, Prewitt, Roberts, and Sobel filtering techniques for the three databases.

- b. The second evaluation parameter, i.e., the metric referring to the goodness of the edge pixels finding the ideal polyp contour (derived from the ground truth mask), is given by the ratio between the number of edge pixels in the ring mask and the number of pixels in the database polyp mask contour,

$$R_{edge} = \frac{(No.of\ edge\ pixels\ in\ the\ ring\ mask)}{(No.of\ pixels\ in\ mask\ contour)}. \quad (3.15)$$

In ideal case, this ratio should be as close to 1 as possible. Figure 3-8 displays the values of this metric resulting from all edge detection techniques for all studied databases.

We have to note that the detected edge pixels in the ring mask may not exactly be the same as the edge pixels in the manually drawn database mask contour, but they still could be used for finding the polyp.

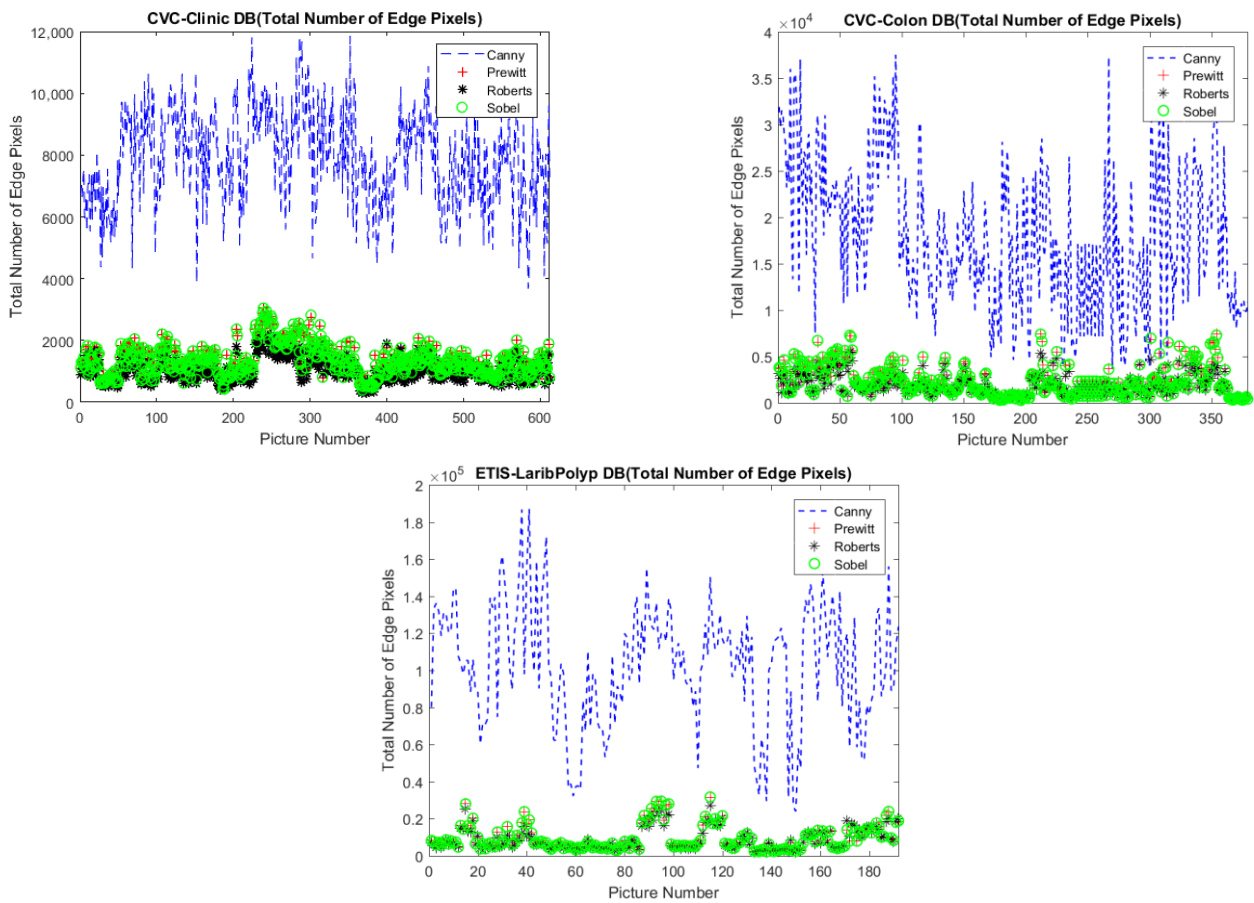


Figure 3-6 The total number of edge pixels resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases

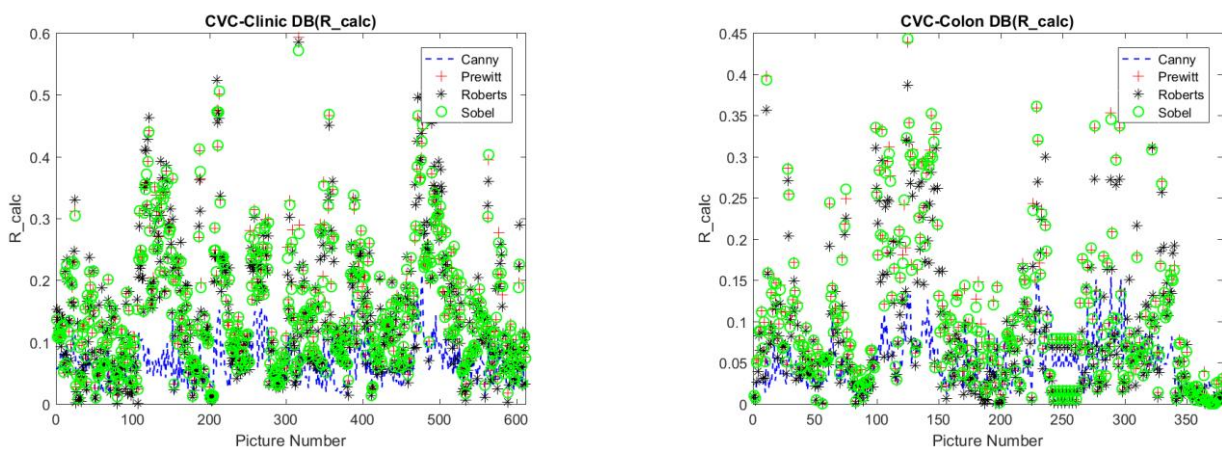


Figure 3-7 Metric  $R_{calc}$  values resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases. Databases CVC-Clinic (Bernal J. , et al., 2015) and CVC-Colon (Bernal, Sanchez, & Vilariño, 2012) are shown here

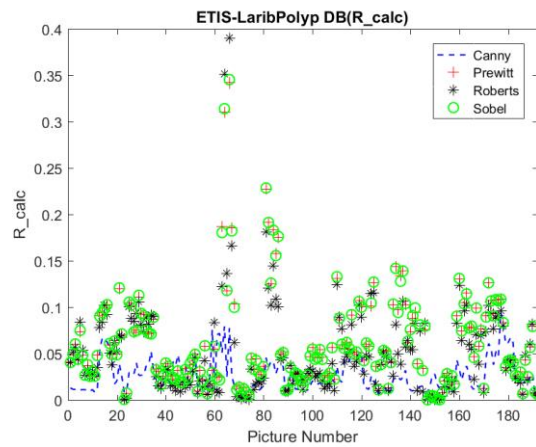


Figure 3-7 (continued): Metric  $R_{calc}$  values resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases. Database ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014) is shown here

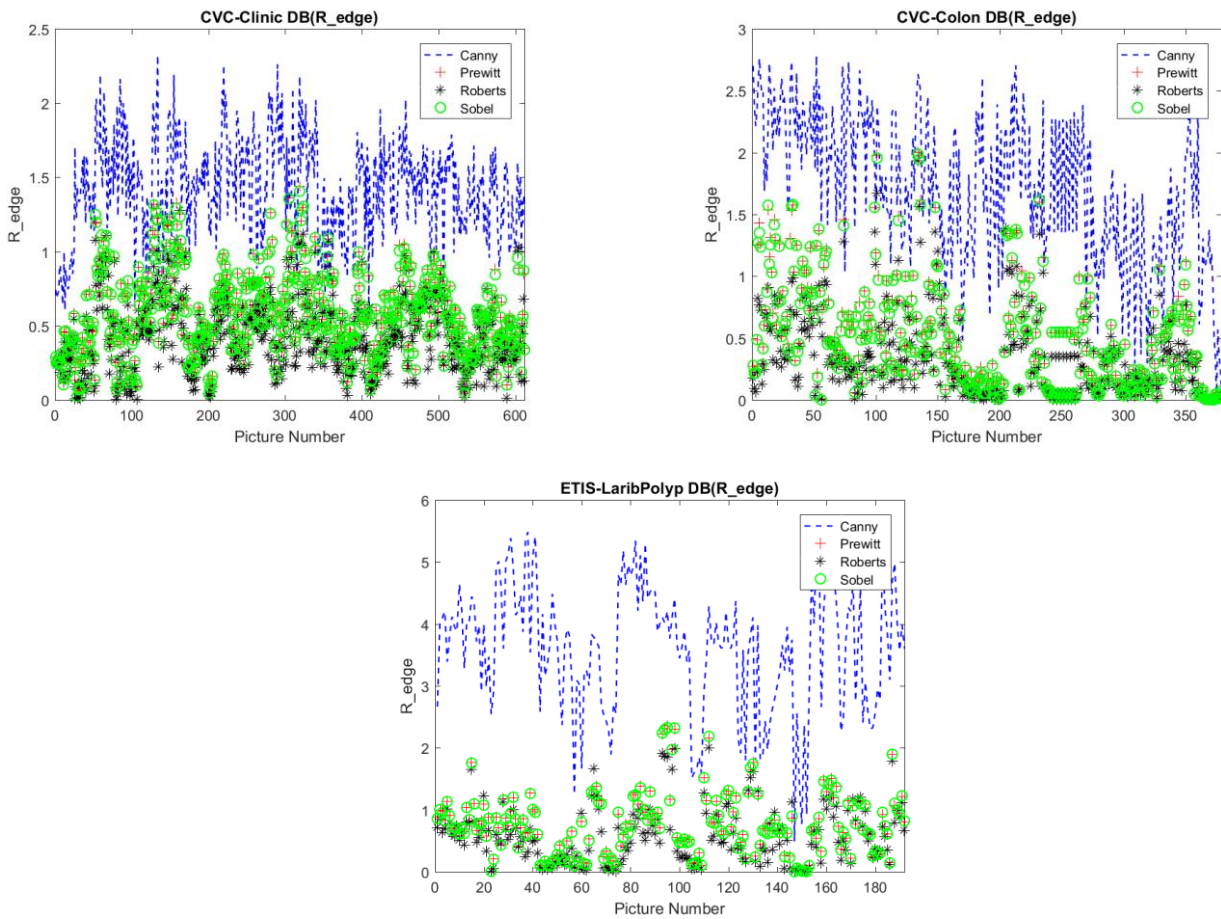


Figure 3-8 Metric  $R_{edge}$  values resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases

11. Selecting the most appropriate edge detection technique: Canny method detects a wide range of fine edges and gives a dense detailed edges map. It also tends to connect edge pixels to continuous edge lines, in contrast with the other three techniques. Figure 3-6 clearly shows the large difference between the total number of edge pixels resulting from Canny and the other three edge detection methods. Prewitt, Roberts, and Sobel have very similar results for the majority of samples. As the research focus is not just on decreasing the number of edge points scanned by the Hough transform, but also in increasing the efficiency of finding the colorectal polyp, the basis for selecting the most appropriate edge detection technique was derived from the definitions of the previous two evaluation metrics  $R_{calc}$  and  $R_{edge}$  as following.

- a. **Selection of the most appropriate edge detection technique using  $R_{calc}$ :** Two different selection strategies using metric  $R_{calc}$  were tested. Based on the definition of  $R_{calc}$ , the nearest the value  $R_{calc}$  to 1 is, the better the filter is.

According to the 1st strategy, the filter that has the mean of the metric  $R_{calc}$  values closest to 1 can be selected. For each database and each type of the four edge detection methods, the mean of the  $R_{calc}$  values was calculated.

However, for the values of metric  $R_{calc}$ , in all databases, for all the four edge detection techniques, there was many samples between  $[0, 0.1]$ , as it is visible in Figure 3-7. This is the reason why another strategy has to be considered as well.

According to the 2nd strategy, the filter which has the most samples close to the ideal value can be selected. For this purpose, a goodness interval can be defined, and the number of samples within that interval can be calculated. For every edge detection technique, the number of samples in each database that has a  $R_{calc}$  value greater than 0.1 was checked as a measure of the filter suitability. Accordingly, the higher the number of resulting samples, the better the filter is. Of course, the percentage of this number within the total number of images in each database must be considered; database CVC-Clinic has 612 images (Bernal J. , et al., 2015), database CVC-Colon has 379 images (Bernal, Sanchez, & Vilariño, 2012), and database ETIS-Larib has 196 images (Silva, Histace, Romain, Dray, & Granado, 2014). Table 3-2 lists the total results of this step.



Table 3-2 Results of selection the most appropriate edge detection technique using metric  $R_{calc}$

	<b>Canny</b>	<b>Prewitt</b>	<b>Roberts</b>	<b>Sobel</b>	
1st strategy: $R_{calc}$ Mean	0.063001	<b>0.153718</b>	0.147087	0.152823	CVC-Clinic
	0.04806	<b>0.099965</b>	0.081437	0.09925	CVC-Colon
	0.026186	<b>0.06001</b>	0.049776	0.059926	ETIS-Larib
2nd strategy: Num. of samples with $R_{calc} > 0.1$	70	<b>394</b>	360	<b>394</b>	CVC-Clinic
	29	138	108	<b>141</b>	CVC-Colon
	0	<b>35</b>	22	33	ETIS-Larib

- b. **Selection of the most appropriate edge detection technique using  $R_{edge}$ :** Two different choosing strategies based on metric  $R_{edge}$  were also tested.

According to the 1st strategy, similar to the test which was executed for  $R_{calc}$  values, a test of how similar the  $R_{edge}$  results to their ideal value can be employed. However, instead of calculating the mean value, the mean absolute error (MAE) value for metric  $R_{edge}$  from its ideal value 1 was calculated. For this criterion, the smallest MAE value nominates the better filter.

According to the 2nd strategy, as metric  $R_{edge}$  should be as close to 1 as possible, finding the goodness interval around  $R_{edge} = 1$  was suggested, and each sample that has  $R_{edge}$  value within  $[0.5, 1.5]$  was considered among the good samples. Consequently, the higher the number of samples within the goodness interval, the better the filter is. Table 3-3 arranges the total results of this step.

Table 3-3 Results of selection the most appropriate edge detection technique using metric  $R_{edge}$

	<b>Canny</b>	<b>Prewitt</b>	<b>Roberts</b>	<b>Sobel</b>	
1st strategy: $R_{edge}$ MAE	0.459919	<b>0.447193</b>	0.579451	0.44937	CVC-Clinic
	0.779827	<b>0.582246</b>	0.683746	0.588328	CVC-Colon
	2.609909	<b>0.44718</b>	0.536933	0.451125	ETIS-Larib
2nd strategy: Num. of samples with $0.5 < R_{edge} < 1.5$	346	<b>353</b>	221	349	CVC-Clinic
	132	<b>156</b>	90	152	CVC-Colon
	4	<b>122</b>	88	121	ETIS-Larib

The highlighted results presented in the above two tables show that the Prewitt filter is the most appropriate edge detection technique based on the proposed selection strategies using both metrics  $R_{calc}$  and  $R_{edge}$  in most cases.

It should be noted from Table 3-2 (2nd strategy) that the results of Sobel filter are very close to the Prewitt ones, especially for database CVC-Clinic: they are the same. Moreover, for database

CVC-Colon, the number of samples that had an  $R_{calc}$  value greater than 0.1 is three samples more in the Sobel filter than in the Prewitt case.

The flowchart in Figure 3-9 summarizes the overall edge detection method selection procedure.



Figure 3-9 Flowchart of the edge detection method selection procedure

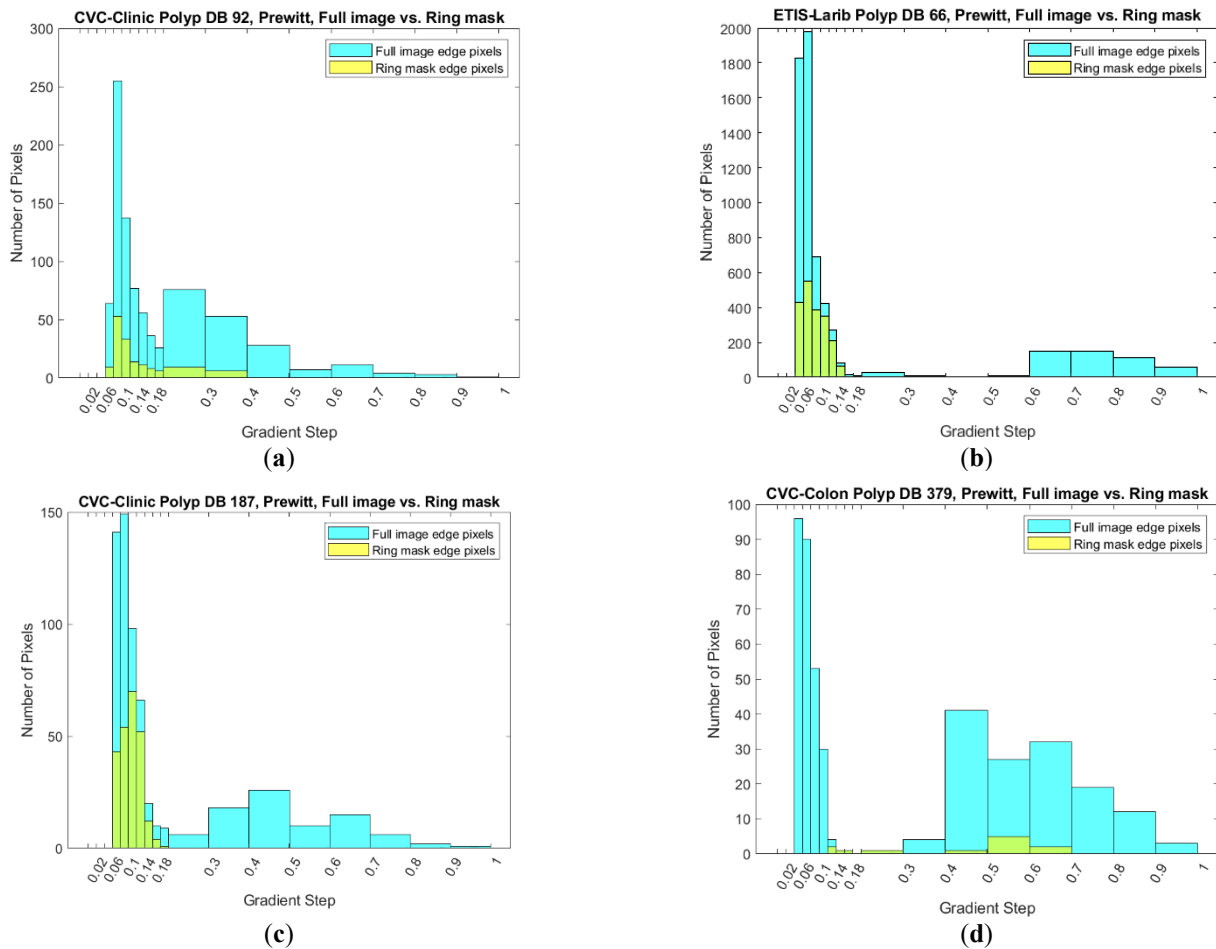
### 3.3.2 Gradient-based thresholding for Prewitt edge detection results

In this part of the proposed method, the dynamic distribution of the normalized gradient-weighted edges pixels resulting from the Prewitt method of both the ring mask (i.e., the area surrounding the ground truth mask contour) and the full image was studied by generating the individual histograms of all images in all databases.

Four individual histograms of four different samples are plotted in Figure 3-10. As in most of the cases the gradient intensities were in the first 20%, the density of the histogram bins in that domain were made to be larger. This dividing of the normalized gradient range resulted in



a more detailed tendency view in the domain with dense bin distribution. Additionally, as the column height in the dense bin distribution domain became relatively smaller, thus, the visibility of the small column magnitudes in the other parts of the histogram (i.e., in the higher normalized gradient domain) became better.

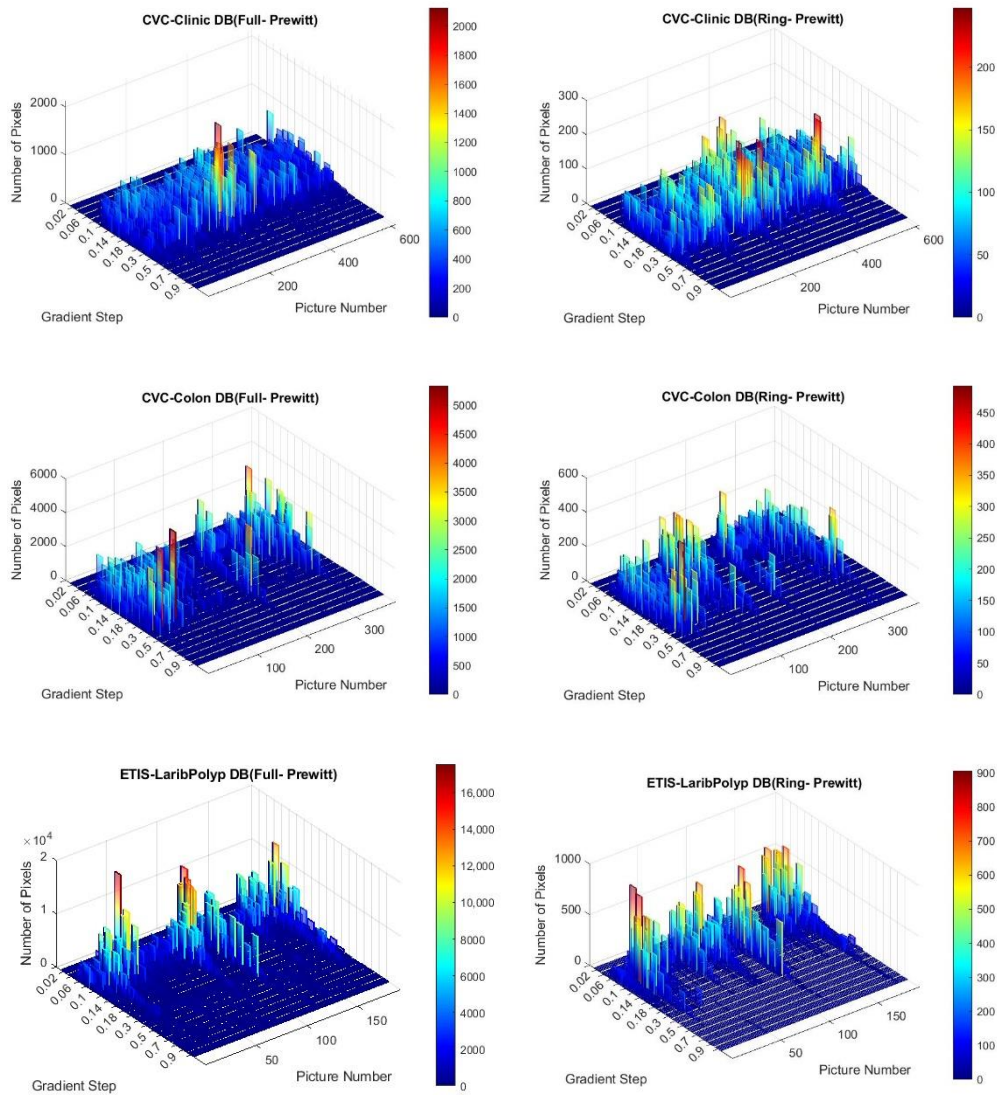


*Figure 3-10 Histograms of the normalized gradient-weighted edge pixels resulting from Prewitt method for four different samples (examples from Database CVC-Clinic (a), (c), CVC-Colon (d), and ETIS-Larib (b), are selected so, that typical behaviors would be visible). The interval  $[0, 0.2]$  has narrower bins, as most of the values in most of the samples were concentrated there, and higher resolution results were needed in this domain. (The yellow histograms for the ring masks are semitransparent, plotted in front of the teal columns for the full images)*

In most of the cases, the distributions of the full image edges (cyan) and the ring mask edges (yellow) showed a very similar tendency with different magnitude, like in the case of the first example, subplot (a) of Figure 3-10. Very often besides the similar tendency in the low gradient domain, the higher gradient parts were missing in the case of the ring mask edges, like

in the 2nd and 3rd examples (subplots (b) and (c)). In some other cases, the distributions of the full picture and the ring mask had different tendencies, like in the case of the 4th example, in subplot (d).

To summarize the results of the individual histograms, 3D-histograms were created for the three databases with the picture number being the 3rd dimension. These total histograms are plotted in Figures 3-11 and 3-12.



*Figure 3-11 The total histograms (linear scale) of the normalized gradient-weighted edges pixels resulted from Prewitt edge detection method. The 1st column is for the edge points in the full images, and the 2nd column is for the edge points in the corresponding ring masks, for the databases CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014). The horizontal axes are the normalized gradient magnitudes intervals and the picture number in the given database*

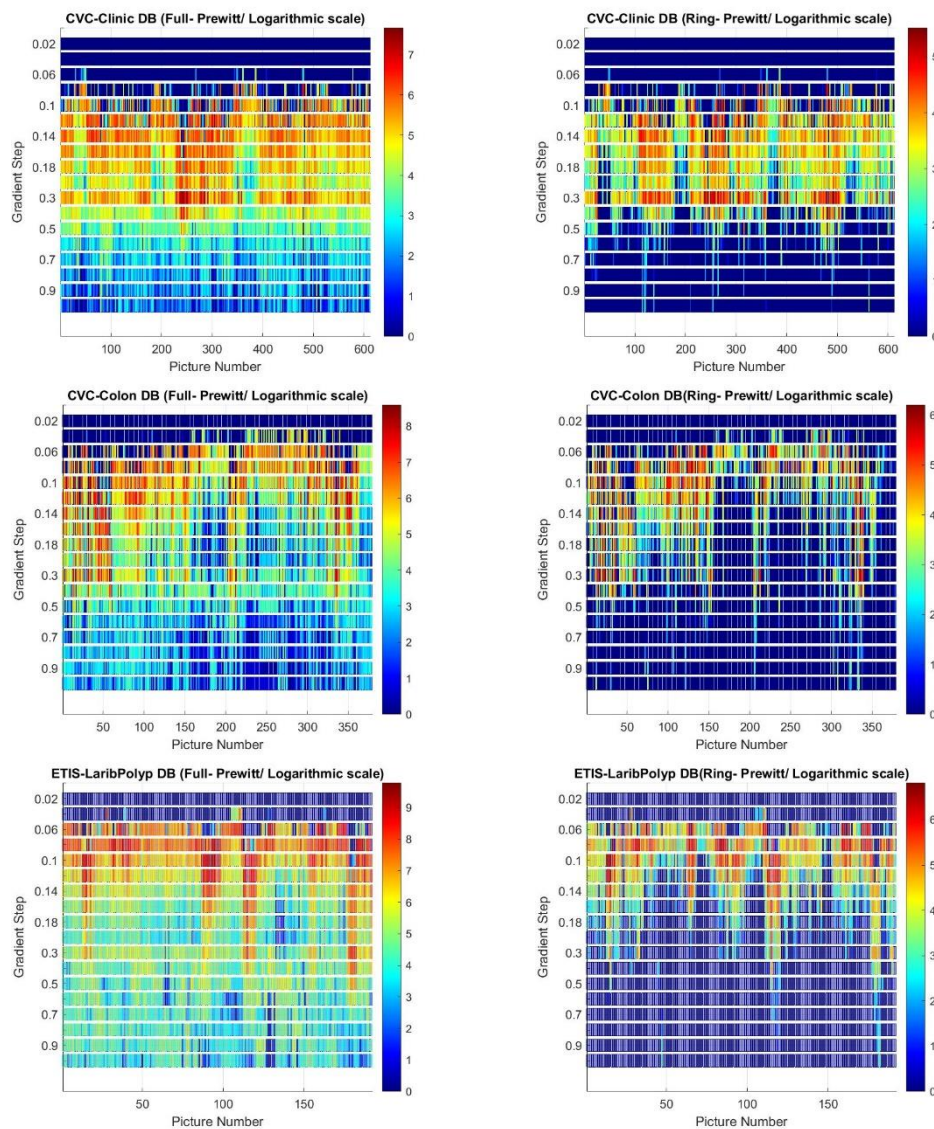


Figure 3-12 The top view of the logarithmic scale plots for the total histograms of the normalized gradient-weighted edges pixels resulted from Prewitt edge detection method. The 1st column is for the edge points in the full images, and the 2nd column is for the edge points in the corresponding ring masks, for the databases CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilarino, 2012), and ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014). The horizontal axis is the picture number in the given database and the vertical axis is the normalized gradient magnitudes intervals

Figure 3-11 shows the perspective view in linear scale, while Figure 3-12 gives the top view of the same histograms in logarithmic scale. It is worth mentioning that the linear and logarithmic scale plots of the same histogram are both given, because the smaller valued histogram parts at the lowest and the highest normalized gradient bins cannot be observed well in the linear scale plots of the total histograms. On the other hand, the top view of the

logarithmic scale plots makes the entire set of data viewable without columns blocking the ones behind them.

Based on results in Figures 3-11 and 3-12, it is visible that the edges with very high gradient values do not belong to the polyp edge, i.e., most of the ring mask edge results have 0 values over 0.3. Additionally, they have zero values below 0.06. Moreover, in many cases this lower limit can be brought up to 0.08, and the upper limit can be as low as 0.2. This property was used, and the edges with too high and too low gradient magnitudes were omitted to both reduce the number of pixels to be Hough transformed and to sort out those pixels that certainly do not belong to the polyp edges. Thus, in the next calculations, beside the full Hough transforms for all the edges, a restricted Hough transform for the edges with normalized gradient values within a wide threshold interval  $[0.06, 0.3]$  and a thin threshold interval  $[0.08, 0.2]$  was performed. The full Hough transform results can be interpreted as a reference, to analyze how the restricted transforms influence the results.

### **3.3.3 Gradient-weighted voting approach for classical and fuzzy Hough transforms**

In addition to excluding the edge points with too high or too low contrast (i.e., gradient thresholding), the gradient values can be used for another purpose, namely, to modify the voting process.

Even after the thresholding, polyp contours usually have segments with higher gradient magnitudes, mixed with lower gradients, whereas some of the background patterns have very low gradients. To decrease the influence of these lower gradient background edges on the Hough transform results, the following weighted voting approach was introduced.

During the original Hough transform, all edge points receive the same vote, no matter how strong these edges are. In the proposed method, instead of 1, each point uses its normalized gradient magnitude as a vote. In the case of the fuzzy Hough transforms, the whole voting membership function is multiplied by the gradient magnitude of the given pixel. As a result, the smaller the step in the intensity at the edge is, the smaller the weight in the voting becomes for the edge point, no matter if it is a classical or a fuzzy point. To clarify better, the pseudocode of the fuzzy circular Hough transform with gradient-weighted voting approach is given in Algorithm 3.

The performance of the proposed voting technique was evaluated using selected samples from each database for three main cases. The first case was the original case where the full Hough transforms (classical and fuzzy) for all the edges without any thresholding were applied.

**Algorithm 3: Fuzzy Hough transform for a circle with parameters a, b, and r with gradient-weighted voting approach****Requirements:**

- an edge image  $I[i_x, i_y]$  with size  $L_x, L_y$
  - a gradient magnitude image  $G[i_x, i_y]$  with size  $L_x, L_y$
  - a finite parameter space  $V[j_a, j_b, j_r]$  with size  $L_a, L_b, L_r$  with initial values of 0
  - a threshold for peak percentage  $P_p$ ,
  - a threshold interval for the gradient magnitudes  $[g_{min}, g_{max}]$
  - a voting membership matrix  $\mu[j_a, j_b, j_r]$  with size  $2d_a + 1, 2d_b + 1, 2d_r + 1$
  - a result image  $R[i_x, i_y]$ , with size  $L_x, L_y$ , with initial values of 0
- 1: **for** each image row  $i_x$  from 1 to  $L_x$
  - 2:     **for** each image column  $i_y$  from 1 to  $L_y$
  - 3:         **for** each parameter space row  $j_a$  from 1 to  $L_a$
  - 4:             **for** each parameter space column  $j_b$  from 1 to  $L_b$
  - 5:                 **for** each parameter space 3rd dimension  $j_r$  from 1 to  $L_r$
  - 6:                     **if**  $j_r^2 = (i_x - j_a)^2 + (i_y - j_b)^2$  **and**  $g_{min} \leq G[i_x, i_y] \leq g_{max}$   
 $V[\{j_a - d_a, \dots, j_a + d_a\}, \{j_b - d_b, \dots, j_b + d_b\}, \{j_r -$
  - 7:                      $d_r, \dots, j_r + d_r\}] = \dots$   
 $V[\{j_a - d_a, \dots, j_a + d_a\}, \{j_b - d_b, \dots, j_b + d_b\}, \{j_r -$   
 $d_r, \dots, j_r + d_r\}] + G[i_x, i_y] \cdot \mu[j_a, j_b, j_r]$
  - 8:                     **end if**
  - 9:                 **end for**
  - 10:             **end for**
  - 11:         **end for**
  - 12:     **end for**
  - 13: **end for**
  - 14: compute the global maximum  $M_G$  in  $V[j_a, j_b, j_r]$
  - 15: compute local maxima  $M(k) = V[j_{a,k}, j_{b,k}, j_{r,k}]$
  - 16: select local maxima with  $M(k) > P_p \cdot M_G$
  - 17: calculate the number  $N_M$  of the local maxima from line 16
  - 18: **for** each local maximum  $k$  from 1 to  $N_M$
  - 19:     **for** each result image row  $i_x$  from 1 to  $L_x$
  - 20:         **for** each result image column  $i_y$  from 1 to  $L_y$
  - 21:             **if**  $j_{r,k}^2 = (i_x - j_{a,k})^2 + (i_y - j_{b,k})^2$
  - 22:                  $R[i_x, i_y] = 1$
  - 23:             **end if**
  - 24:         **end for**
  - 25:     **end for**
  - 26: **end for**

The traditional voting technique was followed, where each edge point obtains one vote if it is eligible. The other two cases were the restricted Hough transforms for the edges with normalized gradient values within the wide threshold interval [0.06, 0.3] and the thin threshold

interval [0.08, 0.2]. For these two cases, the modified, gradient-weighted voting technique was used.

In addition to the classical Hough transform, the fuzzy Hough transform was also studied in each case, with  $\sigma = 3, 5, 7, 9, 11, 13$  and 15 from the voting membership function (3.5). (In order to simplify the referring to these transform types, the classical transform may be considered as a fuzziness parameter  $\sigma = 0$ ).

When selecting the final circles, i.e., the local maxima in the transformed image, a threshold  $P_p$  is to be set within the Hough transforms. The arising circles (their number and location) were studied for different local maximum thresholds, namely for  $P_p = 50\%, 60\%, 70\%, 80\%$ , and 90% of the global maximum of the votes.

The goal of this evaluation step is to determine, how many of the final circles are around the polyp for the various  $\sigma$ s and local maximum thresholds  $P_p$ s. For this purpose, the total number of circles  $N_{total}$  was counted. A circle was considered to be around the polyp by testing, if its center was inside the ground truth mask, and it had points within the ring mask around the polyp contour. The number of such circles within the ring mask  $N_{ring}$  was also counted. The ratio

$$A_r = \frac{N_{ring}}{N_{total}} \quad (3.16)$$

was used as a metric for effectiveness of finding circles related to the polyp. If  $A_r$  is 0, then the polyp is not found, if  $A_r$  is too small, then too many other circles are found, and if  $A_r$  is around 1, then the circle(s) around the polyp are found, but not many other circles can be seen in the final results.

To test the different types of Hough transforms' tolerance degree to the deviation from the circle, it is necessary to know the size of the polyps as well as their roundness. To calculate the average radius of the polyp  $r_{avg}$ , the maximum and minimum coordinates of the polyp mask in  $x$  and  $y$  direction were used as follows,

$$r_{avg} = \frac{(x_{max}-x_{min})+(y_{max}-y_{min})}{2}. \quad (3.17)$$

The center of the polyp was determined as the average of these coordinates,

$$(c_x, c_y) = \left( \frac{(x_{max}+x_{min})}{2}, \frac{(y_{max}+y_{min})}{2} \right). \quad (3.18)$$

This center point and the average radius were used to determine the radial displacement for each mask contour point  $(m_x, m_y)$  by

$$\Delta_{radial} = \left| \sqrt{(m_x - c_x)^2 + (m_y - c_y)^2} - r_{avg} \right|. \quad (3.19)$$

The roundness error can be measured by the maximum of the ratio of the radial displacements and the average radius:

$$\delta_{roundness} = \max \left( \frac{\Delta_{radial}}{r_{avg}} \right). \quad (3.20)$$

The larger the roundness error  $\delta_{roundness}$ , the more the shape deviates from circle. Please note, the roundness error of a circle with diameter of 30 pixels can still be as large as  $\delta_{roundness,30} = 0.04$ , which decreases for a circle twice as large (i.e., of diameter 60 pixels) to  $\delta_{roundness,60} = 0.03$ . Typically, an ellipse with large axis being double the size of the smaller axis have roundness error of about  $\delta_{roundness} \approx 0.33$ , while with axis ratio 1:3, this increases to  $\delta_{roundness} \approx 0.45$ , depending on the size of the ellipse in pixels.

### 3.4 Results and discussion

#### 3.4.1 Number of circles found by the algorithm

In this section, the values of  $A_r$  by  $\sigma$ s and  $P_p$ s are given for databases CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014). The total number of circles found on the images  $N_{total}$  is also important, because as few circles, and thus, ROIs as possible are needed for the steps after the Hough transform. For this reason,  $N_{total}$  is also plotted for the three databases. In Figures 3-13, and 3-14, the values  $A_r$  and  $N_{total}$  are given for CVC-Clinic, then in Figures 3-15 and 3-16 for CVC-Colon, and finally for ETIS-Larib in Figures 3-17 and 3-18. Figures 3-14, 3-16 and 3-18 together with Figures 3-13, 3-15 and 3-17, completely cover the number of circles found in the complete image and in the ring masks.

The test images were selected so that their  $R_{calc}$  values would be larger than 0.1. Both roundish and elongated polyps were selected from all three databases. The images with their ground truth masks are given in Appendix C. The image numbers and their respective  $R_{calc}$  and  $R_{edge}$  values together with the total number of Prewitt edge pixels for each tested sample can be found in Table 3-4. To make the chapter more accessible to the reader, abbreviations and nomenclature section of the proposed approach is also presented in Appendix D.

In order to check the applicability of the proposed method, some images (two from each database) that have unfavorable, very small  $R_{calc}$  values were selected. These “bad” images



are shown as the first two images of the figures illustrating the tests samples in Appendix C, as well as the first two lines of Table 3-4. These images all found zero circles in the ring mask, except for one case CVC-Colon 255, for the original, non-gradient-weighted voting method. This is why their results are not shown in Figures {3-13, to 3-18}.

In Figures 3-13, 3-15 and 3-17 the values of the ratio  $A_r = N_{ring}/N_{total}$  are shown for databases CVC-Clinic (Bernal J. , et al., 2015), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and ETIS-Larib (Silva, Histace, Romain, Dray, & Granado, 2014), respectively.

In Figures {3-13, to 3-18}, each plot has three segments along the vertical axis. The uppermost segment shows the original Hough transform's  $A_r$  values. These are the reference values to which the other results should be compared. The next, middle segment corresponds to the gradient-weighted voting with wider threshold, and the lowermost segment belongs to the gradient-weighted voting with a thinner threshold. Each line in each segment belongs to one image, similar to the order of Table 3-4 (except for the first two lines in each database, i.e., the "bad" images).

The plots have segments along the horizontal axes, too: the upper subplot has the results grouped by the peak percentage  $P_p$  (i.e., threshold for the local maxima of the voting map  $(a, b, r)$  to be considered as the circles), and each segment lists the  $\sigma$ s from classical (i.e.,  $\sigma = 0$ ) till the widest fuzziness parameter  $\sigma = 15$ . The lower subplot orders the values the opposite way: the segments belong to various fuzziness parameters  $\sigma$ , while the columns within the segments belong to the peak percentages  $P_p = 50\%, 60\%, 70\%, 80\%$ , and  $90\%$ . The colored squares mean the number of circles found in the ring mask versus the total number of circles for a given image, for a given  $(P_p, \sigma)$  pair. The darker red the little square is, the fewer circles outside of the polyp contour domain are found. If the little square is bright green, then no circle is found in the polyp contour's ring mask.

From Figure 3-13 we can conclude that for database CVC-Clinic the fuzzy Hough transforms can mostly find circles belonging to polyp contours (most of the little squares in the maps are red).

The classical transform's results are either very good (dark red) or very bad (green).

Mostly, increasing the threshold improves the results, the circles will be more in the area of the polyp contour, but the 90% threshold might be too much. It loses the circles in the ring mask for some cases. Thus, the most ideal value for  $P_p$  is 70–80% of the global maximum of the votes.



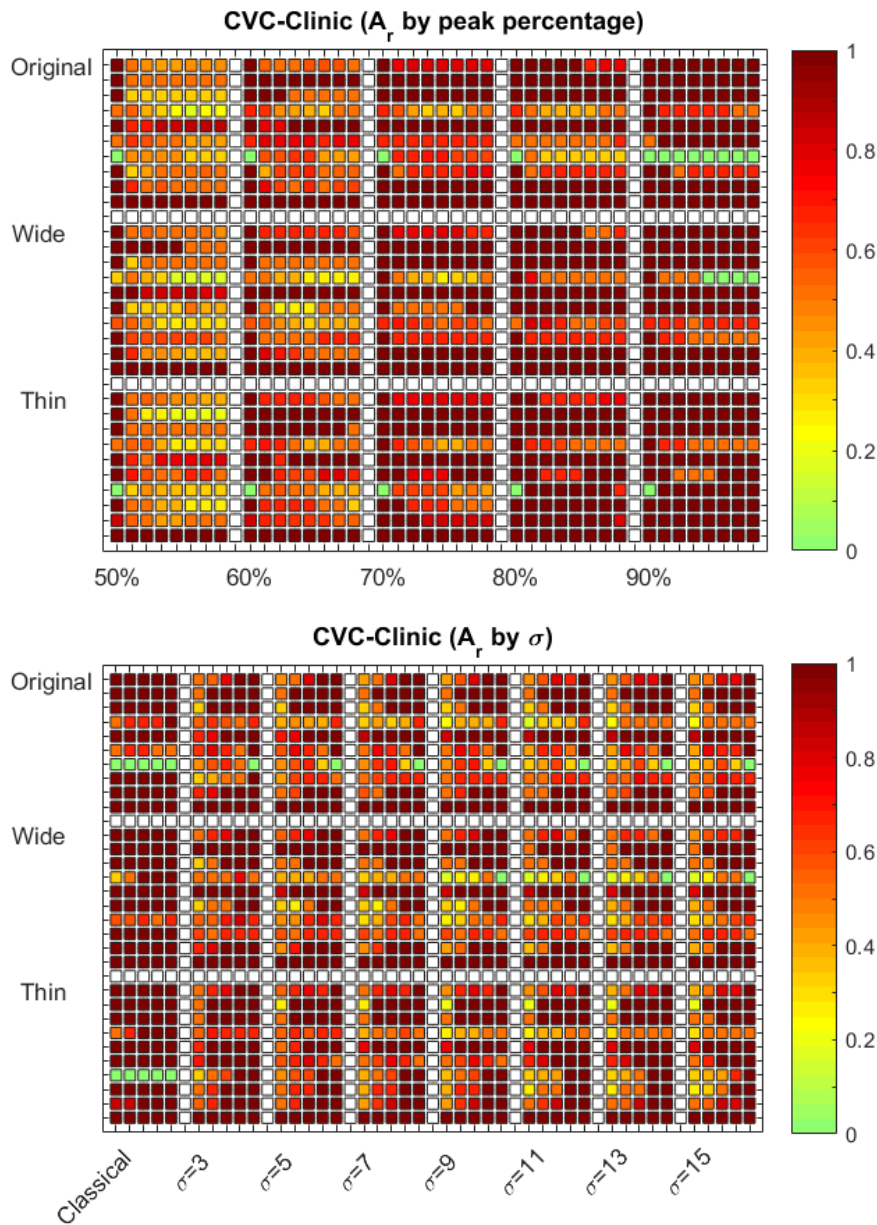


Figure 3-13 The ratio  $A_r$  for the database CVC-Clinic (Bernal J. , et al., 2015). The vertical axis represents the various sample pictures (images 111, 150, 188, 217, 265, 390, 475, 480, 503, and 504) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. Each section has all the sample images in the above given order, and there is a line of white points to indicate the borders between the sections. The horizontal axes have the local maximum threshold percentages  $P_p$ , as well as the various fuzziness parameters  $\sigma$  of the Hough transform. The 1st subplot shows the different values of  $\sigma$  for each local maximum threshold level in increasing order. Here also a column of white dots separates the sections belonging to the given threshold rates. The 2nd subplot gives the same results as the 1st one, only the horizontal axis is grouped the opposite way: the main groups belong to the various  $\sigma$  values, and within each segment the local maximum threshold values increase

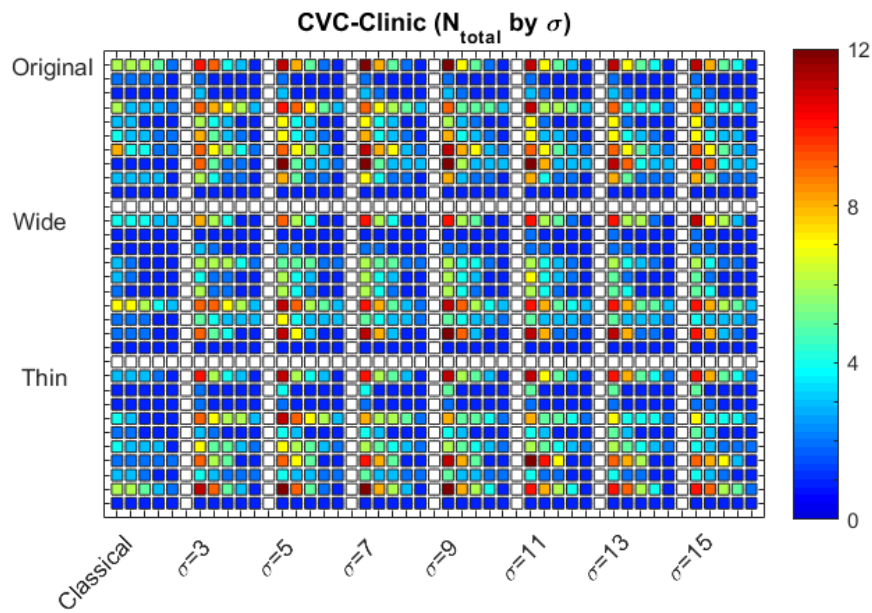


Figure 3-14 The total number of circles  $N_{total}$  found in the sample images of database CVC-Clinic (Bernal J. , et al., 2015). The vertical axis represents the various sample pictures (images 111, 150, 188, 217, 265, 390, 475, 480, 503, and 504) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. The horizontal axis is grouped the following way: the main groups belong to the various  $\sigma$  values, and within each segment the local maximum threshold values increase

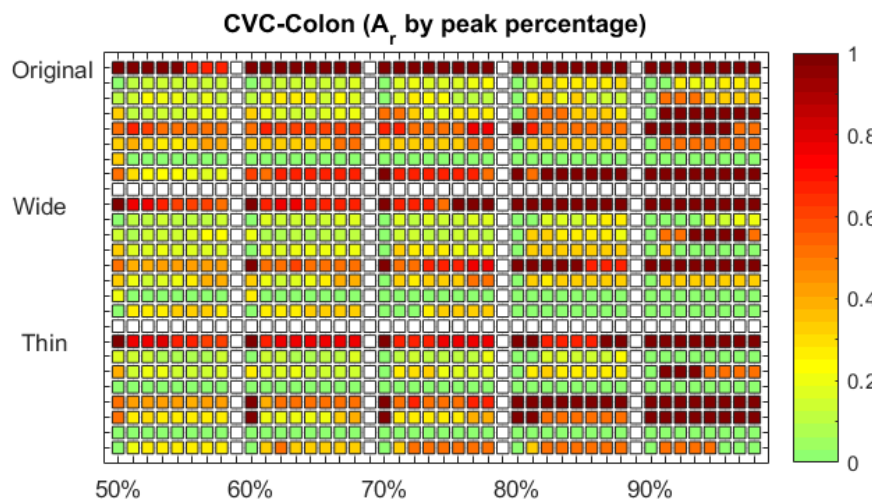


Figure 3-15 The ratio  $A_r$  for the database CVC-Colon (Bernal, Sanchez, & Vilariño, 2012). The vertical axis represents the various sample pictures (images 62, 74, 101, 128, 149, 220, 230, and 283) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. The horizontal axis has the local maximum threshold percentages  $P_p$ , similar to first plot in Figure 3-13

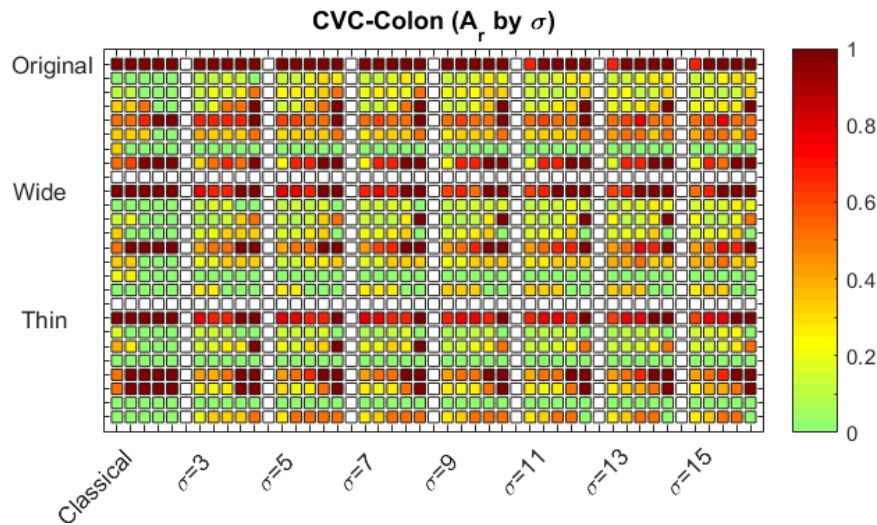


Figure 3-15 (continued): The ratio  $A_r$  for the database CVC-Colon (Bernal, Sanchez, & Vilariño, 2012). The vertical axis represents the various sample pictures (images 62, 74, 101, 128, 149, 220, 230, and 283) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. The horizontal axis has the various fuzziness parameters  $\sigma$  of the Hough transform, similar to second plot in Figure 3-13

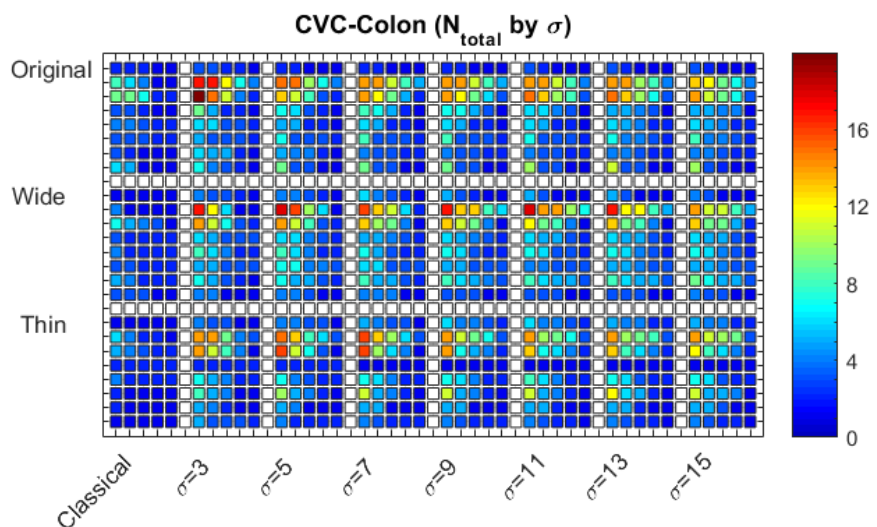


Figure 3-16 The total number of circles  $N_{total}$  found in the sample images of database CVC-Colon (Bernal, Sanchez, & Vilariño, 2012). The vertical axis represents the various sample pictures (images 62, 74, 101, 128, 149, 220, 230, and 283) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. The horizontal axis is grouped similarly to Figure 3-14

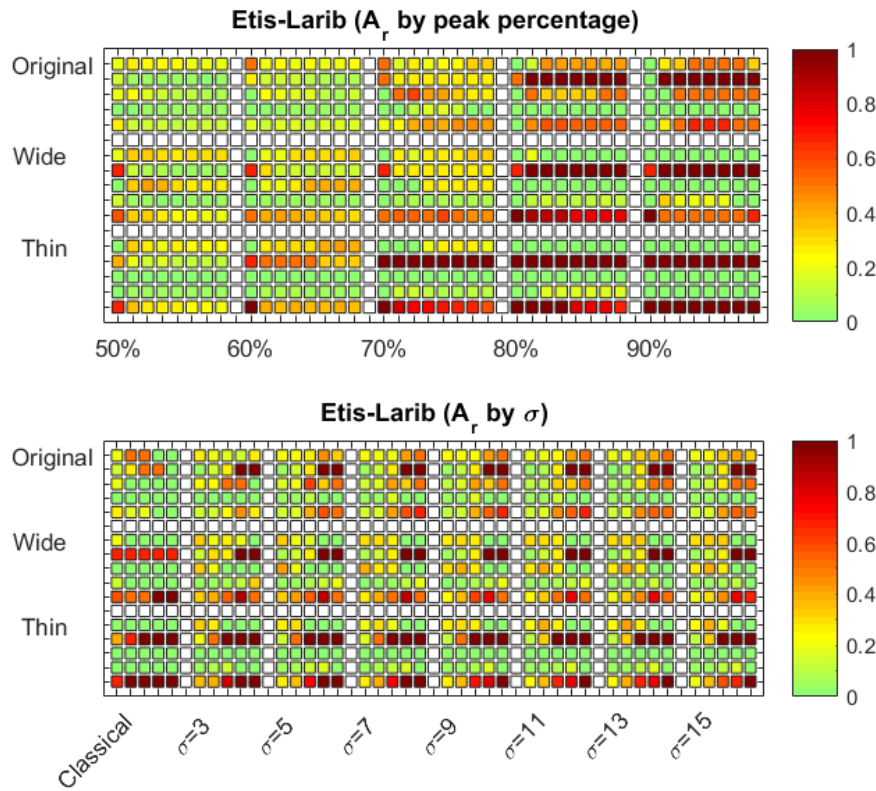


Figure 3-17 The ratio  $A_r$  for the database *ETIS-Larib* (Silva, Histace, Romain, Dray, & Granado, 2014). The vertical axis represents the various sample pictures (images 25, 65, 82, 138, and 160) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. The horizontal axes have the local maximum threshold percentages  $P_p$ , as well as the various fuzziness parameters  $\sigma$  of the Hough transform, similar to Figures 3-13 and 3-15

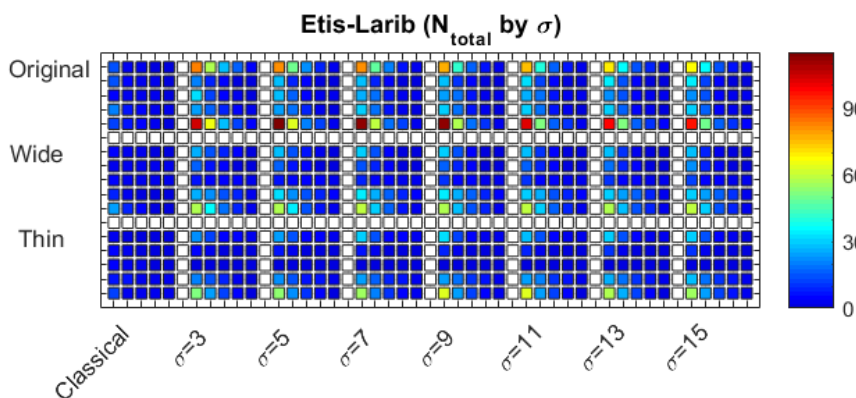


Figure 3-18 The total number of circles  $N_{total}$  found in the sample images of database *ETIS-Larib* (Silva, Histace, Romain, Dray, & Granado, 2014). The vertical axis represents the various sample pictures (images 25, 65, 82, 138, and 160) for both the non-gradient-weighted (original), and gradient-weighted voting process with wide and thin thresholds, respectively. The horizontal axis is grouped similarly to Figures 3-14 and 3-16

Table 3-4 The selected samples and their metrics  $R_{calc}$  and  $R_{edge}$  together with the total number of Prewitt edge pixels. The first two lines in each case are marked by gray color as they are the “bad” samples with low values of  $R_{calc}$ , the others have  $R_{calc}$  larger than 0.1. The last two columns contain the metrics regarding the size and roundness of the polyp

Database	Image	$R_{calc}$	$R_{edge}$	No. of Edge Pixels	$r_{avg}$	$\delta_{roundness}$
CVC-Clinic	29	0.004918033	0.013274336	610	38.75	0.358
	201	0.008460237	0.038314176	1182	44.25	0.222
	111	0.247629671	0.733884298	1793	92.75	0.380
	150	0.251902588	1.054140127	1314	52	0.261
	188	0.364678899	0.424	436	66.5	0.173
	217	0.230191827	0.74393531	1199	64.5	0.409
	265	0.260652765	0.871212121	2206	110	0.775
	390	0.297808765	0.641630901	1004	75.75	0.366
	475	0.367805755	0.653354633	1112	106.5	0.757
	480	0.222554145	0.685057471	1339	72.75	0.315
	503	0.310214376	0.976190476	1586	84.75	0.151
504	0.180173092	0.68358209	1271	56.75	0.201	
CVC-Colon	51	0.004398827	0.035087719	1364	27	0.278
	255	0.002773925	0.006779661	721	52.25	0.299
	62	0.242270224	0.801120448	2361	117	0.337
	74	0.218066743	1.457692308	1738	46.25	0.694
	101	0.180649379	1.978520286	4589	71.5	0.711
	128	0.182273052	0.44973545	1399	84	0.483
	149	0.215956424	1.099510604	3121	98	0.334
	220	0.149488927	0.966942149	2348	59.5	0.202
	230	0.318594104	0.641552511	882	74.25	0.442
283	0.194486983	0.404458599	653	52.25	0.132	
ETIS-Larib	24	0.007230077	0.204545455	6224	39.75	0.135
	151	0.0010755	0.012437811	4649	70.75	0.507
	25	0.104344123	0.879186603	7044	153.25	0.549
	65	0.117593198	1.24	7645	125.25	0.249
	82	0.191721133	1.269230769	4131	113.25	0.321
	138	0.102428256	0.649859944	2265	63.5	0.449
	160	0.129833607	1.297423888	8534	151	0.430

The fuzziness has a kind of optimal value around  $\sigma = 5$  (there the images have the largest number of dark red points and darkest red points), but this optimum is not very sharp, the neighboring  $\sigma$  values have very similar results in the domain of  $\sigma = \{3,5,7\}$ . Even for those images that have no circle found in the ring mask for the classical case, the results can be improved by the fuzzy version of the transform.

Regarding the gradient-weighted voting, and the gradient-based thresholding, the results improve compared to the original voting scheme, but the difference is not extremely large.

There is one sample image where the circles found are not around the polyp contour for the classical transform: image No. 475. However, this particular polyp has barely visible circular contour segments. The results improve even for this case for the gradient-weighted voting processes (except for the classical Hough transform in the thin threshold case). Here, we have to note that in order to have the transform calculable and also to remove the very large and very small circles from the results, the search for the  $r$  domain was limited between 1% and 25% of the longer image dimension, so the longer edges of the very elongated polyps, like the one on image No. 475 could not be detected.

From Figure 3-15, for database CVC-Colon, it can be seen that more circles are found outside of the polyp contour ring mask, i.e., the little squares in the plots tend towards the yellow and greenish part of the color palette.

The value of the ideal sigma seems to be a little bit higher, around  $\sigma = 7$  or  $\sigma = 9$ , but there is much more variation than in the case of the previous database. For some images the higher  $\sigma$  domains give better results, for others the lower  $\sigma$  domains.

The thin threshold loses the circles in the ring mask for more images, so it is not a good option for this database.

For the peak percentage  $P_p$  this database behaves similarly to the previous one.

Again, there is an image that does not show circles in the ring mask for most of the three cases: image No. 230, which has a brightly lighted front part, and the polyp is at the background.

From Figure 3-17 it can be seen that for database ETIS-Larib, the ratio of circles outside the ring mask is further increased, i.e., the  $A_r$  became even smaller and the plots became even less red.

Here, the higher  $\sigma$  values perform slightly better, but this tendency is even less clear than in the case of CVC-Colon.

The  $P_p$  value here is also like the previous cases, 70–80%.

The thin threshold is worsening the results: even though there are pictures, where the number of circles outside the ring mask decreases, there are more images, where the circles within the ring mask at the polyp contour disappear.

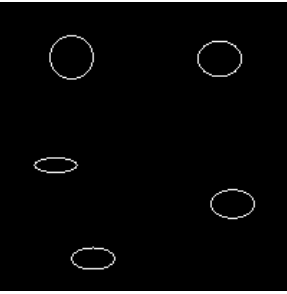
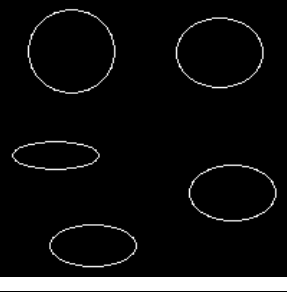
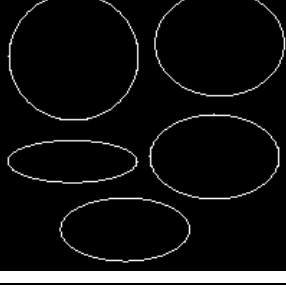
For image No. 82, where the polyp contour is very sharp, the polyp's circle gets lost due to the gradient thresholding for the thin threshold.



### 3.4.2 Roundness metrics evaluation

In order to test how tolerant the classical and fuzzy Hough transforms for the deviation from the circle, a test image of five ellipses in an image of size  $200 \times 200$  pixels was used. The larger axis of the ellipse was the same for each of the objects in one image, and the ratio of the two axes were 1:1, 5:6, 4:6, 3:6, and 2:6. Three images were used, one with a larger axis of 30 pixels, one with 60 pixels and one with 90 pixels. The properties of the ellipses are given in Table 3-5.

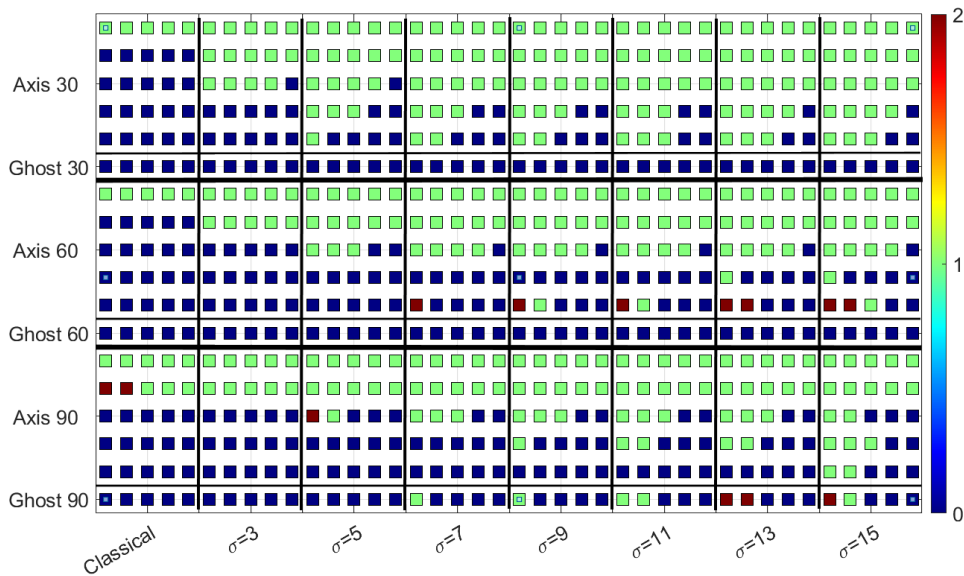
*Table 3-5 Properties of the ellipses of the artificial test images of size  $200 \times 200$  pixels. The order of the ellipses is the same as the order in the vertical axis of Figure 3-19, except for the ghost circles*

Artificial Image	Axes (Pixels)	$r_{avg}$	$\Delta_{radial}$	$\delta_{roundness}$
	30, 30	15.5	0.6508	0.042
	30, 25	14	1.8902	0.135
	30, 20	13	2.8758	0.221
	30, 15	11.75	4.7322	0.403
	30, 10	10.5	5.0081	0.477
	60, 60	30.5	0.7931	0.026
	60, 50	28	2.8902	0.103
	60, 40	25.5	5.2002	0.204
	60, 30	23	7.6023	0.331
	60, 20	20.5	10.0369	0.490
	90, 90	45.5	1.6081	0.035
	90, 75	41.5	4.2220	0.102
	90, 60	38	8.4958	0.224
	90, 45	34.25	12.1989	0.356
	90, 30	30.5	15.0686	0.494

In Figure 3-19, the number of circles that were found by the Hough transforms are summarized for each of the ellipses separately. As the centers of the ellipses were rather close to each other especially on the largest axis case, for lower thresholds one or more unphysical

“ghost” circles appeared, typically at the regions between 2 or 3, almost overlapping ellipses. The number of these ghost circles are also given under the numbers belonging to the real, physical ellipses.

It can be seen from Figure 3-19 that of the artificial image with objects of the smallest size, the classical Hough transform could only detect the precise circle, even the roundness error  $\delta_{roundness} = 0.135$  was not tolerable. As the fuzziness parameter increased, the peaks belonging to more and more elongated ellipses grew above the half of the global maximum of the votes, i.e., above  $P_p = 50\%$ . Already for  $\sigma = 5$  all the ellipses were found in the  $P_p = 50\%$  case. The corresponding radii were smaller than the larger axis of the ellipse.



*Figure 3-19 The number of circles found in the artificial images containing ellipses of large axes with size 30, 60, and 90 pixels. In the horizontal axis, the fuzziness parameter  $\sigma$  is the main ordering parameter, and within the cluster of each  $\sigma$ , the peak percentages  $P_p$  increase from 50% to 90%. In the vertical direction the three different greater axis sizes are denoted. Within each group, the topmost row belongs to the circle (axis ratio 1:1), the next row to the axis ratio 5:6, then 4:6, 3:6, 2:6. The last row, separated by a horizontal line, represents the non-physical ghost circles*

As the size of the ellipses is increased, the limit for finding all the ellipses went higher. Additionally, due to the closeness of the ellipses compared to the full width at half maximum of the voting matrix, overlaps arose between the individual ellipses and thus, false peaks appeared in the accumulator space, either in the area at the center of the image, which is circumvented by the ellipses, or at the two most elongated ellipses, that sometimes appeared as two twin peaks at the vote maps.



For the larger sized ellipse, even the classical Hough transform has a small roundness error tolerance:  $\delta_{roundness} = 0.1$  is still found, although as two circles.

If the roundness of the tested colonoscopy images is studied in the last two columns of Table 3-4, we can see that extremely elongated polyps also exist with  $\delta_{roundness} > 0.75$ , but the typical roundness error is between 0.15 and 0.5. This is the reason why this domain was selected for the ellipses to be studied in the artificial examples.

Similar to the case of the artificially generated test images, in the cases of very much elongated shapes  $\delta_{roundness} > 0.4$ , the polyps were very often not found if their sizes were large, or found only for very low threshold  $P_p$ , like in the cases of ETIS 1st and 4th images, or CVC-Clinic 4th.

Additionally, the smaller polyps with radii around 60 are usually detected, regardless of their roundness.

### 3.4.3 Time evaluation

To demonstrate the effect of the running time, the ratio between the total running times of the gradient-weighted method with wide and thin thresholds to the original Hough transform is given in Figure 3-20.

It can be seen that in the case of most of the images, the wide threshold decreases the running time of the Hough transform algorithms by about 50%, and the thin threshold decreases the runtime by another roughly 50%, but in some cases, it can go down to lower than 10% of the original transform's running time. For database CVC-Colon, the results are a bit better than for the other databases. From the 2nd subplot of Figure 3-20, we can conclude that these results are almost size independent, though for a smaller number of edge pixels, the results are usually a little bit better.

It is important to note that the number of pixels to be transformed has an almost linear effect on the transformation time, so using Prewitt edge detection is a really important step, as it can be seen from Figure 3-6. As the continuity of the edges is not a key point in Hough transform, and by giving it up, we can get rid of a lot of weaker edges in the background. This step has a double advantage.

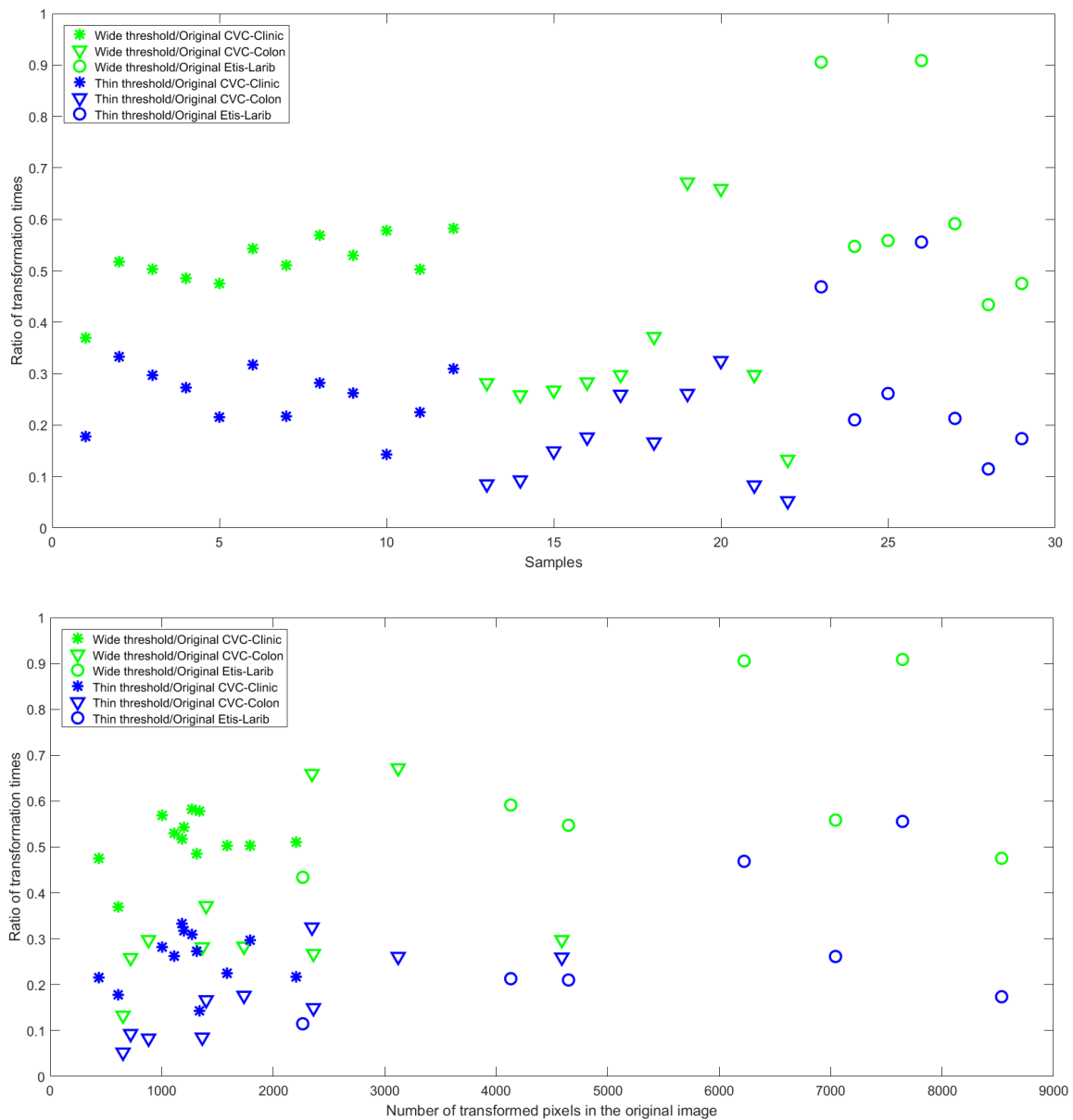


Figure 3-20 The ratio of the classical and fuzzy Hough transformations' total runtime for the wide and thin gradient thresholds compared to the original, full Hough transform times. The two subplots have different horizontal axes: the first subplot just lists the images according to their order in Table 3-4, the second subplot's horizontal axis is the number of pixels to be transformed in the original image (i.e., the 5th column of Table 3-4)

### 3.5 Thesis statements

#### 3.5.1 Thesis 3

In order to increase the classical Hough transform's tolerance to deviations from circle, I proposed to apply fuzzy Hough transform in the colorectal polyp localization task instead of the classical one, and I compared their performance.

- A. As the detection success also depends on the size of the colorectal polyp and its roundness, I introduced the roundness error together with the radial displacements and the average radius of the polyp to test the different types of Hough transforms' tolerance degree to the deviation from the circle.
- B. Using artificial images, I showed that if the Hough transform's fuzzyness parameter increases, or the local maximum threshold of the transformed image decreases, or the size of the object decreases, the tolerance for shape deviations increases as well.  
For classical Hough transform the tolerance increases with increasing the size of the object.

#### 3.5.2 Thesis 4

For colorectal polyp localization purposes, I developed a novel gradient-weighted voting approach for classical and fuzzy circular Hough transforms.

- A. As a preprocessing method for Hough transforms, I studied four different edge detection algorithms. I developed and applied two evaluation metrics using the number of pixels in the polyp contours, and the number of edge pixels detected by the algorithms both in the full image and in the environment of the polyp contour for all the images of 3 publicly available databases. The Prewitt edge detection algorithm proved to be the optimal choice for colorectal polyp localization purposes.
- B. In order to further decrease the number of unnecessary edge points to be processed by Hough transforms, I suggested an experimentally validated gradient-based thresholding method. I proposed two different gradient threshold domains based on the typical gradient domain of the polyp edge points, a wider and a thinner one with relative gradient values  $[0.06, 0.3]$ , and  $[0.08, 0.2]$ . The wider gradient threshold was proven to give more reliable results. Using this method, the runtime of the algorithms decreased by about 50% compared to the full Hough transforms. The thinner threshold

further decreased the runtime, but it removed so many valuable edge points which made the detection of the polyp impossible for several images.

- C. I introduced a gradient-weighted voting approach to both classical and fuzzy circular Hough transforms based on the probability distribution of the polyp contour pixels' gradients of images from 3 publicly available databases. The normalized gradient values serve as weights of the edge points in the Hough transformation step. For 8 grades of fuzzyness and 5 local maximum thresholds of the global maximum of the votes, I performed the Hough transforms for both the original and the gradient-weighted voting processes with wide and thin gradient thresholds. The gradient-weighted voting mechanism improved the ratio of the circles near the polyp among all the circles for the Hough transforms, especially for the fuzzy Hough transform. For general practice, I suggest using 80% of the global maximum of the votes as local maximum threshold and the application of fuzzy Hough transform with fuzzyness parameter  $\sigma = 7$  together with the wide gradient-weighted voting process.

## 4. Active contour methods in colonoscopy image segmentation

### 4.1 Introduction

Image segmentation is the process of partitioning an image into multiple meaningful and homogeneous regions or objects. It is a fundamental task in various applications of computer vision and image processing such as object recognition, tracking, and measurement.

Medical images segmentation has an exceptionally important role in medical diagnosis and classification, treatment planning, and image-guided interventions. It helps with identifying the location, size, and shape of several anatomical structures, tumors, lesions, and abnormalities in medical images.

Segmentation methods are generally classified into two main categories: edge-based and region-based segmentation methods. Edge-based segmentation methods use the discontinuities in the local properties like (brightness, color, or texture) as a criterion for determining the boundaries of the regions to be segmented, such as Canny approach (Canny, 1986) that depends on the detection of discontinuities in the brightness to build the connected curves. On the other hand, region-based segmentation methods depend on grouping pixels of similar local properties to define the boundaries of the areas to be segmented, like region growing (Adams & Bischof, 1994) and region merging (Brice & Fennema, 1970) approaches. Although these algorithms have not been widely successful in more complex tasks due to the absence of a well-defined optimization criterion, they have played a crucial role in developing more advanced approaches that utilize energy constraints and forces within images to separate the region of interest and rely on variational principles to achieve superior results (Ismail & Nagy, 2021).

Variational methods in image segmentation are mathematical frameworks used to define and optimize the boundaries between different regions or objects in an image, and consequently find a closed contour that divides the image into subregions. These methods use partial differential equations (PDEs) to model the evolution of the image over time, with the aim of finding a segmentation that minimizes an energy functional which captures the properties of the desired segmentation, such as smoothness and boundary adherence. By minimizing this energy functional, an optimal segmentation of the image can be obtained, where each region or object is separated by a well-defined boundary.

By the end of the 80's, the first variational formulations for image segmentation appeared, especially, the Snakes approach by Kass, Witkin, and Terzopoulos in 1988 (Kass, Witkin, &

Terzopoulos, 1988) which is an edge-based segmentation method, and the Mumford-Shah formulation in 1989 (Mumford & Shah, 1989) which is a region-based segmentation method.

Researchers' interest in active contour techniques for medical image segmentation has grown recently. (Yang & Jiang, 2020) constructed a new level set model based on non-local means (NLM) filtering with a new edge-stop criterion. A partially automatic hybrid active contour method using a local bias correction function and probability score was proposed in (Fang, Pan, Yao, Zhang, & Guo, 2020) for segmentation of 2D ultrasound images with noise, blurry boundaries, and intensity heterogeneity. (Bhat & Kumar, 2019) introduced a three phases approach to segment the optic disc in retinal fundus images. A localized active contour step was used for the 3rd, refined segmentation.

Segmenting colonoscopy images presents a demanding task in accurately outlining and separating colorectal polyps. Numerous complexities contribute to this challenge. To begin with, the quality of colonoscopy images exhibits significant variability due to factors like motion artifacts, poor illumination, and low contrast. Additionally, the edges of polyps' manifest diverse degrees of strength and prominence, attributed to variations in lesions morphology and their location on or within the bowel wall. Lastly, the diversity in colon anatomy among individuals creates a challenge in developing a segmentation algorithm that can effectively detect and segment the colorectal polyps across a wide range of populations.

Even though deep learning algorithms are often used for colonoscopy polyp segmentation, active contour methods have recently shown high efficiency in this field as well. In (Georgieva & Petrov, 2017), (Georgieva, Petrov, Nagy, & Sziová, 2018), (Sasmal, Iwahori, Bhuyan, & Kasugai, 2018), and (Dutta, Sasmal, Bhuyan, & Iwahori, 2018) researchers used the active contour without edges approach, i.e., Chan-Vese segmentation model to build segmentation schemes for polyps with not strong enough edges. The proposed methods were accompanied with various pre-processing steps and different evaluation metrics.

This is the reason why the main objective of this chapter is to conduct a systematic and comprehensive assessment of the effectiveness of not only the Chan-Vese active contour method, but another powerful variational segmentation method, namely, the geodesic active contour method for colorectal polyps segmentation purposes (Ismail & Nagy, 2022). Due to the complex bowel environment, non-uniform properties of polyps, and variations in their edge strength, it is interesting to investigate and compare the feasibility of the geodesic and Chan-Vese active contour methods application, especially because they operate through two different mechanisms.

In the following sections, first, I give a mathematical background about the variational methods and the level set methods for image segmentation. Then, I introduce the metric used in evaluating the performance of the segmentation methods. Finally, I apply and evaluate the geodesic and Chan-Vese active contour methods for a carefully selected set of sample images with different types of preprocessing steps and initial masks.

## 4.2 Variational methods for image segmentation – mathematical background

### 4.2.1 Snakes: active contour models

Active contour, also known as "snake", is one of the first variational approaches for image segmentation and it is considered among the most effective models. It was first introduced by Kass, Witkin, and Terzopoulos in 1988 (Kass, Witkin, & Terzopoulos, 1988), and since then, it has been developed and modified by many researchers (Caselles, Kimmel, & Sapiro, 1997), (Chan & Vese, 2000), (Chan & Vese, 2001). The classical snake approach is briefly summarized as follows:

1. Initialization: the active contour is defined by drawing an initial contour or curve that roughly outlines the object of interest within the image.
2. Energy calculation: the snake calculates the energy of the curve, which is the sum of internal energy and external energy.
3. Minimization: the snake iteratively moves the contour to a more optimal place by decreasing its energy in each sequential step. This movement is controlled by the forces generated by the energy function.
4. Convergence: the snake continues to move until it reaches convergence, which is when the energy of the curve is minimized, and the contour has settled on the boundary of the object of interest (Ismail & Nagy, 2021).

In (Kass, Witkin, & Terzopoulos, 1988), authors proposed to minimize the following energy functional which has two parts

$$E(C) = E_{ext}(C) + E_{int}(C). \quad (4.1)$$

The external and internal energies are formulated by the following equations respectively

$$E_{ext}(C) = - \int_0^1 |\nabla I(C(s))|^2 ds, \quad (4.2)$$

$$E_{int}(C) = \int_0^1 \left\{ \frac{\alpha}{2} |C_s(s)|^2 + \frac{\beta}{2} |C_{ss}(s)|^2 \right\} ds, \quad (4.3)$$

where,  $I : \Omega \subset R^2 \rightarrow R$  is the input image, and  $C : [0,1] \rightarrow \Omega$  denotes a parametric curve on a grid  $\Omega$ . The weights  $\alpha \geq 0$  and  $\beta \geq 0$  deal with the elasticity and the stiffness of the curve respectively.  $C_s$  and  $C_{ss}$  are the first and second derivative of the curve  $C$  with respect to its parameter  $s$ .

The previous equations show that the energy state of the contour at any time is controlled by both external and internal energies. The external energy is related to the properties of the image, like the image gradient, and measures for the computed curve  $C$  how well it agrees with the maximum of the brightness gradient  $|\nabla I|$ . On other hand, the internal energy is defined by the curve shape characteristics such as: length, smoothness, stiffness, elasticity, curvature, and continuity.

Minimizing the total energy function (4.1) creates curves that are short, smooth, and stiff while passing through locations with large gradient values. Gradient descent minimization is one of the simplest techniques which can be used to minimize the total energy function.

Back to the equation (4.1) and substitute the external and internal energies by their values, we get

$$E(C) = - \int_0^1 |\nabla I(C)|^2 ds + \int_0^1 \left\{ \frac{\alpha}{2} |C_s(s)|^2 + \frac{\beta}{2} |C_{ss}(s)|^2 \right\} ds. \quad (4.4)$$

The corresponding second-order partial differential equation (Euler-Lagrange equation) which must be satisfied at the minima of (4.4) is given by

$$\alpha C_{ss} - \beta C_{ssss} - \nabla E_{ext}(C) = 0. \quad (4.5)$$

Equation (4.5) represents a force balance equation

$$f_{int}(C) + f_{ext}(C) = 0, \quad (4.6)$$

where  $f_{int}(C) = \alpha C_{ss} - \beta C_{ssss}$  is the two-parts internal force, the tension energy  $\alpha C_{ss}$  which tends to smooth out the curve and prevent it from having sharp corners, together with the bending energy  $\beta C_{ssss}$  which penalizes large curvatures and helps to maintain the overall shape of the curve.  $f_{ext}(C) = -\nabla E_{ext}(C)$  is the external force that pulls the contour toward the desired location.

To solve equation (4.5),  $C(s)$  is treated as a function of temporal and spatial changes, i.e.,  $C(s, t)$ . The solution is obtained at the stable-state solution of the following gradient descent equation



$$\frac{\partial C(s,t)}{\partial t} = \alpha C_{ss}(s,t) - \beta C_{ssss}(s,t) + f_{ext}(C(s,t)). \quad (4.7)$$

We can find a numerical solution to (4.7) by discretizing the equation, after that solving the discrete system iteratively (Kass, Witkin, & Terzopoulos, 1988).

The limitations of these models (i.e., snakes, active contour models) are well defined. They tend to be sensitive to initial conditions and are typically placed close to the boundaries of the objects of interest. Moreover, these models use an explicit parameterization that makes them unable to handle significant topological changes.

#### 4.2.2 Mumford-Shah approach

The Mumford-Shah approach is a popular image segmentation method that was first introduced by the mathematicians David Mumford and Jayant Shah in 1989 (Mumford & Shah, 1989). It is a variational method that seeks to partition an image into regions or segments based on the image's intensity values and the boundaries between these regions.

Mumford and Shah proposed to compute the segmentation of the input image  $I : \Omega \subset R^2 \rightarrow R$  by minimizing a cost functional which has three main parts as follows

$$E(u, C) = \int_{\Omega} (I(x) - u(x))^2 dx + \lambda \int_{\Omega/C} |\nabla u(x)|^2 dx + \nu |C|. \quad (4.8)$$

This relates to the piecewise smooth approximation  $u : \Omega \rightarrow R$  of the input image  $I$ , in addition to a one-dimensional discontinuity set  $C \subset \Omega$ . The three components of equation (4.8) have the following definitions:

- The first part is a quadratic data term which demonstrates that  $u$  is a convenient approximation of the input image  $I$ .
- The second one is a smoothness term, with weight  $\lambda > 0$ , confirms that the entire  $u$  is smooth except for the boundary discontinuity set  $C$ .
- The third one is another regularizer, with weight  $\nu > 0$ ; it limits the number of pixels where the discontinuities are allowed by assuring that the minimal length of discontinuity set is  $|C|$ .

To obtain  $u(x)$  which is constant in each of the regions separated by the boundary  $C$ , the piecewise-constant Mumford-Shah functional is defined as

$$E(u, C) = \int_{\Omega} (I(x) - u(x))^2 dx + \nu |C|. \quad (4.9)$$

If we refer to these regions by  $\{ \Omega_1, \dots, \Omega_n \}$  and the constants by  $u_i$ , we will have the following general formula

$$E(\{u_1, \dots, u_n\}, C) = \sum_{i=1}^n \int_{\Omega} (I(x) - u_i)^2 dx + v|C|. \quad (4.10)$$

The last formula is the backbone of many methods used in this field. The Chan-Vese model is a prominent example of the most influential methods. The optimization problem involved in the Mumford-Shah approach can be solved using a variety of numerical techniques, such as gradient descent or level set methods.

### 4.3 Level Set Methods for image segmentation– mathematical background

During the early 1990s, numerous research papers were published on the use of Level Set Methods (LSMs) for image segmentation. LSMs concept was first published in 1979 by Dervieux and Thomasset (Dervieux & Thomasset, 1979), and subsequently reinvented in 1988 by Osher and Sethian (Osher & Sethian, 1988). LSMs are important tools for numerical analysis and representation of curves, shapes, and surfaces, and their application has become common in image segmentation and curves evolution related domains.

The topological changes of the curve are treated automatically since the temporal evolution of a curve  $C(t)$  is implicitly represented using the zero-level set of a function  $\phi(x, t)$ , with a parameterization free formulation, such that

$$C(t) = \{x \in \Omega \mid \phi(x, t) = 0\}. \quad (4.11)$$

As long as the curve evolves over time, the values of the function  $\phi(x, t)$  will also be updated at each time and position to create some desired motion for the curve. The fundamental question that arises here is: how to evolve the function  $\phi$  to mimic the curve evolution. When we have a curve evolution, the motion of the curve  $C(t)$  is expressed by

$$\frac{dC}{dt} = Fn, \quad (4.12)$$

where  $F$  is a local speed along with the outer normal  $n$ . Based on the previous speed of the curve, we should know how to evolve the function  $\phi$  to represent the curve evolution successfully. One basic requirement, at any time, is the function  $\phi$  must be zero for all points of the curve  $C$

$$\phi(C(t), t) = 0 \quad \forall t. \quad (4.13)$$

If the previous equation is fulfilled, then its derivative with respect to time must vanish

$$0 = \frac{d}{dt} \phi(C(t), t) = \nabla \phi \cdot \frac{dC}{dt} + \frac{\partial \phi}{\partial t}. \quad (4.14)$$

We can find the temporal evolution of  $\phi$  by substituting each of the equation (4.12) and the outer normal ( $n = -\frac{\nabla \phi}{|\nabla \phi|}$ ) in equation (4.14) as follows

$$\frac{\partial \phi}{\partial t} = -\nabla \phi \cdot \frac{dC}{dt} = \nabla \phi \cdot F \frac{\nabla \phi}{|\nabla \phi|} = F |\nabla \phi|. \quad (4.15)$$

As a consequence, for a curve evolution in a normal direction with speed  $F$ , the function  $\phi$  at the zero level must follow the following level set equation

$$\frac{\partial \phi}{\partial t} = F |\nabla \phi|. \quad (4.16)$$

As the motion of the function  $\phi$  at the boundary of the curve is defined using the level set equation, the evolution outside the specific location of the curve boundaries will often take place randomly. Typically, one treats the level set function  $\phi(x, t)$  as a signed distance function, i.e.

$$\phi(x, t) = \pm \text{dist}(x, C). \quad (4.17)$$

Mostly,  $\phi(x, t)$  is positive inside the curve and negative outside it, or vice versa, it is a matter of choice.

One key advantage of applying LSMs in image segmentation is their ability to handle complex and evolving boundaries between regions of interest. Unlike traditional segmentation techniques that rely on explicit parameterization of the boundaries or intensity thresholds, LSMs use implicit representations of the segmentation boundaries via the zero-level set function. This allows LSMs to capture topological changes and treat them automatically, handle concavities, and merge/split regions naturally without the need for re-initialization or manual intervention.

Using LSMs concept, the geodesic and the Chan-Vese active contours approaches were proposed to find a level set representation for each of snake model and Mumford-Shah model, respectively.

### 4.3.1 Geodesic active contours

In 1997, the geodesic active contours model, derived from "snakes", was first introduced by Caselles, Kimmel, and Sapiro (Caselles, Kimmel, & Sapiro, 1997), (Kichenassamy, Kumar, Olver, Tannenbaum, & Yezzi, 1995). The edge-based segmentation energy functional for the geodesic active contours has the following form

$$E(C(s)) = \int_0^{L(C)} g(|\nabla I(C(s))|) ds. \quad (4.18)$$

The geodesic length is controlled by the edge strength measurement function or the stopping function  $g(I)$ . The basic aim of  $g(I)$  is to limit the contour evolution and stop it when the contour gets close enough to the desired edge. It is defined as follows

$$g(I) = \frac{1}{1+|\nabla \hat{I}|^p}, \quad (4.19)$$

where  $\hat{I}$  is the smoothed image of the considered image  $I$  and  $p = 1$  or  $2$ . It is obvious that the value of the function  $g(I)$  will decrease when we have high values of the gradient. The gradient descent of the curve  $C$  is given by the following expression

$$\frac{\partial C(t)}{\partial t} = g(I)\kappa \bar{N} - (\nabla g \cdot \bar{N})\bar{N}, \quad (4.20)$$

where  $\kappa = \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right)$  is the Euclidean curvature and  $\bar{N}$  is the unit inward normal. Representing the previous equation using the level set equation will be as follows

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \text{div} \left( g(I) \frac{\nabla \phi}{|\nabla \phi|} \right) = g(I)|\nabla \phi| \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) + \nabla g(I) \cdot \nabla \phi, \quad (4.21)$$

where  $\phi$  is a function representing the curve  $C$  as a zero-level set.

Although the geodesic active contours approach has been successful in certain areas, it has limited applicability in other specific scenarios. This is because it depends on image gradient information and may not be the best choice for processing images that are noisy or contain objects with poorly defined edges.

### 4.3.2 Chan-Vese active contours

An alternative approach to the edge-based methods, Chan and Vese proposed a two-phase region-based segmentation method known as active contour without edges (Chan & Vese, 2000), (Chan & Vese, 2001). This approach reformulates the variational principle of the piecewise constant Mumford-Shah model using the level set method. Chan and Vese introduced the following basic energy functional

$$F(c_1, c_2, C) = \mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(\text{inside}(C)) \\ + \lambda_1 \int_{\text{inside}(C)} |I(x, y) - c_1|^2 dx dy + \lambda_2 \int_{\text{outside}(C)} |I(x, y) - c_2|^2 dx dy, \quad (4.22)$$

where  $C$  is the curve that splits the image  $I$  into two regions of nearly constant intensity, the constants  $c_1$  and  $c_2$  are the averages of  $I$  inside and outside the contour  $C$  respectively, and  $(\mu \geq 0; \nu \geq 0; \lambda_1, \lambda_2 > 0)$  are fixed parameters. The default values most often used are:  $\lambda_1 = \lambda_2 = 1$  and  $\nu = 0$ . Then, the minimization problem is presented as

$$\inf_{c_1, c_2, C} F(c_1, c_2, C). \quad (4.23)$$

Equation (4.22) is a special case of the piecewise-constant Mumford Shah equation (4.10) for two phases. Chan and Vese worked on finding the best approximation  $u$  of the image  $I$ , as a function taking only two values

$$u = \begin{cases} \text{average}(I) \text{ inside } C \\ \text{average}(I) \text{ outside } C \end{cases} \quad (4.24)$$

The authors reformulated (4.22) using the level set method to implicitly represent the contour  $C$  by the zero-level set in terms of a  $\phi$  function as following

$$\begin{cases} C = \partial\omega = \{(x, y) \in \Omega: \phi(x, y) = 0\} \\ \text{inside}(C) = \omega = \{(x, y) \in \Omega: \phi(x, y) > 0\} \\ \text{outside}(C) = \Omega \setminus \bar{\omega} = \{(x, y) \in \Omega: \phi(x, y) < 0\} \end{cases} \quad (4.25)$$

With the evolving curve  $C \subset \Omega$ , as the boundary of an open subset  $\omega \subset \Omega$ ,  $\omega$  refers to the region inside the curve  $C$ , and  $\Omega \setminus \bar{\omega}$  refers to the region outside the curve  $C$ .

The Heaviside step function was their choice, and since it is not a differentiable function, so they assumed that it is slightly smoothed and by this way they were able to use the gradient descent to find the local minimization of the Chan Vese energy functional. Thus, its derivative will be a smoothed Dirac delta function. Generally, the Heaviside and Dirac delta functions are defined as follows

$$H(z) = \begin{cases} 1, & \text{if } z \geq 0, \\ 0, & \text{if } z < 0, \end{cases} \quad \delta_0(z) = \frac{d}{dz} H(z). \quad (4.26)$$

After that, they redefined the energy functional using  $H(\phi)$  by the following way

$$\begin{aligned} F(c_1, c_2, \phi) &= \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy \\ &\quad + \nu \int_{\Omega} H(\phi(x, y)) dx dy \\ &\quad + \lambda_1 \int_{\Omega} |I(x, y) - c_1|^2 H(\phi(x, y)) dx dy \\ &\quad + \lambda_2 \int_{\Omega} |I(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy. \end{aligned} \quad (4.27)$$

The final segmentation curve could be obtained from the minimizer of the energy functional (4.27) with respect to  $\phi$ . The standard Euler-Lagrange calculus is used to compute the gradient descent equation which defining the initial contour as the following

$$\frac{\partial \phi}{\partial t} = \delta_{\varepsilon}(\phi) \left[ \mu \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - \nu - \lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2 \right] = 0 \text{ in } \Omega,$$

$$\frac{\delta_{\varepsilon}(\phi)}{|\nabla \phi|} \frac{\partial \phi}{\partial \vec{n}} = 0 \text{ on } \partial \Omega, \quad (4.28)$$

where  $\vec{n}$  denotes the exterior normal to the boundary, and  $\partial \phi / \partial \vec{n}$  denotes the normal derivative of  $\phi$  at the boundary.

#### 4.4 Practical implementation and results

##### 4.4.1 Evaluation criteria

Practically, random image samples from colonoscopy images databases ETIS-Larib, (Silva, Histace, Romain, Dray, & Granado, 2014), CVC-Colon (Bernal, Sanchez, & Vilariño, 2012), and CVC-Clinic (Bernal J. , et al., 2015) were tested. Each of the calculations started from a rough initial mask – rectangular or circular–, based on the ground truth mask of the examined image. The initial mask can be also the output of any of the existing polyp localization methods (Silva, Histace, Romain, Dray, & Granado, 2014), (Nagy, Lilik, & Kóczy, 2017), (Yuji, et al., 2015), (Bernal J. ,et al., 2017).

To evaluate the results of the active contour segmentation methods, the Sørensen–Dice Similarity Coefficient was selected as a performance measurement metric. The Sørensen–Dice index (SDI) is a widely applied metric for quantitatively validating of medical image segmentation algorithms. Moreover, it directly compares between the resulted segmented area and the mask provided in the database, i.e., the so-called ground-truth mask (Dutta, Sasmal, Bhuyan, & Iwahori, 2018), (Taha & Hanbury , 2015), (Carass, Roy, Gherman, & et al., 2020). It is mathematically computed as

$$Dice(A, B) = \frac{2 \times |A \cap B|}{|A| + |B|}, \quad (4.29)$$

where  $A$  and  $B$  are the two sets being compared,  $|A|$  and  $|B|$  are the number of elements in  $A$  and  $B$  respectively, and  $|A + B|$  is the number of elements that are common to both  $A$  and  $B$ .

The SDI ranges from 0 to 1, with 0 indicating no overlap between the sets and 1 indicating complete overlap or similarity. It is a symmetric measure, meaning that the SDI between set  $A$  and set  $B$  is the same as the SDI between set  $B$  and set  $A$ .

For this study,  $A$  represents the active contour output binary segmented area and  $B$  represents the ground-truth mask.

##### 4.4.2 Reference case

As a reference case, geodesic and Chan-Vese active contour methods were tested on the 2D grayscale version of the colonoscopy original images and their original masks. Since the initial mask plays a critical role in the contour evolution and the accuracy of the segmentation process, an initial mask was generated based on the corresponding given mask in the related database. It is a rectangle that encompasses the polyp area, to simulate the output of a previous, rough

polyp localization method. Moreover, a solid initial mask will provide a clear definition of the region of interest and help the approaches converge to the desired solution.

In MATLAB, the smoothness factor and contraction bias parameters are two important parameters used in the active contour function to modify the behavior of the active contour during the segmentation process. By adjusting the values of these parameters, the user can control the trade-off between contour smoothness and flexibility as well as the degree of expansion or contraction of the contour, which improves the accuracy and robustness of the segmentation results.

The default parameters utilized in the active contours' methods provided by MATLAB were selected. The smoothness factor is 1 and 0, and the contraction bias is 0.3 and 0 for the geodesic and the Chan–Vese active contours respectively, and the number of iterations is 100 for both. Certain samples were selected and analyzed as representatives of the broader sample pool, and Table 4-1 lists the Sørensen–Dice Coefficient for these images from the three databases.

*Table 4-1 Sørensen–Dice Coefficient for the image samples in the reference case*

<b>Database</b>	<b>Image Number</b>	<b>Geodesic</b>	<b>Chan–Vese</b>
<b>ETIS-Larib Polyp DB</b>	<b>49</b>	0.8688	0.6771
	<b>65</b>	0.8922	0.7111
	<b>135</b>	0.8659	0.3950
	<b>185</b>	0.8702	0.6350
<b>CVC-Colon Polyp DB</b>	<b>28</b>	0.9225	0.7442
	<b>67</b>	0.9097	0.6016
	<b>150</b>	0.8832	0.7116
	<b>202</b>	0.9235	0.4318
<b>CVC-Clinic Polyp DB</b>	<b>127</b>	0.8865	0.5657
	<b>151</b>	0.9330	0.6524
	<b>302</b>	0.8583	0.6073
	<b>405</b>	0.8968	0.6512

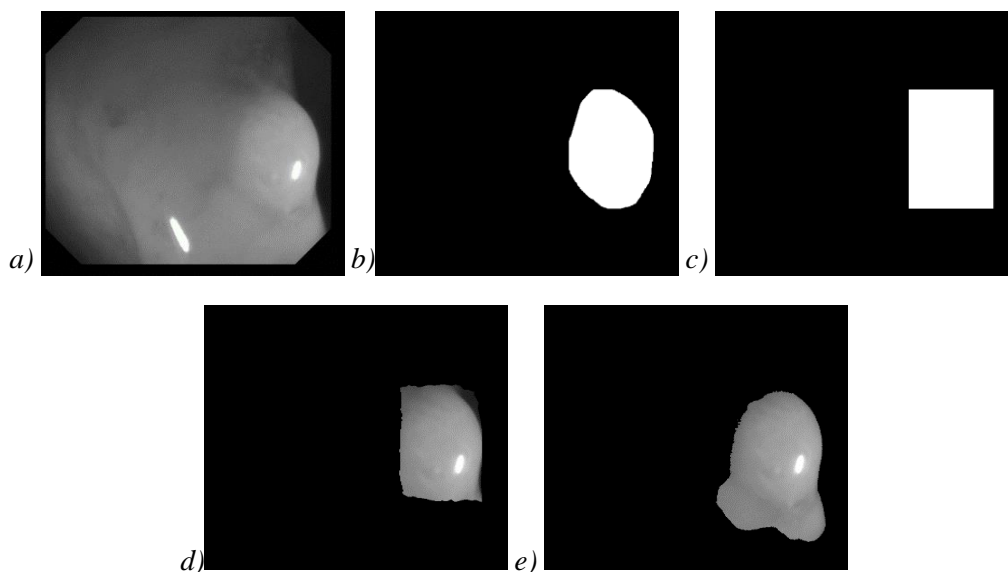
The numerical results of the Sørensen–Dice Coefficient showed that the geodesic method performs better than the Chan–Vese method using the first default settings. Moreover, the geodesic initial mask and thus the resulting segmented area showed a tendency to shrink or contract towards the polyp, while for the Chan–Vese, it tended to grow or expand outwards, away from the polyp. It is also worth mentioning that the geodesic segmented area has an almost

irregular rectangular shape, thus the contour fails to mimic the exact smooth and elliptical polyp shape.

The different behaviors of the geodesic and Chan-Vese initial masks in the segmentation results may be attributed to their underlying algorithms and energy minimization strategies. Geodesic active contours rely on an initial seed point and minimize an energy functional towards the object boundaries, resulting in the mask shrinking towards the polyp. On the other hand, Chan-Vese active contours partition the image into object and background regions based on their intensity values and seek to find the optimal contour that separates these regions, causing the mask to expand away from the polyp searching for this contour.

Figure 4-1 represents the reference case segmentation results of the grayscale version of image (28) in database CVC-ColonDB (Bernal, Sanchez, & Vilariño, 2012).

To enhance the overall segmentation accuracy, a series of main pre-processing steps were suggested and their effect in comparison with the reference case results was studied.



*Figure 4-1 The reference case segmentation results of the grayscale version of image (28) in database CVC-ColonDB (Bernal, Sanchez, & Vilariño, 2012). (a) The original image, (b) the given database mask, (c) the generated initial mask, (d) the geodesic segmented area, and (e) the Chan–Vese segmented area*

#### 4.4.3 Pre-processing phase

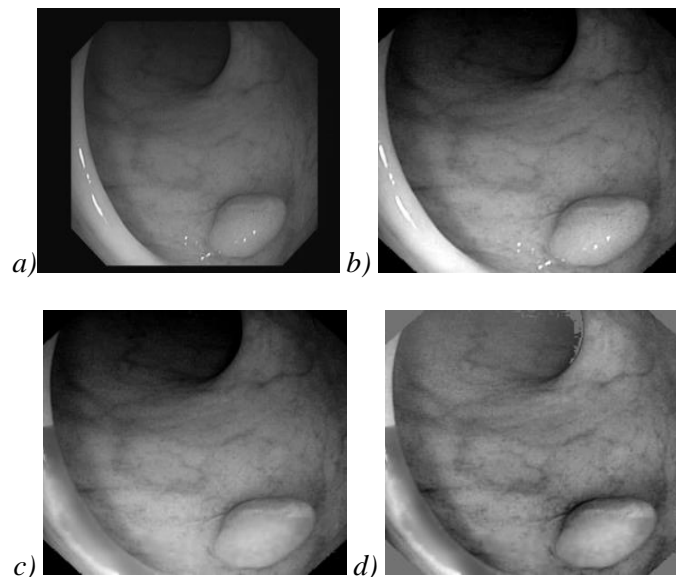
The following image pre-processing steps were carried out on all the colour channels to make the images more segmentable (Solecki, Nagy, & Sziová, 2020). First, the black frames of the images were cut off to reduce unnecessary information, then the reflections were smoothed into their environment together with a histogram stretching the following way. As all the images



contained remainings of the frames at the corners, there was a large black peak at the black side of the image histogram; these pixels were collected into a black mask. As all the pictures contained bright whitish reflections, there was a peak, lower in magnitude, but much broader in width at the other, white side of the histogram, too. The pixels with this intensity were collected into a white mask. The middle part of the image histogram that was the useful information part, was stretched to the 0 – 255 intensity domain linearly.

Two steps were executed on the reflection pixels to smoothen them into their environment. First, the white or reflection mask was extended to its neighbouring pixels, if the average of the unmasked environment of the given pixel was much smaller than the actual pixel value. Second, the pixels in the reflection domains (inside the extended white mask) were substituted by the average of their unmasked environment (only those pixels counted into the environment, where neither the white, nor the black masks were present).

As a last step a large-scale average of the image, i.e., the background, was subtracted to provide a more even density distribution, where only the patterns that are of the same size-scale of the expected polyps are visible, the larger tendencies are washed out. This step of course was together with a histogram stretching. Figure 4-2 demonstrates the pre-processing steps for image (127) in database CVC-ClinicDB (Bernal J. , et al., 2015).



*Figure 4-2 The pre-processing steps of image (127) in database CVC-ClinicDB (Bernal J. , et al., 2015). (a) The original image, (b) the image with the frame cut-off and histograms stretched, (c) the removed reflection image, and (d) the background subtracted image*

#### 4.4.4 Post pre-processing case

Since the characteristics of the colonoscopy images databases are not uniform, and the applicability of the segmentation methods varies according to the used pictures features, there is a need to investigate whether the performance of the segmentation process is affected by the application of the pre-filtering techniques. Along with the background-subtracted image, various pre-filtering methods' outputs were examined and tested as input images for the active contour methods.

When choosing filtering techniques for colonoscopy pictures, various factors must be considered. As these images are often noisy, reducing noise components can lead to more accurate segmentation results. Linear filters, such as mean or Gaussian filters, can be efficient options for this purpose. In some cases, it may be necessary to reduce noise while preserving edges, which is particularly useful for the edge-based segmentation methods, like the geodesic method. The median filter is less sensitive to outliers than the mean filter and can remove the outliers' pixels without reducing the strength of the image edges, as well as the Wiener filter can reduce noise levels while maintaining edges. While certain denoising filters may result in blurred images with smoother edges, which can be beneficial when implementing the Chan-Vese method as it looks for a smooth contour between objects of varying intensity to start the segmentation process with.

The first set of selected filters consisted of the basic statistical filters: the mean, median, and standard deviation filters. Furthermore, the difference between the mean and median was also evaluated, as it provides some detailed information about rapidly changing parts of the image (Ismail & Nagy, 2021).

As gradients are the basis of many edge detecting methods and their values are higher around large variations, i.e., around edges, thus the second suggested group of filters was the image gradients. The gradient magnitude, the gradient direction, and the  $x$  and  $y$  direction components of the gradient were implemented.

The third group of filters was based on the Rényi entropies (Rényi, 1960) of the image. The structural entropy and the spatial filling factor (mathematically explained in Subsection 2.2.2 in Equations (2.10) to (2.12)) were calculated as suggested filtration techniques as well. The filter size for the three filters groups previously listed was  $5 \times 5$ .

The Gaussian filter with a standard deviation of 0.5, and the Wiener filter with a neighborhood size  $3 \times 3$  were also applied as the last two filtering choices.

Practically, the 2D grayscale version of the background-subtracted image and its fourteen filtered images were evaluated as different input pictures for the geodesic and Chan-Vese active contour methods. We kept using the rectangular initial mask with the same active contour default settings in this case.

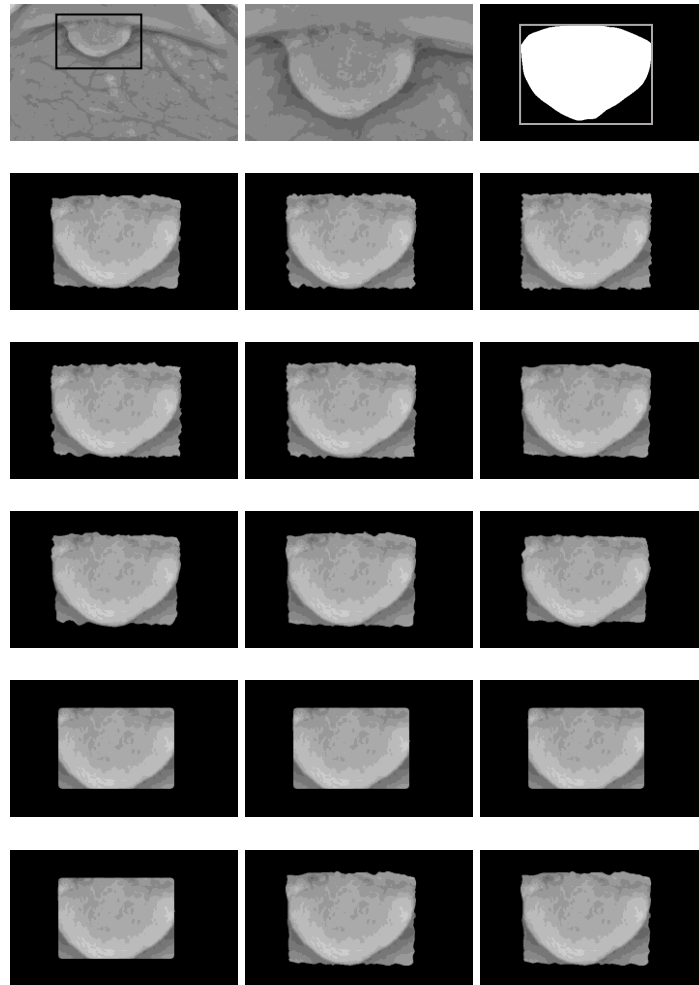
Compared to the reference case, the results confirmed that using background-subtracted versions instead of the grayscale versions of the original colonoscopy images was highly beneficial and improved the segmentation results in 75% of the samples using the Chan-Vese method (9 out of 12). Whereas the geodesic method results showed that only 42% of the samples (5 out of 12) were positively affected. The Sørensen-Dice Coefficient values for the remaining image samples (3 out of 12 for the Chan-Vese method, and 7 out of 12 for the geodesic method) showed only a slight decrease compared to the reference case results. In certain samples, the pre-processing phase generated images with smoother edges, which had a negative impact on the edge-based geodesic method's performance.

The former results are clearly illustrated by comparing the Sørensen–Dice Coefficient values of using the background-subtracted image and the rectangular initial mask for each tested sample in Tables 4-2, 4-3, and 4-4 with the corresponded reference case result (the tables are listed in the forthcoming part, in Subsection 4.4.5 for better comparability of the results). The highlighted value indicates the value that improved after the pre-processing phase and using the background-subtracted version as an input for the active contour methods.

The outcomes of the pre-filtering methods demonstrated that both the geodesic and Chan-Vese active contour methods had superior performance with the mean, median, Gaussian, and Wiener filters across almost all samples. Moreover, the Chan-Vese segmented area matched the elliptical polyp borders very well. Whereas the geodesic resulted area was still far from the ideal requested shape.

Despite the nearly identical performance of the geodesic method with almost all filters, for our next investigations, we will keep on just the mean, median, Gaussian, and Wiener filters, because the best segmentation results were obtained using these filters for both the geodesic and Chan-Vese methods.

Figures 4-3 and 4-4 give the resulting masked images for the background-subtracted image (65) in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014) using the geodesic and Chan-Vese methods respectively.



*Figure 4-3 The masks resulting from the geodesic method for one input image with different prefiltering methods. The original input image is the background subtracted image (65) in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014); it appears in the top left, in position 1), next, in position 2) is the zoomed version of the original, unfiltered image for better visualization (indicated in image 1) by a black rectangle), and 3) the ground truth mask belonging to the zoomed partition; the gray rectangle denotes the initial mask boundaries we used in the calculation. The order of appearance for the geodesic masks: 4) original, unfiltered image, 5-8) Gradient filtered images (magnitude, direction, x and y component), 9) mean filtered image, 10) standard deviation filtered image, 11) median filtered image, 12) difference of mean and median filtered image, 13)  $S_1$ , 14)  $S_2$ , 15)  $S_{str}$ , 16)  $\ln_q$  filtered images, 17) Gaussian filtered image, and 18) Wiener filtered image*

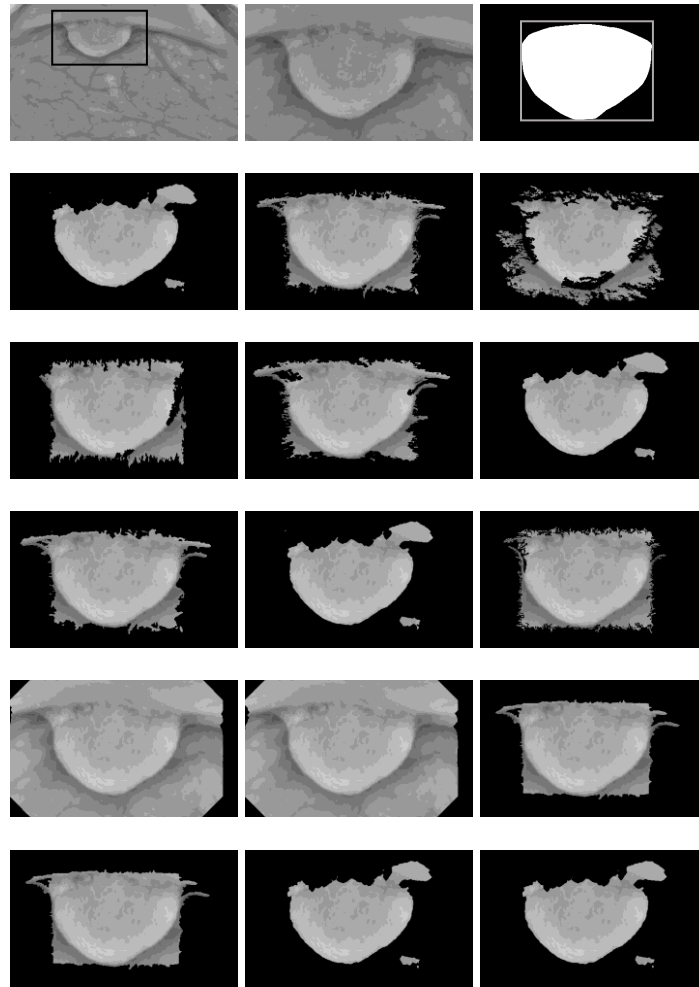


Figure 4-4 The masks resulting from the Chan-Vese method for one input image with different prefiltering methods. The original input image is the background subtracted image (65) in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014); it appears in the top left, in position 1), next, in position 2) is the zoomed version of the original, unfiltered image for better visualization (indicated in image 1) by a black rectangle), and 3) the ground truth mask belonging to the zoomed partition; the gray rectangle denotes the initial mask boundaries we used in the calculation. The order of appearance for the geodesic masks: 4) original, unfiltered image, 5-8) Gradient filtered images (magnitude, direction,  $x$  and  $y$  component), 9) mean filtered image, 10) standard deviation filtered image, 11) median filtered image, 12) difference of mean and median filtered image, 13)  $S_1$ , 14)  $S_2$ , 15)  $S_{str}$ , 16)  $\ln_q$  filtered images, 17) Gaussian filtered image, and 18) Wiener filtered image

#### 4.4.5 Circular initial mask case

To examine the impact of the initial mask shape on the performance of geodesic and Chan-Vese methods, a circular initial mask as an alternative to the previously used rectangular initial mask was proposed to be utilized. This investigation was motivated by the observation that certain polyp localization algorithms, such as those based on Hough transforms, provide a circle as the initial estimate of the polyp's location.

For this goal, three different circular initial masks were assessed. The centers of the circular initial masks were similar to the center of the previously used rectangular initial mask, but the diameters were different. The width, length, and diameter of the rectangular mask were considered as potential diameters for three circular initial masks.

According to the analysis, the most reasonable results were achieved using a circular initial mask with a diameter equal to the length of the original rectangular mask. Therefore, only the results obtained using this particular circular mask will be presented and discussed.

The values of Sørensen–Dice similarity coefficient for geodesic and Chan-Vese methods using both rectangular and circular initial masks and default parameters for the five selected input images were compared. The comparison results are illustrated as charts in Figure 4-5, and numerically listed in the Tables 4-2, 4-3, and 4-4.

Generally speaking and based on the geodesic comparison results, it was found that the rectangular initial mask yielded higher similarity coefficients in 91.67% of the cases, whereas the circular initial mask only worked better in only 8.33% of the cases. Furthermore, the Chan-Vese results showed that the rectangular initial mask was better suited for approximately 50% of the total similarity coefficients, while the circular initial mask was more appropriate for the remaining 50%. Overall, the Chan-Vese method demonstrated lower sensitivity to the shape of the initial mask than the geodesic method, as previously noted.

To further qualify how much lower the sensitivity of the Chan-Vese method is to the shape of the initial mask compared to the geodesic method and whether or not this difference is significant, the relative difference between rectangular and circular initial masks' Sørensen–Dice indices results was calculated. It is defined the following way

$$Relative\ Difference = \frac{Dice_{rect} - Dice_{circ}}{\left(\frac{Dice_{rect} + Dice_{circ}}{2}\right)} \quad (4.30)$$

The Sørensen–Dice index for geodesic and Chan-Vese methods using both rectangular and circular initial masks of the 12 sample images with all the 5 selected preprocessing methods was plotted in Figure 4-6 in two different ways, as bars in the 1st plot and continuous lines in the 2nd one. As the trendlines plotted with continuous lines are more visible than the bar plots, they were used to formulate the final conclusions. Also the difference between the rectangular and the circular initial masks (relative to the average of their Sørensen–Dice indices) was plotted in Figure 4-7.

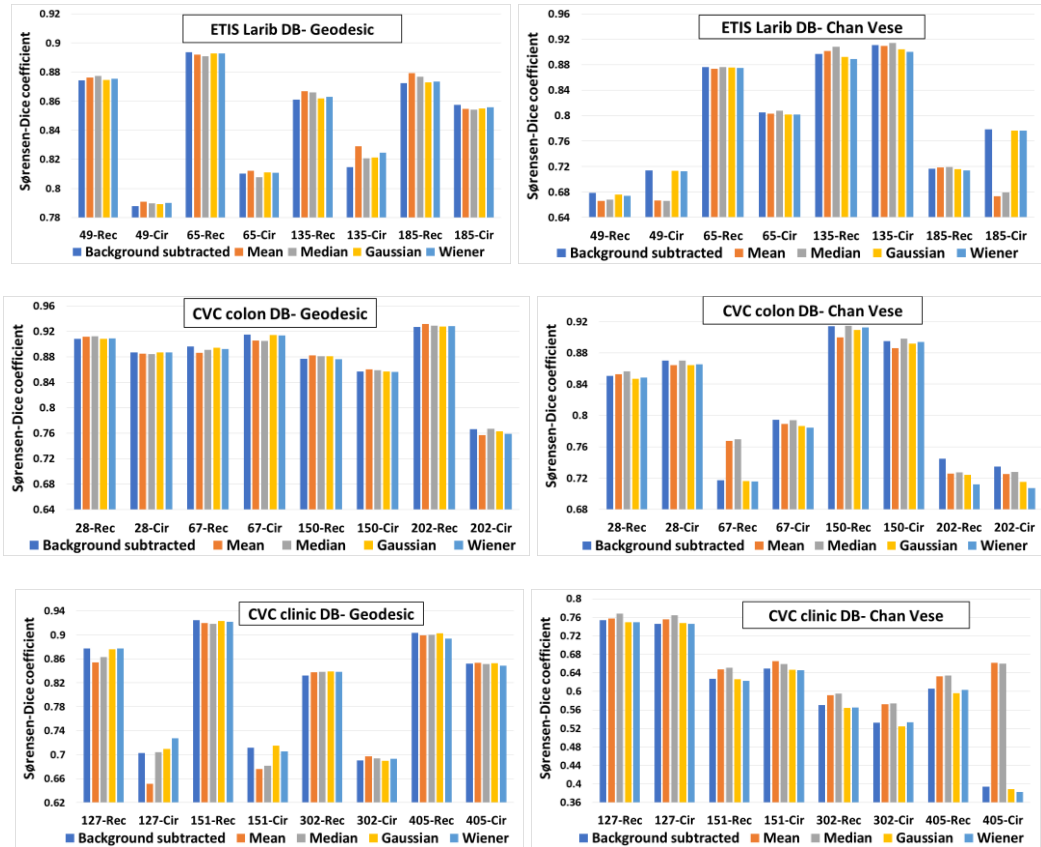


Figure 4-5 Sørensen–Dice similarity coefficient comparison results for geodesic and Chan-Vese methods using both rectangular and circular initial masks for the image samples of the three studied databases ( (Silva, Histace, Romain, Dray, & Granado, 2014); (Bernal, Sanchez, & Vilariño, 2012); (Bernal J. , et al., 2015)). The rectangular and circular masks are indicated by (Rec) and (Cir) respectively. The order of the columns’ colors is in the same order as in the bottom of the figures: background subtracted, mean, median, Gaussian and Wiener filtered images

The following observations can be formulated based on both Figure 4-6 and Figure 4-7

- a) The 5 preprocessing methods that are used for one image give almost the same results (the lines within the vertical gridlines are almost always close to horizontal in the 2nd plot of Figure 4-6).
- b) The Geodesic rectangular (the blue line in the 2nd plot of Figure 4-6) is the best almost always. Wherever the Chan-Vese method is better, all 4 results are about the same (Figure 4-7 has small columns in E-L 135 and Col 150).
- c) The Chan-Vese is less sensitive to the shape of the initial mask, i.e., the yellow and grey lines run much closer to each other in the 2nd plot of Figure 4-6, and the orange columns are usually smaller in Figure 4-7.

Table 4-2 Sørensen–Dice Coefficient of the post pre-processing case for the image samples in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014) using both rectangular and circular initial masks

Image Number	Active Contour Input Image	Geodesic Rectangular	Geodesic Circular	Chan–Vese Rectangular	Chan–Vese Circular
49	Background subtracted	0.8743	0.7879	0.6787	0.7137
	Mean	0.8763	0.7910	0.6661	0.6662
	Median	0.8773	0.7897	0.6678	0.6655
	Gaussian	0.8745	0.7892	0.6761	0.7129
	Wiener	0.8755	0.7902	0.6737	0.7122
65	Background subtracted	0.8937	0.8102	0.8763	0.8047
	Mean	0.8921	0.8121	0.8736	0.8028
	Median	0.8908	0.8079	0.8760	0.8074
	Gaussian	0.8929	0.8111	0.8755	0.8017
	Wiener	0.8927	0.8107	0.8747	0.8020
135	Background subtracted	0.8610	0.8148	0.8969	0.9111
	Mean	0.8669	0.8290	0.9016	0.9098
	Median	0.8660	0.8208	0.9080	0.9140
	Gaussian	0.8618	0.8212	0.8923	0.9044
	Wiener	0.8630	0.8247	0.8890	0.9001
185	Background subtracted	0.8724	0.8574	0.7167	0.7786
	Mean	0.8793	0.8547	0.7184	0.6730
	Median	0.8768	0.8542	0.7190	0.6791
	Gaussian	0.8730	0.8551	0.7155	0.7763
	Wiener	0.8735	0.8559	0.7137	0.7763



*Table 4-3 Sørensen–Dice Coefficient of the post pre-processing case for the image samples in database CVC-ColonDB (Bernal, Sanchez, & Vilariño, 2012) using both rectangular and circular initial masks*

<b>Image Number</b>	<b>Active Contour Input Image</b>	<b>Geodesic Rectangular</b>	<b>Geodesic Circular</b>	<b>Chan–Vese Rectangular</b>	<b>Chan–Vese Circular</b>
<b>28</b>	<b>Background subtracted</b>	0.9080	0.8867	0.8503	0.8702
	<b>Mean</b>	0.9117	0.8848	0.8527	0.8644
	<b>Median</b>	0.9122	0.8845	0.8562	0.8702
	<b>Gaussian</b>	0.9084	0.8867	0.8470	0.8641
	<b>Wiener</b>	0.9090	0.8870	0.8482	0.8651
<b>67</b>	<b>Background subtracted</b>	0.8960	0.9149	0.7172	0.7943
	<b>Mean</b>	0.8863	0.9053	0.7674	0.7893
	<b>Median</b>	0.8907	0.9049	0.7694	0.7938
	<b>Gaussian</b>	0.8941	0.9144	0.7162	0.7867
	<b>Wiener</b>	0.8925	0.9133	0.7157	0.7845
<b>150</b>	<b>Background subtracted</b>	0.8772	0.8572	0.9140	0.8948
	<b>Mean</b>	0.8823	0.8605	0.8998	0.8861
	<b>Median</b>	0.8808	0.8589	0.9146	0.8981
	<b>Gaussian</b>	0.8808	0.8572	0.9091	0.8916
	<b>Wiener</b>	0.8764	0.8565	0.9122	0.8940
<b>202</b>	<b>Background subtracted</b>	0.9266	0.7665	0.7447	0.7345
	<b>Mean</b>	0.9314	0.7567	0.7257	0.7252
	<b>Median</b>	0.9287	0.7666	0.7272	0.7279
	<b>Gaussian</b>	0.9274	0.7632	0.7242	0.7150
	<b>Wiener</b>	0.9283	0.7590	0.7118	0.7072

Table 4-4 Sørensen–Dice Coefficient of the post pre-processing case for the image samples in database CVC-ClinicDB (Bernal J. , et al., 2015) using both rectangular and circular initial masks

Image Number	Active Contour Input Image	Geodesic Rectangular	Geodesic Circular	Chan–Vese Rectangular	Chan–Vese Circular
127	Background subtracted	0.8769	0.7024	0.7541	0.7461
	Mean	0.8537	0.6510	0.7576	0.7558
	Median	0.8626	0.7039	0.7684	0.7646
	Gaussian	0.8757	0.7094	0.7493	0.7476
	Wiener	0.8770	0.7275	0.7492	0.7464
151	Background subtracted	0.9246	0.7117	0.6275	0.6496
	Mean	0.9198	0.6762	0.6473	0.6658
	Median	0.9185	0.6814	0.6514	0.6589
	Gaussian	0.9234	0.7152	0.6262	0.6465
	Wiener	0.9217	0.7055	0.6225	0.6455
302	Background subtracted	0.8317	0.6906	0.5705	0.5327
	Mean	0.8375	0.6970	0.5921	0.5723
	Median	0.8383	0.6939	0.5953	0.5739
	Gaussian	0.8391	0.6899	0.5641	0.5247
	Wiener	0.8383	0.6931	0.5654	0.5335
405	Background subtracted	0.9032	0.8522	0.6065	0.3941
	Mean	0.8990	0.8531	0.6330	0.6621
	Median	0.8998	0.8514	0.6342	0.6603
	Gaussian	0.9026	0.8526	0.5964	0.3891
	Wiener	0.8937	0.8482	0.6036	0.3832

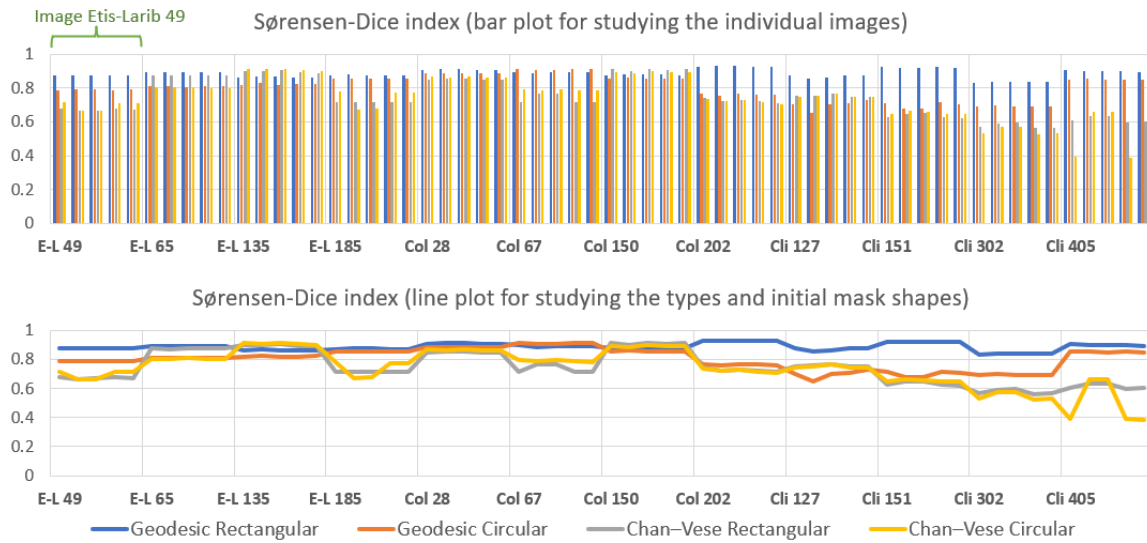


Figure 4-6 Sørensen–Dice coefficients for geodesic and Chan-Vese methods using both rectangular and circular initial masks of the 12 sample images with all the 5 selected preprocessing methods

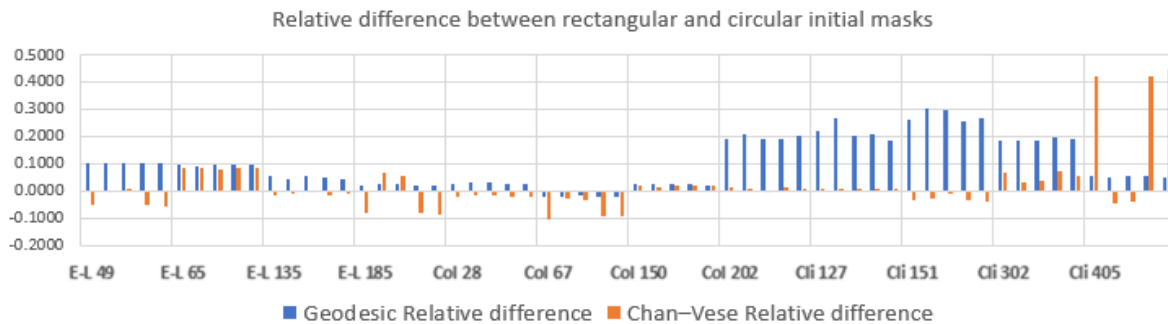


Figure 4-7 The relative difference of the Sørensen–Dice coefficients for geodesic and Chan-Vese methods using both rectangular and circular initial masks of the 12 sample images with all the 5 selected preprocessing methods

#### 4.4.6 Modified rectangular initial mask and tuned parameters case

In this case, the effect of the rectangular initial mask size with the tuned parameters was verified based on many experiments. Two different masks and different parameters were chosen to initialize the geodesic and Chan-Vese methods. The geodesic method was initialized using the same rectangular mask that encompasses the polyp area, while the Chan-Vese method was initialized using a decreased size rectangular mask located within the polyp area.

For the active contour function parameters, the smoothness factor plays a critical role in determining the smoothness of the resulting contour. By increasing the smoothness factor, the contour becomes smoother, reducing its susceptibility to noise and local variations in the image. Conversely, decreasing the smoothness factor enhances the contour's flexibility, allowing it to accurately fit to complex shapes' details and contours within the image. In other words, the smoothness factor involves a trade-off between generating a smoother segmented contour with a higher value vs. fitting finer details of the input image more precisely with a lower smoothness value.

The contraction bias parameter controls the expansion or contraction of the resulting contour. The contour contracts and fits tightly around the object boundaries in the image when the contraction bias parameter has a positive value. In contrast, if the contraction bias parameter has a negative value, the contour will extend and expand to cover more area inside the object boundaries.

The previously discussed two parameters were empirically selected. It was found that the same default parameters, i.e., a smoothness factor of 1 and a contraction bias of 0.3, continued to work well with the geodesic method. Whereas, the Chan-Vese method performed better with a smoothness factor equal to 1 and a contraction bias equal to  $-0.5$ . The performance of both methods was evaluated using the modified rectangular initial masks and different tuned parameters for 100, 200, and 300 iterations.

Figures 4-8, 4-9 and 4-10 illustrate the overall results of this case. Table 4-5 gives the numerical results for the image (151) of database CVC-ClinicDB (Bernal J. , et al., 2015), as an example.

In the case of 100 iterations and compared to the results of the post pre-processing case with rectangular initial mask, the results of the geodesic method remained unchanged in this sub-case, as the same rectangular initial mask and default parameters were used. For the Chan-Vese method, approximately 88.34% of the total similarity coefficients were enhanced with the modified rectangular initial mask and tuned parameters case.

For comparison, Figure 4-11 shows the Chan-Vese segmentation results of the image (135) in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014), in both cases, the post pre-processing with rectangular initial mask, and the modified rectangular initial mask and tuned parameters with 100 iterations.

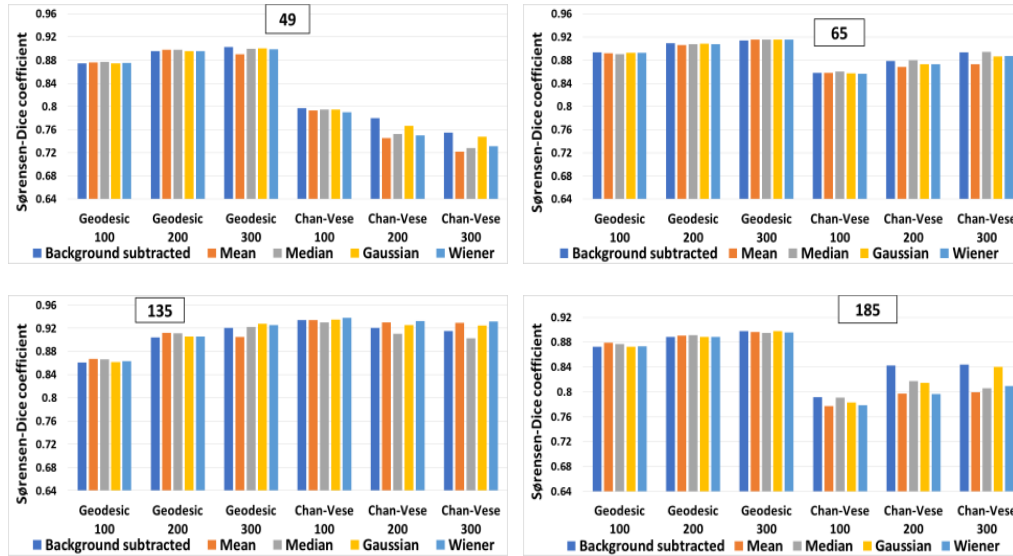


Figure 4-8 Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image samples in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014), with 100, 200, and 300 iterations. The order of the columns’ colors is in the same order as in the bottom of the figures: background subtracted, mean, median, Gaussian and Wiener filtered images

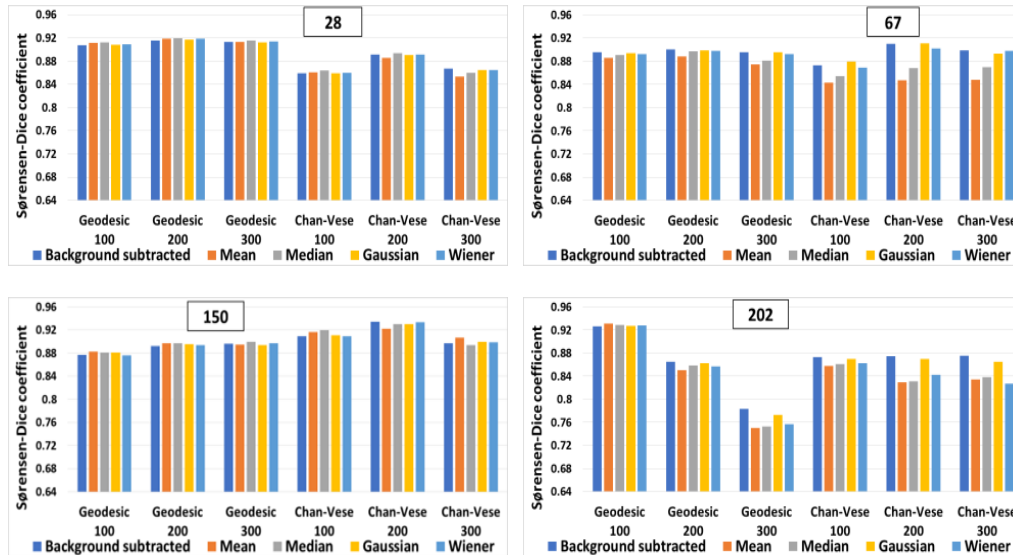


Figure 4-9 Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image samples in database CVC-ColonDB (Bernal, Sanchez, & Vilariño, 2012), with 100, 200, and 300 iterations. The order of the columns’ colors is in the same order as in the bottom of the figures: background subtracted, mean, median, Gaussian and Wiener filtered images

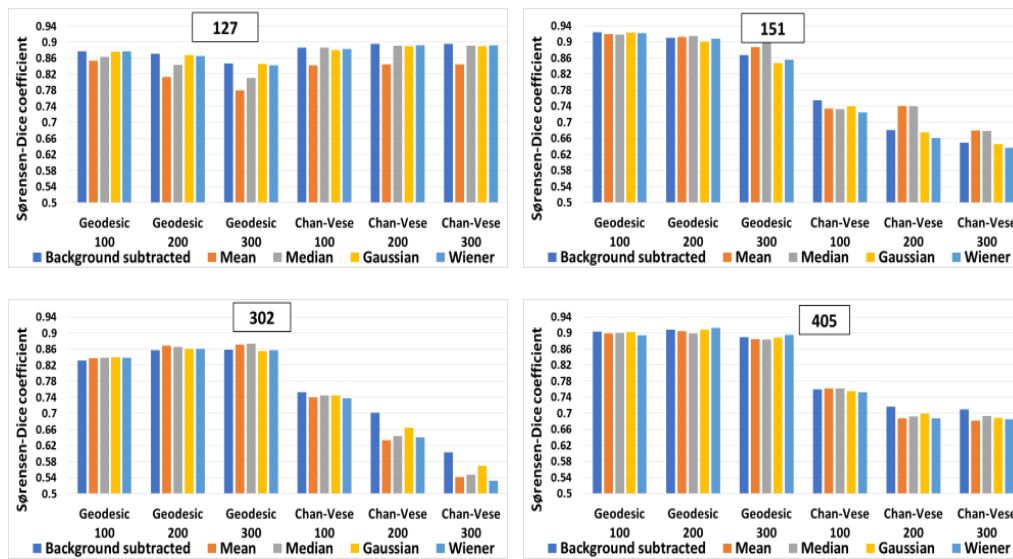
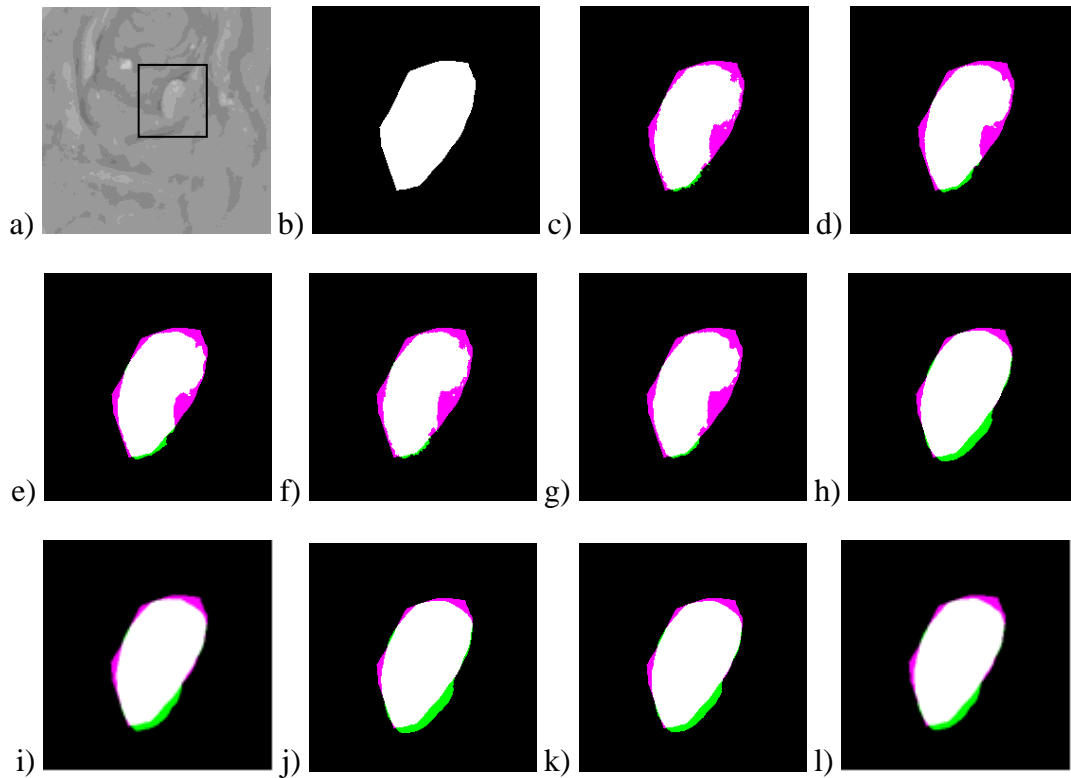


Figure 4-10 Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image samples in database CVC-ClinicDB (Bernal J. , et al., 2015), with 100, 200, and 300 iterations. The order of the columns’ colors is in the same order as in the bottom of the figures: background subtracted, mean, median, Gaussian and Wiener filtered images

Table 4-5 Numerical values for Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image (151) in database CVC-ClinicDB (Bernal J. , et al., 2015), with 100, 200, and 300 iterations. The highlighted values refer to the Chan–Vese improved values compared with the post pre-processing results given in the Table 4-4.

Image Number	Active Contour Input Image	Geodesic 100	Geodesic 200	Geodesic 300	Chan–Vese 100	Chan–Vese 200	Chan–Vese 300
151	Background subtracted	0.9246	0.9100	0.8674	0.7552	0.6810	0.6496
	Mean	0.9198	0.9127	0.8867	0.7338	0.7412	0.6792
	Median	0.9185	0.9154	0.8994	0.7332	0.7400	0.6778
	Gaussian	0.9234	0.9007	0.8473	0.7403	0.6753	0.6454
	Wiener	0.9217	0.9083	0.8554	0.7249	0.6607	0.6370

The results of the segmentation methods with higher number of iterations (200 and 300) varied according to the used pictures and to the different non-uniform characteristics of the databases. One of the most significant reasons for the inconsistent tendency of the results was the difference in the studied image size across the various databases, as well as the variation in polyp size compared to its image size. Additional factors such as the polyp shape (elliptical, circular, irregular), polyp viewpoint (top view, lateral, semi-lateral), and the strength of the



*Figure 4-11 The Chan–Vese comparison segmentation results of the image (135) in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014). (a) The original background subtracted image, with the zoomed part indicated by a black square around the polyp. (b) The given database mask. (c, e, g, i, k) The original (unfiltered image), mean, median, Gaussian, and Wiener filtered images results in the post pre-processing with rectangular initial mask case. (d, f, h, j, l) The original (unfiltered image), mean, median, Gaussian, and Wiener filtered images results in the modified rectangular initial mask and tuned parameters case with 100 iterations. The colors connotations: white is the matching area between the given database mask and the resulted segmented area, pink is the remaining area of the given mask that was not covered by the segmented area, and green is an additional area of the segmented area*

polyp's edges could also contribute to this irregular performance. Moreover, the masks given by the databases were drawn manually and some of them did not match the borderline of the polyps perfectly, thus if our method went closer to the ideal contour than the manually drawn mask, then the Sørensen–Dice metric showed a worse result than expected.

For the database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014) images which have the largest size, the improvement in the performance of the geodesic method with increasing the iterations number was evident in (38 out of 40) values of the total similarity coefficients in the case of 200 and 300 iterations. The Chan–Vese method improved just for (19) values of the total coefficients. However, as the decrease in the performance of the Chan-

Vese method in the case of image 49 was due to a reflection at the borderline of the polyp, a general conclusion can be drawn; the higher number of iterations is optimal for this database.

For the database CVC-ColonDB (Bernal, Sanchez, & Vilariño, 2012) images which have a medium size, most of the similarity coefficients values increased in the case of 200 iterations, (15 out of 20) and (17 out of 20) for the geodesic and Chan–Vese methods respectively. Then a decrease followed in the case of 300 iterations: (17 out of 20) and (15 out of 20) of the total similarity coefficients values decreased using the geodesic and Chan–Vese methods respectively. In this case, the 200 iterations seem to be an optimum. The exception for this tendency is the blurry image 202 with not strong enough edges, thus the geodesic method cannot really find optimum.

For the database CVC-ClinicDB (Bernal J. , et al., 2015) images which have the smallest size, 70% and 75% of the total similarity coefficients became lower in the case of applying 200 and 300 iterations compared to the case of 100 iterations with the geodesic and Chan–Vese methods respectively. Here, less clearly than in the previous case, the 100 iterations seem to be the optimal choice.

#### **4.4.7 Special cases**

Many cases were studied to extend the experiments and expand the field of investigation, by using additional samples with different characteristics and features. All of the following special cases were analyzed using the previously selected tuned parameters.

Firstly, it was studied what happens if the masks are not located around the polyps, by testing the performance of three alternative initial mask locations.

In the first case, the starting mask was initialized around some fairly smooth region far away from the polyp area to be segmented. This case produced undesired segmentation results, the geodesic method could not find any contour to shrink on, so it remained more or less the same as the initial mask. In the case of the Chan-Vese method, the resulting masks were mostly covering one or multiple regions between visible veins or folds, and the method failed both to find the polyp, and to indicate that the mask is located in a wrong area.

In the second case, when the initial and ground-truth masks only partially overlapped, the Chan-Vese method expanded toward the polyp area. However, if one or more, fairly homogeneous areas were included into the initial mask, it kept evolving towards these areas, too, and they remained as separate parts of the result. The geodesic method, similarly to the



first case (without overlap), did not change much compared to the initial mask, only on the side of the polyp, if the borderlines were strong enough.

If multiple polyps were present in the image, and the initial mask was one joint mask instead of two separate ones, the methods also failed to find the polyps. In the case of multiple polyps, at least one of the polyps had only partially visible borderlines, and very similar colour to the background. As more prominent edges and colour regions were due to the folding of the bowel than due to the polyps themselves, these methods were not suitable to separate them, even if theoretically they can distinguish multiple objects.

These studies showed the sensitivity of the methods to the initial mask location.

As a final step, the colour sensitivity was also studied. In the case of an advanced stage polyp, the polyp colour is different from the colour of the background, whereas in early stages, the polyp and the background usually have the same colour. Analysis showed that for darker colored polyps, even the background subtraction is not necessary, the methods find the polyps rather well for both the original, and the background subtracted images. As an example of an advanced stage colorectal polyp, the results for image 516 with dark polyp from the database CVC-ClinicDB (Bernal J. , et al., 2015) are shown in Figure 4-12. The other examples in this chapter were all on images with polyps having the same colour as the bowel around them.

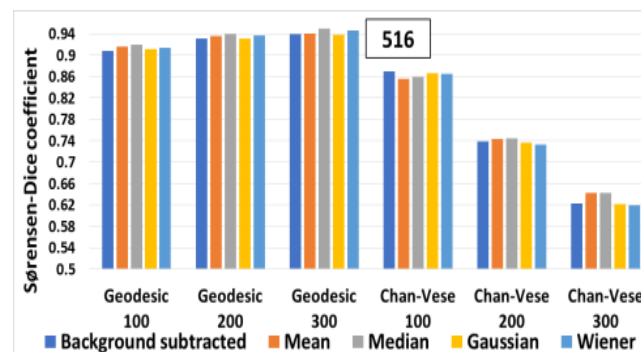


Figure 4-12 Darker polyp example: image 516 in database CVC-ClinicDB (Bernal J. , et al., 2015). Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case with 100, 200, and 300 iterations. Separate initial masks were used for the Chan–Vese and for the geodesic method. The order of the columns' colors is in the same order as in the bottom of the figures: background subtracted, mean, median, Gaussian and Wiener filtered images

## 4.5 Thesis statements

### 4.5.1 Thesis 5

For colorectal polyp segmentation purposes, I systematically evaluated the performance of the geodesic and Chan–Vese as two variational segmentation methods using image samples of 3 publicly available databases (including some special problematic cases).

- A. I tested the outputs of different pre-filtering techniques as potential inputs for both segmentation methods using Sørensen-Dice Index (SDI) metric. I found that the background-subtracted, mean, median, Gaussian, and Wiener filtered images are the best candidates for improving the segmentation results.
- B. I investigated the effect of the initial mask shape on the segmentation results using Sørensen-Dice index metric. Circular and rectangular initial masks were evaluated. I concluded that for the geodesic method in 91.67% of the cases the rectangular masks were more beneficial, while for the Chan Vese method, in 50% of the cases. The Chan-Vese method demonstrated lower sensitivity to the shape of the initial mask than the geodesic method.

The Geodesic method with the rectangular initial mask provided the best Sørensen-Dice index for most of the cases. If the Chan-Vese method gave better result, the difference between the geodesic and Chan–Vese Sørensen-Dice indices was very small.

- C. I studied the influence of the initial mask size together with the smoothness factor, contraction bias, and the number of iterations on the segmentation results. The geodesic approach achieved the best performance using an initial mask encompassing the boundaries of the polyp to be segmented with a contraction bias value which makes the contour shrink towards the polyp area, while the Chan–Vese approach accomplished the best results using an initial mask located within the polyp region with a contraction bias value which makes the contour expand towards the polyp. The suitable number of iterations ranged from 100 to 300 depending on different image and polyp sizes, although other factors including the strength of colorectal polyp edges or presence of veins also played a significant role.
- D. I concluded that, in most of the cases, the Chan–Vese method performed better than the geodesic method in matching the actual contour of the polyps, but it depended more on the effectiveness of the pre-processing procedure.

## 5. Outlook and future work

In the future the fuzzy classification method needs a more thorough investigation on the selection of the antecedents, probably also the introduction of some new antecedents, or using antecedents from both RGB and HSV images. It seems, that using the characteristic points of the  $\alpha = 0$  and  $\alpha = 1$  cuts of the rules with the same weight in the selection of the reduced number of antecedents is not effective enough, another weighting, or another selection criterion is needed. Using clustering methods, such as fuzzy c-means clustering can improve the fitting of the rules to the data of the training set.

Also, a study on the borderlines of the output classes based on their polyp content might be necessary, as an image segment with very small percentage of polyp might be dominated by the statistics of the polypless class for some of the antecedents, while in other antecedents the polyp contour's appearance might be the most important factor. Clustering methods might also improve the selection of the number of necessary consequents.

Moreover, as the number of available databases increased during my study, it is necessary to check the applicability of the method to the new databases.

In the future, a parallelization of the Hough transform is needed to decrease the calculation time and especially the optimization of its 3D local maximum search is necessary. As the fuzzy and classical transforms are built from very basic program elements, and no advanced components are used, it is worth implementing them in C, python or Cuda to allow to run them on parallel computers.

Finally, as this field of medical image processing is currently dominated by artificial intelligence methods, especially by neural network-based approaches, it might be necessary to develop and test alternatives of some components of my methods using neural networks.

## 6. Publications

- Ismail, R., & Nagy, S. (2023). A Novel Gradient-Weighted Voting Approach for Classical and Fuzzy Circular Hough Transforms and Their Application in Medical Image Analysis—Case Study: Colonoscopy. *Appl. Sci*, 13, 9066. <https://doi.org/10.3390/app13169066>.
- Ismail, R., Prukner, P., & Nagy, S. (2023). On Applying Gradient Based Thresholding on the Canny Edge Detection Results to Improve the Effectiveness of Fuzzy Hough Transform for Colonoscopy Polyp Detection Purposes. In R. M. Kountchev (Eds.), *New Approaches for Multidimensional Signal Processing. NAMSP 2022. Smart Innovation, Systems and Technologies* (Vol. 332, pp. 110-121). Springer: Singapore. [https://doi.org/10.1007/978-981-19-7842-5\\_10](https://doi.org/10.1007/978-981-19-7842-5_10).
- Ismail, R., Sziova, B., Taha, H., & Nagy, S. (2022). The effect of different consequent setting on the effectiveness of Kóczy-Hirota fuzzy rule interpolation in colonoscopy image classification. *2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, (pp. 1-4). Ohrid, North Macedonia. <https://doi.org/10.1109/ICEST55168.2022.9828693>.
- Ismail, R., & Nagy, S. (2022). Ways of improving of active contour methods in colonoscopy image segmentation. *Image Analysis & Stereology*, 41, 7–23. <https://doi.org/10.5566/ias.2604>.
- Ismail, R., & Nagy, S. (2021). On Metrics Used in Colonoscopy Image Processing for Detection of Colorectal Polyps. In R. M. Kountchev (Ed.), *New Approaches for Multidimensional Signal Processing. Smart Innovation, Systems and Technologies* (Vol. 216, pp. 137-151). Springer: Singapore. [https://doi.org/10.1007/978-981-33-4676-5\\_10](https://doi.org/10.1007/978-981-33-4676-5_10).
- Nagy, S., Ismail, R., Sziová, B., & Kóczy, L. (2021). On classical and fuzzy Hough transform in colonoscopy image processing. *IEEE AFRICON 2021, Virtual Conference*, (pp. 124–129). 13 - 15 September, Arusha, Tanzania. <https://doi.org/10.1109/AFRICON51333.2021.9570897>.
- Sziová, B., Ismail, R., Lilik, F., Kóczy, L., & Nagy, S. (2020). Fuzzy rulebase parameter determination for stabilized KH interpolation based detection of colorectal polyps on colonoscopy images. *IEEE International Conference on Fuzzy Systems (FUZZ IEEE)*, (pp. 1-6). Glasgow, United Kingdom. <https://doi.org/10.1109/FUZZ48607.2020.9177839>.

---

## References

- Adams, R., & Bischof, L. (1994). Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(6), 641-647. doi:10.1109/34.295913
- Ahmad, O., Brandao, P., Sami, S., & et al. (2019). Artificial intelligence for real-time polyp localization in colonoscopy withdrawal videos. *Gastrointestinal Endoscopy*, *89*, AB647.
- Alam, M., & Fattah, S. (2023). SR-AttNet: An Interpretable Stretch-Relax Attention based Deep Neural Network for Polyp Segmentation in Colonoscopy Images. *Computers in biology and medicine*, *160*, 106945. <https://doi.org/10.1016/j.compbimed.2023.106945>
- Aliyi, S., Dese, K., & Raj, H. (2023). Detection of gastrointestinal tract disorders using deep learning methods from colonoscopy images and videos. *Scientific African*, *20*, e01628. <https://doi.org/10.1016/j.sciaf.2023.e01628>.
- Ballard, D. (1981). Generalizing the Hough Transform to detect arbitrary shapes. *Pattern Recognition*, *13*, 111-122. [https://doi.org/10.1016/0031-3203\(81\)90009-1](https://doi.org/10.1016/0031-3203(81)90009-1).
- Barbosa, W., & Vieira, A. (2019). On the Improvement of Multiple Circles Detection from Images using Hough Transform. *TEMA (São Carlos)*, *20*(2), 331-342. <https://doi.org/10.5540/tema.2019.020.02.0331>
- Bernal J., et al. (2017). Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results From the MICCAI 2015 Endoscopic Vision Challenge. *IEEE Transactions on Medical Imaging*, *36*(6), 1231-1249. doi:10.1109/TMI.2017.2664042
- Bernal, J., Sanchez, F., & Vilariño, F. (2012). Towards automatic polyp detection with a polyp appearance model. *Pattern Recognition*, *45*(9), 3166-3182. <https://doi.org/10.1016/j.patcog.2012.03.002>
- Bernal, J., Sanchez, F., Fernández-Esparrach, G., Gil, D., Rodríguez, C., & Vilariño, F. (2015). WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics*, *43*, 99-111. doi:10.1016/j.compmedimag.2015.02.007
- Bhat, S., & Kumar, P. (2019). Segmentation of Optic Disc by Localized Active Contour Model in Retinal Fundus Image. In S. T. Tiwari (Ed.), *Smart Innovations in Communication and Computational Sciences. Advances in Intelligent Systems and Computing*. 851, pp. 35-44. Springer, Singapore. [https://doi.org/10.1007/978-981-13-2414-7\\_4](https://doi.org/10.1007/978-981-13-2414-7_4)

- Billah, M., Waheed, S., & Rahman, M. (2017). An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features. *Int J Biomed Imaging*, 9545920. doi:10.1155/2017/9545920
- Bonyár, A. (2016). AFM characterization of the shape of surface structures with localization factor. *Micron*, 87, 1-9. <https://doi.org/10.1016/j.micron.2016.05.002>
- Bonyár, A., Molnár, L., & Harsányi, G. (2012). Localization factor: A new parameter for the quantitative characterization of surface structure with atomic force microscopy (AFM). *Micron*, 43(2-3), 305-310. <https://doi.org/10.1016/j.micron.2011.09.005>
- Brice, C., & Fennema, C. (1970). Scene analysis using regions. *Artificial Intelligence*, 1(3-4), 205-226. [https://doi.org/10.1016/0004-3702\(70\)90008-1](https://doi.org/10.1016/0004-3702(70)90008-1)
- Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6), 679–698. doi:10.1109/TPAMI.1986.4767851
- Carass, A., Roy, S., Gherman, A., & et al. (2020). *Evaluating White Matter Lesion Segmentations with Refined Sørensen-Dice Analysis*. <https://doi.org/10.1038/s41598-020-64803-w>
- Caselles, V., Kimmel, R., & Sapiro, G. (1997). Geodesic Active Contours. *International Journal of Computer Vision* 22, 61–79. <https://doi.org/10.1023/A:1007979827043>
- Chan, T., & Vese, L. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2), 266-277. doi:10.1109/83.902291
- Chan, T., & Vese, L. (2000). *Image segmentation using level sets and the piecewise-constant Mumford-Shah model*. Computational Applied Math Group.
- Chen, J., Qiang, H., Wu, J., Xu, G., & Wang, Z. (2021). Navigation path extraction for greenhouse cucumber-picking robots using the prediction-point Hough transform. *Computers and Electronics in Agriculture*, 180, 105911. <https://doi.org/10.1016/j.compag.2020.105911>.
- Chuquimia, O., Pinna, A., Dray, X., & Granado, B. (2020). A Low Power and Real-Time Architecture for Hough Transform Processing Integration in a Full HD-Wireless Capsule Endoscopy. *IEEE transactions on biomedical circuits and systems.*, 14, 646–657. <https://doi.org/10.1109/TBCAS.2020.3008458>.
- Csimadia, G., & Nagy, S. (2014). The Effect of the Contrast Enhancement Processes on the Structural Entropy of Colonoscopic Images. *ICEST 2014*. June 25-27, Nis, Serbia.

- Cucchiara, R., & Filicori, F. (1998). The Vector-Gradient Hough Transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, 20, 746–750. <https://doi.org/10.1109/34.689304>.
- Dervieux, A., & Thomasset, F. (1979). A finite element method for the simulation of a Rayleigh-Taylor instability. In R. Rautmann (Ed.), *Approximation Methods for Navier–Stokes Problems. Lecture Notes in Mathematics* (pp. 145–158). Berlin, Heidelberg: Springer. <https://doi.org/10.1007/BFb0086904>
- Djekoune, A., Messaoudi, K., & Amara, K. (2017). Incremental circle hough transform: An improved method for circle detection. *Optik*, 133, 17–31. <https://doi.org/10.1016/j.ijleo.2016.12.064>
- Dutta, S., Sasmal, P., Bhuyan, M., & Iwahori, Y. (2018). Automatic Segmentation of Polyps in Endoscopic Image Using Level-Set Formulation. *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, (pp. 1-5). Chennai, India. doi:10.1109/WiSPNET.2018.8538615
- Fang , L., Pan, X., Yao, Y., Zhang, L., & Guo , D. (2020). A hybrid active contour model for ultrasound image segmentation. *Soft Comput* 24, 18611–18625. <https://doi.org/10.1007/s00500-020-05097-y>
- Georgieva, V., & Petrov, P. (2017). An Approach for Colorectal Polyp Segmentation. *Conference on Communication, Electromagnetics and Medical Application (CEMA'2017)*. Sofia, Bulgaria.
- Georgieva, V., Nagy, S., Kamenova, E., & Horváth, A. (2015). An Approach for pit pattern recognition in colonoscopy images. *Egyptian Computer Science Journal*, 39(2), 72–82.
- Georgieva, V., Petrov, P., Nagy, S., & Sziová, B. (2018). Detecting contours of pathological forms in colonoscopy images using a hybrid method. *Communication, Electromagnetics and Medical Application (CEMA'2018)*. Sofia, Bulgaria.
- Häfner, M., Gangl, A., Liedlgruber, M., Uhl, A., Vécsei, A., & Wrba, F. (2010). Classification of endoscopic images using delaunay triangulation-based edge features. *Campilho, A., Kamel, M. (eds) Image Analysis and Recognition. ICIAR 2010. Lecture Notes in Computer Science, Springer*. Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-13775-4\\_14](https://doi.org/10.1007/978-3-642-13775-4_14)
- Han, J., Kóczy, L., & Poston, T. (1994). Fuzzy Hough Transform. *Pattern Recognition Letters*, 44, 649–658. [https://doi.org/10.1016/0167-8655\(94\)90068-X](https://doi.org/10.1016/0167-8655(94)90068-X).

- Hapsari, R., Utoyo, M., Rulaningtyas, R., & Suprajitno, H. (2020). Iris segmentation using Hough Transform method and Fuzzy C-Means method. *Journal of Physics: Conference Series.*, 1477, 022– 037. doi:10.1088/1742-6596/1477/2/022037
- Horváth, A., Spindler, S., Szalai, M., & Rácz, I. (2016). Preprocessing endoscopic images of colorectal polyps. *Acta Technica Jaurinensis*, 9(1), 65–82.
- Hough, P. (1959). Machine Analysis of Bubble Chamber Pictures. In *Proceedings of the 2nd International Conference on High Energy Accelerators and Instrumentation, HEACC1959*. Geneva, Switzerland: CERN.
- Ismail, R., & Nagy, S. (2021). On Metrics Used in Colonoscopy Image Processing for Detection of Colorectal Polyps. In R. M. Kountchev (Ed.), *New Approaches for Multidimensional Signal Processing. Smart Innovation, Systems and Technologies* (Vol. 216, pp. 137-151). Springer: Singapore. [https://doi.org/10.1007/978-981-33-4676-5\\_10](https://doi.org/10.1007/978-981-33-4676-5_10).
- Ismail, R., & Nagy, S. (2022). Ways of improving of active contour methods in colonoscopy image segmentation. *Image Analysis & Stereology*, 41, 7–23. <https://doi.org/10.5566/ias.2604>.
- Ismail, R., & Nagy, S. (2023). A Novel Gradient-Weighted Voting Approach for Classical and Fuzzy Circular Hough Transforms and Their Application in Medical Image Analysis—Case Study: Colonoscopy. *Appl. Sci*, 13, 9066. <https://doi.org/10.3390/app13169066>
- Ismail, R., Prukner, P., & Nagy, S. (2023). On Applying Gradient Based Thresholding on the Canny Edge Detection Results to Improve the Effectiveness of Fuzzy Hough Transform for Colonoscopy Polyp Detection Purposes. In R. M. Kountchev (Ed.), *New Approaches for Multidimensional Signal Processing. NAMSP 2022. Smart Innovation, Systems and Technologies* (Vol. 332, pp. 110-121). Springer: Singapore. [https://doi.org/10.1007/978-981-19-7842-5\\_10](https://doi.org/10.1007/978-981-19-7842-5_10).
- Ismail, R., Sziova, B., Taha, H., & Nagy, S. (2022). The effect of different consequent setting on the effectiveness of Kóczy-Hirota fuzzy rule interpolation in colonoscopy image classification. *2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, (pp. 1-4). Ohrid, North Macedonia. 10.1109/ICEST55168.2022.9828693
- Jha, D., Ali, S., Tomar, N., Håvard, D., Johansen, D., Rittscher, J., . . . Halvorsen, P. (2021). Real-Time Polyp Detection, Localization and Segmentation in Colonoscopy Using Deep Learning. *IEEE Access*, 9, 40496-40510. doi:10.1109/ACCESS.2021.3063716



- Kalbasi, M., & Nikmehr, H. (2020). Noise-Robust, Reconfigurable Canny Edge Detection and its Hardware Realization. *IEEE Access*, 8, 39934 – 39945. <https://doi.org/10.1109/ACCESS.2020.2976860>
- Karaman, A., Pacal, I., Basturk, A., Akay, B., Nalbantoglu, U., Coskun, S., & et al. (2023). Robust real-time polyp detection system design based on YOLO algorithms by optimizing activation functions and hyper-parameters with Artificial Bee Colony (ABC). *Expert Systems with Applications*, 221, 119741. <https://doi.org/10.1016/j.eswa.2023.119741>.
- Karkanis, S., Iakodivis, D., Karras, D., & Maroulis, D. (2001). Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures. *Proceedings of the IEEE International Conference in Image Processing*, 2, pp. 833–836. Thessaloniki, Greece. doi:10.1109/ICIP.2001.958623.
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active Contour Models. *International Journal of Computer Vision*, 321–331. <https://doi.org/10.1007/BF00133570>
- Kichenassamy, S., Kumar, A., Olver, P., Tannenbaum, A., & Yezzi, A. (1995). Gradient flows and geometric active contour models. *IEEE International Conference on Computer Vision*, (pp. 810-815). Cambridge, MA, USA. doi:10.1109/ICCV.1995.466855.
- Koczy, L., & Hirota, K. (1993). Approximate reasoning by linear rule interpolation and general approximation. *Int J Approx Reason*, 9(3), 197–225. [https://doi.org/10.1016/0888-613X\(93\)90010-B](https://doi.org/10.1016/0888-613X(93)90010-B)
- Krenzer, A., Banck, M., Makowski, K., Hekalo, A., Fitting, D., Troya, J., . . . Puppe, F. (2023). A Real-Time Polyp Detection System with Clinical Application in Colonoscopy Using Deep Convolutional Neural Networks. *J. Imaging*, 9(26). <https://doi.org/10.3390/jimaging9020026>
- Kudo, S., Hirota, S., Nakajima, T., Hosobe, S., & Kusaka, H. (1994). Colorectal tumours and pit pattern. *J Clin Pathol*, 47(10), 880-885. doi:doi: 10.1136/jcp.47.10.880.
- Lei, X., & Erkki, O. (1993). Randomized Hough Transform (RHT): Basic Mechanisms, Algorithms, and Computational Complexities. *CVGIP Image Underst*, 57, 131–154. <https://doi.org/10.1006/ciun.1993.1009>
- Li, J., Yang, L., Qu, Y., & Sexton, G. (2018). An extended Takagi–Sugeno–Kang inference system (TSK+) with fuzzy interpolation and its rule base generation. *Soft Comput*, 22, 3155–3170. <https://doi.org/10.1007/s00500-017-2925-8>

- Lin, G. T. (2020). Fruit detection in natural environment using partial shape matching and probabilistic Hough transform. *Precision Agric*, 21, 160–177. <https://doi.org/10.1007/s11119-019-09662-w>.
- Liu, W., Zhang, Z., Li, S., & Tao, D. (2017). Road Detection by Using a Generalized Hough Transform. *MDPI, Remote Sensing*, 9, 590. <https://doi.org/10.3390/rs9060590>.
- Maini, R., & Aggarwal, H. (2009). Study and Comparison of Various Image Edge Detection Techniques. *International Journal of Image Processing (IJIP)*, 3(1), 1 - 11.
- Mamdani, E. (1977). Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Transactions on Computers*, 26(12), 1182-1191. doi:10.1109/TC.1977.1674779.
- Mathavan, S., Vaheesan, K., Kumar, A., Chandrakumar, C., Kamal, K., Rahman, M., & Stonecliffe-Jones, M. (2017). Detection of pavement cracks using tiled fuzzy Hough Transform. *J. Electron. Imaging*, 26, 053008. <https://doi.org/10.1117/1.JEI.26.5.053008>.
- Mo, X., Tao, K., Wang, Q., & Wang, G. (2018). An efficient approach for polyps detection in endoscopic videos based on faster R-CNN. *24th International Conference on Pattern Recognition (ICPR)*, (pp. 3929-3934). Beijing, China. doi:10.1109/ICPR.2018.8545174
- Molnár, L., Nagy, S., & Mojzes, I. (2009). Structural entropy in detecting background patterns of AFM images. *Vacuum*, 84(1), 179-183. <https://doi.org/10.1016/j.vacuum.2009.04.025>
- Montseny, E., Sobrevilla, P., & Marès Martí, P. (2003). Edge orientation-based fuzzy Hough transform (EOFHT). *The 3rd Conference of the European Society for Fuzzy Logic and Technology*. 10–12, September, Zittau, Germany.
- Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied*, 42(5), 577-685. doi:10.1002/cpa.3160420503
- Nagy, S., Ismail, R., Sziová, B., & Kóczy, L. (2021). On classical and fuzzy Hough transform in colonoscopy image processing. *IEEE AFRICON 2021, Virtual Conference*, (pp. 124–129). 13 - 15 September, Arusha, Tanzania. <https://doi.org/10.1109/AFRICON51333.2021.9570897>
- Nagy, S., Kovács, M., Sziová, B., & Kóczy, L. (2019). Fuzzy Hough Transformation in aiding computer tomography-based liver diagnosis. *2019 IEEE AFRICON*. 25-27 September, Accra, Ghana. doi:10.1109/AFRICON46755.2019.9133793

- Nagy, S., Lilik, F., & Kóczy, L. (2017). Entropy based fuzzy classification and detection aid for colorectal polyps. *IEEE AFRICON*, (pp. 78-82). Cape Town, South Africa. doi:10.1109/AFRCON.2017.8095459.
- Nagy, S., Solecki, L., Sziova, B., Sarkadi-Nagy, B., & Kóczy, L. (2020). Applying Fuzzy Hough Transform for Identifying Honed Microgeometrical Surfaces. In L. Kóczy, J. Medina-Moreno, E. Ramírez-Poussa, & A. Šostak (Eds.), *Computational Intelligence and Mathematics for Tackling Complex Problems. Studies in Computational Intelligence* (Vol. 819, pp. 35–42). Springer, Cham. [https://doi.org/10.1007/978-3-030-16024-1\\_5](https://doi.org/10.1007/978-3-030-16024-1_5).
- Nagy, S., Sziova, B., & Koczy, L. (2017). The effect of wavelet analysis on entropy based fuzzy classification of colonoscopy images. *5th International Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII 2017)*, (pp. 2-5). Beijing, China.
- Nagy, S., Sziova, B., & Koczy, L. (2018). The effect of image feature qualifiers on fuzzy colorectal polyp detection schemes using KH interpolation - towards hierarchical fuzzy classification of coloscopic still images. *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, (pp. 1-7). Rio de Janeiro, Brazil. doi:10.1109/FUZZ-IEEE.2018.8491479
- Nagy, S., Sziova, B., & Pipek, J. (2019). On Structural Entropy and Spatial Filling Factor Analysis of Colonoscopy Pictures. *Entropy*, 21(3), 256. <https://doi.org/10.3390/e21030256>
- Nahum, K., Eldar, Y., & Bruckstein, A. (1991). A probabilistic Hough Transform. *Pattern Recognit*, 24, 303–316. [https://doi.org/10.1016/0031-3203\(91\)90073-E](https://doi.org/10.1016/0031-3203(91)90073-E)
- Osher, S., & Sethian, J. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1), 12-49. [https://doi.org/10.1016/0021-9991\(88\)90002-2](https://doi.org/10.1016/0021-9991(88)90002-2)
- Pipek, J., & Varga, I. (1992). Universal classification scheme for the spatial-localization properties of one-particle states in finite, d-dimensional systems. *Phys Rev A*, 46(6), 3148-3163. doi:10.1103/physreva.46.3148. PMID: 9908483
- Pogorelov, K., Randel, K., Griwodz, C., Eskeland, S., de Lange, T., Johansen, D., & et, a. (2017). KVASIR: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference 2017*, (pp. 164–169). Taipei, Taiwan. <https://doi.org/10.1145/3083187.3083212>.

- Prewitt, J. (1970). Object enhancement and extraction. *Picture Processing and Psychopictorics, 1st ed.*; Lipkin, B., Rosenfeld, A., Eds.; Academic Press, New York, NY, USA, 75–149.
- Pugin, E., Zhiznyakov, A., & Zakharov, A. (2018). Pipes Localization Method Based on Fuzzy Hough Transform. In A. K. Abraham (Ed.), *Advances in Intelligent Systems and Computing, Proceedings of the Second International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’17)* (pp. 536–544). Springer: Cham, Switzerland. [https://doi.org/10.1007/978-3-319-68321-8\\_56](https://doi.org/10.1007/978-3-319-68321-8_56).
- Rácz, I., Horváth, A., Szalai, M., Spindler, S., Kiss, G., Regöczi, H., & Horváth, Z. (2015). Digital image processing software for predicting the histology of small colorectal polyps by using narrow-band imaging magnifying colonoscopy. *Gastrointestinal Endoscopy, 81*(5), AB259. doi:10.1016/j.gie.2015.03.1345
- Rácz, I., Jánoki, M., & Saleh, H. (2010). Colon cancer detection by rendezvous colonoscopy: successful removal of stuck colon capsule by conventional colonoscopy. *Case Rep Gastroenterol, 4*(1), 19-24. doi:10.1159/000251663
- Rényi, A. (1960). On measures of information and entropy. *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, 4*, pp. 547-561. Berkeley, CA, USA.
- Roberts, L. (1965). *Machine perception of 3-D solids*. Ph.D. thesis, Massachusetts Institute of Technology, Department of Electrical Engineering. Cambridge, MA, USA.
- Ruano, L., Barrera, C., Bravo, D., Gomez, M., & Romero, E. (2019). Localization of Small Neoplastic Lesions in Colonoscopy by Estimating Edge, Texture and Motion Saliency. *In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 5945–5948). Berlin, Germany. <https://doi.org/10.1109/EMBC.2019.8856864>.
- Ruiz, L., Guayacán, L., & Martínez, F. (2019). Automatic polyp detection from a regional appearance model and a robust dense Hough coding. *In Proceedings of the 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, (pp. 1-5). Bucaramanga, Colombia. <https://doi.org/10.1109/STSIVA.2019.8730270>.
- Sánchez-Peralta, L., Bote-Curiel, L., Picón, A., Sánchez-Margallo, F., & Pagador, J. (2020). Deep learning to find colorectal polyps in colonoscopy: A systematic literature review. *Artificial Intelligence in Medicine, 108*:101923. doi:10.1016/j.artmed.2020.101923

- Sasmal, P., Iwahori, Y., Bhuyan, M., & Kasugai, K. (2018). Active contour segmentation of polyps in capsule endoscopic images. *International Conference on Signals and Systems (ICSigSys)*, (pp. 201-204). Bali, Indonesia. doi:10.1109/ICSIGSYS.2018.8372666
- Shaaf, Z., Jamil, M., & Ambar, R. (2022). Automatic Localization of the Left Ventricle from Short-Axis MR Images Using Circular Hough Transform. In M. Kaiser, K. Ray, A. Bandyopadhyay, K. Bandyopadhyay, & K. Long (Eds.), *Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering, Lecture Notes in Networks and Systems* (pp. 501–508). Springer, Singapore. [https://doi.org/10.1007/978-981-16-7597-3\\_41](https://doi.org/10.1007/978-981-16-7597-3_41)
- Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.
- Silva, J., Histace, A., Romain, O., Dray, X., & Granado, B. (2014). Towards embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. *Int J Comput Assisted Radiology and Surgery*, 9, 283-293.
- Sobel, I. (1978). Neighborhood coding of binary images for fast contour following and general binary array processing. *Computer Graphics and Image Processing*, 8127-8135. [https://doi.org/10.1016/S0146-664X\(78\)80020-3](https://doi.org/10.1016/S0146-664X(78)80020-3)
- Solecki, L., Nagy, S., & Sziová, B. (2020). The effect of background and outlier subtraction on the structural entropy of two-dimensional measured data. *International Journal of Reasoning-based Intelligent Systems*, 12(3), 200. doi:10.1504/IJRIS.2020.10031843
- Sornapudi, S., Meng, F., & Yi, S. (2019). Region-based automated localization of colonoscopy and wireless capsule endoscopy polyps. *Appl. Sci.*, 9, 2404. <https://doi.org/10.3390/app9122404>.
- Sziová, B., Ismail, R., Lilik, F., Kóczy, L., & Nagy, S. (2020). Fuzzy rulebase parameter determination for stabilized KH interpolation based detection of colorectal polyps on colonoscopy images. *IEEE International Conference on Fuzzy Systems (FUZZ IEEE)*, (pp. 1-6). Glasgow, United Kingdom. doi:10.1109/FUZZ48607.2020.9177839
- Sziová, B., Nagy, S., & Fazekas, Z. (2021). Application of Structural Entropy and Spatial Filling Factor in Colonoscopy Image Classification. *Entropy*, 23(8), 936. <https://doi.org/10.3390/e23080936>
- Sziová, B., Nagy, S., & Kóczy, L. (2021). The Effects of Preprocessing on Colorectal Polyp Detecting by Fuzzy Algorithm. In *Recent Developments and the New Direction in Soft-Computing Foundations and Applications. Studies in Fuzziness and Soft Computing*

- (Vol. 393, pp. 347–357). Springer, Cham.  
[https://link.springer.com/chapter/10.1007/978-3-030-47124-8\\_28](https://link.springer.com/chapter/10.1007/978-3-030-47124-8_28)
- Taha, A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging* 15, 29. <https://doi.org/10.1186/s12880-015-0068-x>
- Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-15*(1), 116-132. doi:10.1109/TSMC.1985.6313399
- Tan, J., Gao, Y., Liang, Z., Cao, W., Pomeroy, M., Huo Y.M., . . . Pickhardt, P. (2020). 3D-GLCM CNN: A 3-Dimensional Gray-Level Co-Occurrence Matrix-Based CNN Model for Polyp Classification via CT Colonography. *IEEE Transactions on Medical Imaging*, 39, 2013-2024. doi:10.1109/TMI.2019.2963177
- Tikk, D., Joo, I., Koczy, L., Varlaki, P., Moser, B., & Gedeon, T. (2002). Stability of interpolative fuzzy KH-controllers. *Fuzzy Sets and Systems*, 125(1), 105-119. doi:10.1016/S0165-0114(00)00104-4
- Varga, I., & Pipek, J. (2003). Rnyi entropies characterizing the shape and the extension of the phase space representation of quantum wave functions in disordered systems. *Phys Rev E Stat Nonlin Soft Matter Phys*, 68(2), 026202. doi:10.1103/PhysRevE.68.026202
- Vijayarajeswari, R., Parthasarathy, P., Vivekanandan, S., & Basha, A. (2019). Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement*, 146, 800–805. <https://doi.org/10.1016/j.measurement.2019.05.083>
- Vleugels, J., Hazewinkel, Y., & Dekker, E. (2017). Morphological classifications of gastrointestinal lesions. *Best Practice & Research Clinical Gastroenterology*, 31(4), 359-367. <https://doi.org/10.1016/j.bpg.2017.05.005>
- Wittenberg, T., Zobel, P., Rathke, M., & Mühlendorfer, S. (2019). Computer aided detection of polyps in whitelight-colonoscopy images using deep neural networks. *Current Directions in Biomedical Engineering*, 5, 231–234. <https://doi.org/10.1515/cdbme-2019-0059>.
- Xiao, W., Chang, L., & Liu, W. (2018). Semantic Segmentation of Colorectal Polyps with DeepLab and LSTM Networks. *IEEE International Conference on Consumer Electronics*, (pp. 1-2). Taichung, Taiwan. doi:10.1109/ICCE-China.2018.8448568.

- Yang, X., & Jiang, X. (2020). A Hybrid Active Contour Model based on New Edge-Stop Functions for Image Segmentation. *International Journal of Ambient Computing and Intelligence (IJACI)*, 11(1), 87-98.
- Yao, H., Stidham, R., Soroushmehr, R., Gryak, J., & Najarian, K. (2019). Automated Detection of Non-Informative Frames for Colonoscopy Through a Combination of Deep Learning and Feature Extraction. *In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (pp. 2402–2406). Berlin, Germany. <https://doi.org/10.1109/EMBC.2019.8856625>.
- Yue, G., Han, W., Li, S., Zhou, T., Lv, J., & Wang, T. (2022). Automated polyp segmentation in colonoscopy images via deep network with lesion-aware feature selection and refinement. *Biomed. Signal Process. Control.*, 78, 103846. <https://doi.org/10.1016/j.bspc.2022.103846>
- Yuji, I., Akira, H., Yoshinori, A., Bhuyan, M., Robert, J., & Kunio, K. (2015). Automatic detection of polyp using hessian filter and HOG features. *Procedia Computer Science*, 60, 730-739. <https://doi.org/10.1016/j.procs.2015.08.226>
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- Zadeh, L. (1968). Fuzzy algorithms. *Information and Control*, 12(2), 94–102. [https://doi.org/10.1016/S0019-9958\(68\)90211-8](https://doi.org/10.1016/S0019-9958(68)90211-8)
- Zhao, K., Han, Q., Zhang, C., Xu, J., & Cheng, M. (2022). Deep Hough Transform for Semantic Line Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44, 4793–4806. <https://doi.org/10.1109/TPAMI.2021.3077129>.

## List of figures

Figure 2-1 Examples of crisp and fuzzy membership functions .....	9
Figure 2-2 Fuzzy inference system based on 2 antecedents.....	10
Figure 2-3 True positive rates and false positive rates of the stabilized Kóczy-Hirota fuzzy rule interpolation-based colonoscopy image classification .....	15
Figure 2-4 Examples of the RGB color space histograms together with the triangular rules resulting from the three rulebase generation methods of the measured data in the training sets for the three databases .....	20
Figure 3-1 A sample image (220) from database CVC-Colon (Bernal, Sanchez, & Vilariño, 2012) for demonstrating the Hough transform steps .....	40
Figure 3-2 The Hough transformed images of the edge detected picture in subplot (c) of Figure 3-1 .....	41
Figure 3-3 The resulting final circles from the inverse Hough transform of the picture in Figure 3-1 .....	42
Figure 3-4 First 5 steps of the edge detection algorithm selection process.....	46
Figure 3-5 Steps 6 and 7 of the edge detection algorithm selection process .....	47
Figure 3-6 The total number of edge pixels resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases .....	49
Figure 3-7 Metric $R_{calc}$ values resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases .....	49
Figure 3-8 Metric $R_{edge}$ values resulting from Canny, Prewitt, Roberts, and Sobel techniques for the three databases .....	50
Figure 3-9 Flowchart of the edge detection method selection procedure .....	53
Figure 3-10 Histograms of the normalized gradient-weighted edge pixels resulting from Prewitt method for four different samples .....	54
Figure 3-11 The total histograms (linear scale) of the normalized gradient-weighted edges pixels resulted from Prewitt edge detection method .....	55
Figure 3-12 The top view of the logarithmic scale plots for the total histograms of the normalized gradient-weighted edges pixels resulted from Prewitt edge detection method	56
Figure 3-13 The ratio $A_r$ for the database CVC-Clinic.....	62
Figure 3-14 The total number of circles $N_{total}$ found in the sample images of database CVC- Clinic .....	63



---

Figure 3-15 The ratio $A_r$ for the database CVC-Colon.....	63
Figure 3-16 The total number of circles $N_{total}$ found in the sample images of database CVC-Colon .....	64
Figure 3-17 The ratio $A_r$ for the database ETIS-Larib.....	65
Figure 3-18 The total number of circles $N_{total}$ found in the sample images of database ETIS-Larib.....	65
Figure 3-19 The number of circles found in the artificial images containing ellipses of large axes with size 30, 60, and 90 pixels .....	69
Figure 3-20 The ratio of the classical and fuzzy Hough transformations' total runtime for the wide and thin gradient thresholds compared to the original, full Hough transform times ..	71
Figure 4-1 The reference case segmentation results of the grayscale version of image (28) in database CVC-ColonDB.....	85
Figure 4-2 The pre-processing steps of image (127) in database CVC-ClinicDB.....	86
Figure 4-3 The masks resulting from the geodesic method for one input image with different prefiltering methods.....	89
Figure 4-4 The masks resulting from the Chan-Vese method for one input image with different prefiltering methods.....	90
Figure 4-5 Sørensen–Dice similarity coefficient comparison results for geodesic and Chan-Vese methods using both rectangular and circular initial masks for the image samples of the three studied databases .....	92
Figure 4-6 Sørensen–Dice coefficients for geodesic and Chan-Vese methods using both rectangular and circular initial masks of the 12 sample images with all the 5 selected preprocessing methods .....	96
Figure 4-7 The relative difference of the Sørensen–Dice coefficients for geodesic and Chan-Vese methods using both rectangular and circular initial masks of the 12 sample images with all the 5 selected preprocessing methods.....	96
Figure 4-8 Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image samples in database ETIS-Larib Polyp DB .....	98
Figure 4-9 Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image samples in database CVC-ColonDB.....	98
Figure 4-10 Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image samples in database CVC-ClinicDB.....	99

Figure 4-11 The Chan–Vese comparison segmentation results of the image (135) in database  
ETIS-Larib Polyp DB ..... 100

Figure 4-12 Darker polyp example: image 516 in database CVC-ClinicDB..... 102

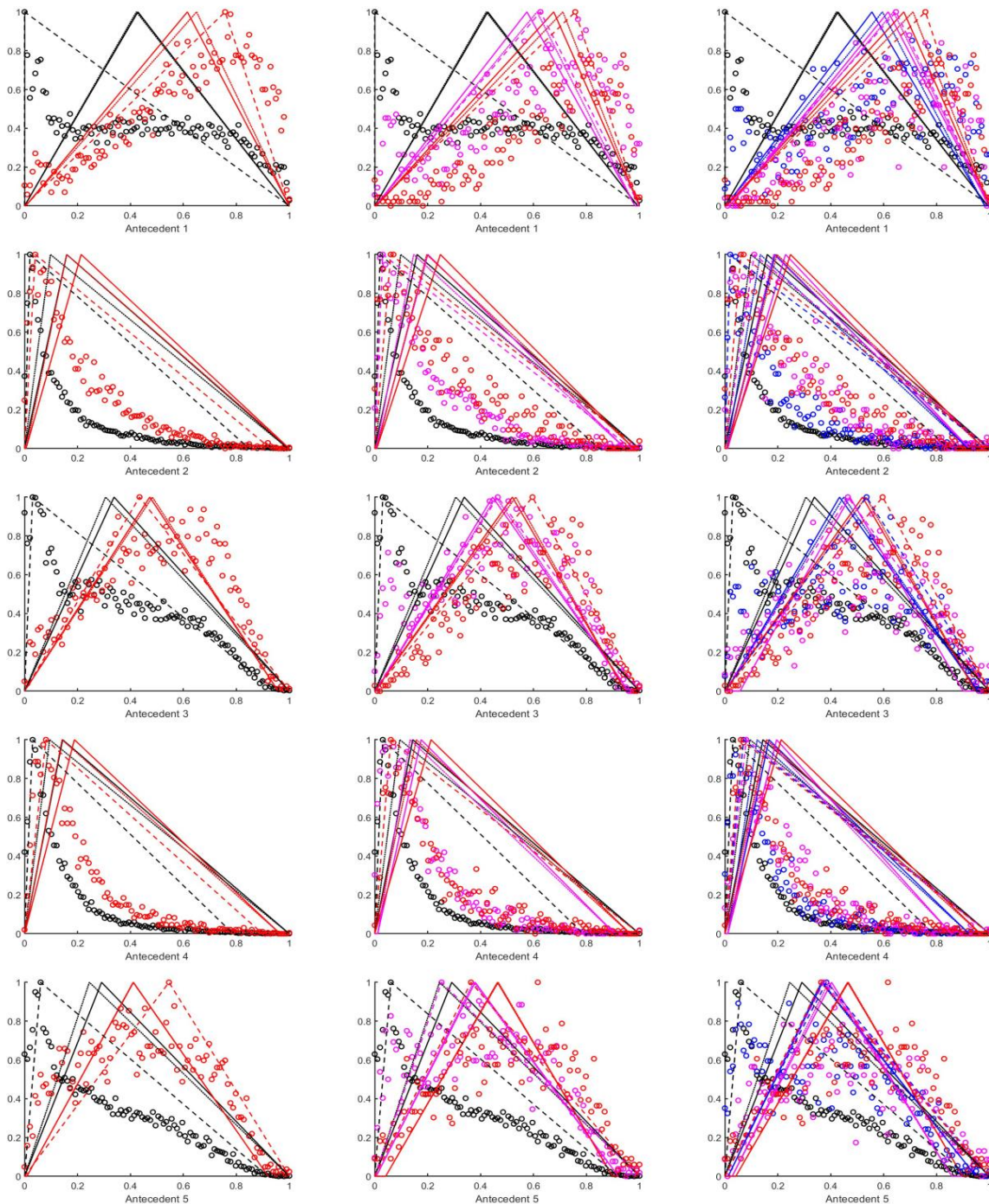
## List of Tables

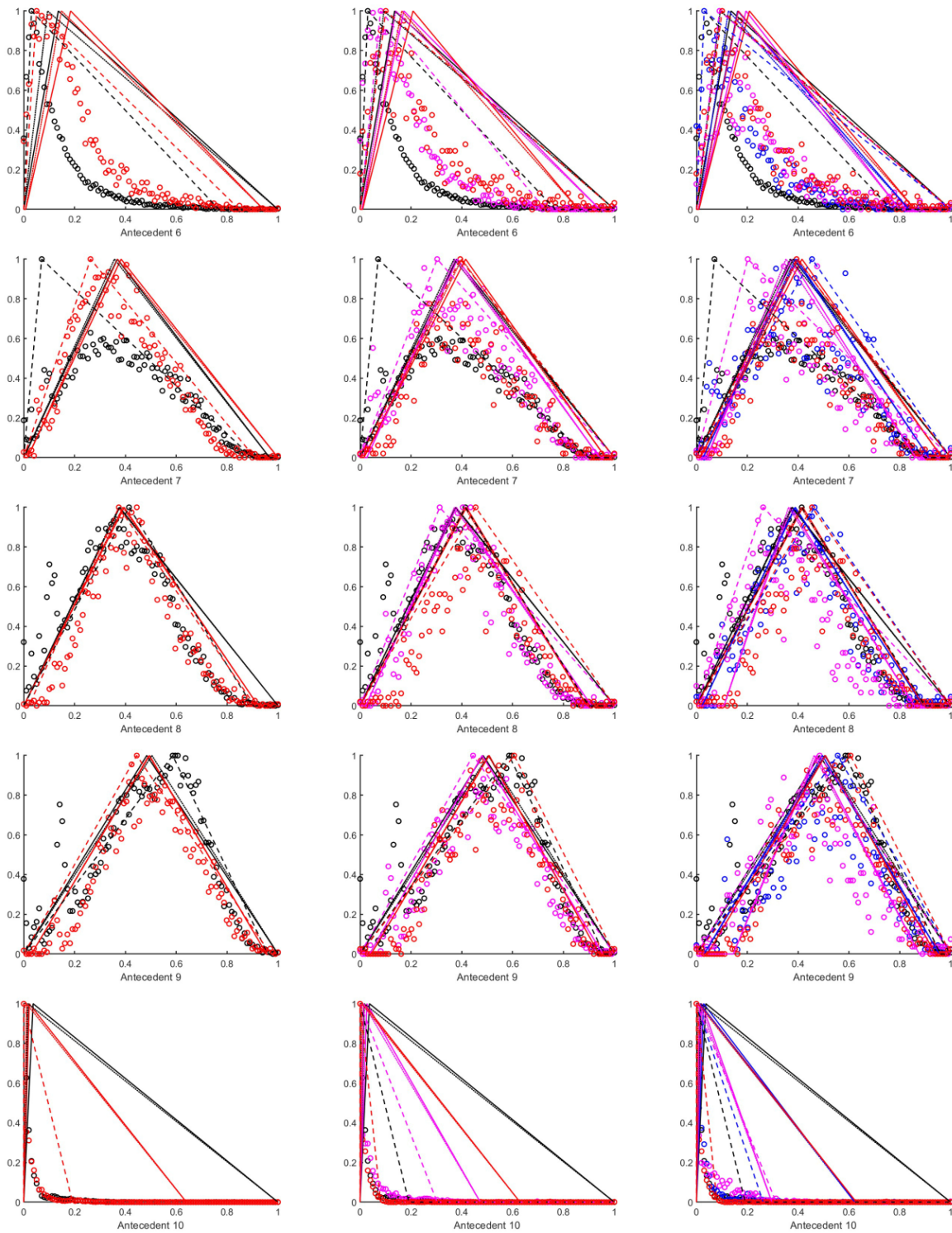
Table 2-1 The numbering of the antecedent parametersr.....	18
Table 2-2 Classification performance for the different methods with different number of selected antecedents.....	23
Table 2-3 Classification performance for different number of selected antecedents with different consequent classes using mean-centered rulebase generation method for RGB and HSV color spaces.....	26
Table 2-4 Classification performance for different number of selected antecedents with different consequent classes using median-centered rulebase generation method for RGB and HSV color spaces.....	27
Table 2-5 Classification performance for different number of selected antecedents with different consequent classes using histogram-fitted 1% rulebase generation method for RGB and HSV color spaces.....	28
Table 3-1 The practical strengths and weaknesses of Hough transform.....	32
Table 3-2 Results of selection the most appropriate edge detection technique using metric $R_{calc}$ .....	52
Table 3-3 Results of selection the most appropriate edge detection technique using metric $R_{edge}$ .....	52
Table 3-4 The selected samples and their metrics $R_{calc}$ and $R_{edge}$ together with the total number of Prewitt edge pixels.....	66
Table 3-5 Properties of the ellipses of the artificial test images of size $200 \times 200$ pixels.....	68
Table 4-1 Sørensen–Dice Coefficient for the image samples in the reference case.....	84
Table 4-2 Sørensen–Dice Coefficient of the post pre-processing case for the image samples in database ETIS-Larib Polyp DB (Silva, Histace, Romain, Dray, & Granado, 2014) using both rectangular and circular initial masks.....	93
Table 4-3 Sørensen–Dice Coefficient of the post pre-processing case for the image samples in database CVC-ColonDB (Bernal, Sanchez, & Vilariño, 2012) using both rectangular and circular initial masks.....	94
Table 4-4 Sørensen–Dice Coefficient of the post pre-processing case for the image samples in database CVC-ClinicDB (Bernal J. , et al., 2015) using both rectangular and circular initial masks.....	95

Table 4-5 Numerical values for Sørensen–Dice Coefficient of the modified rectangular initial mask and tuned parameters case for the image (151) in database CVC-ClinicDB (Bernal J. , et al., 2015), with 100, 200, and 300 iterations .....	99
--	----

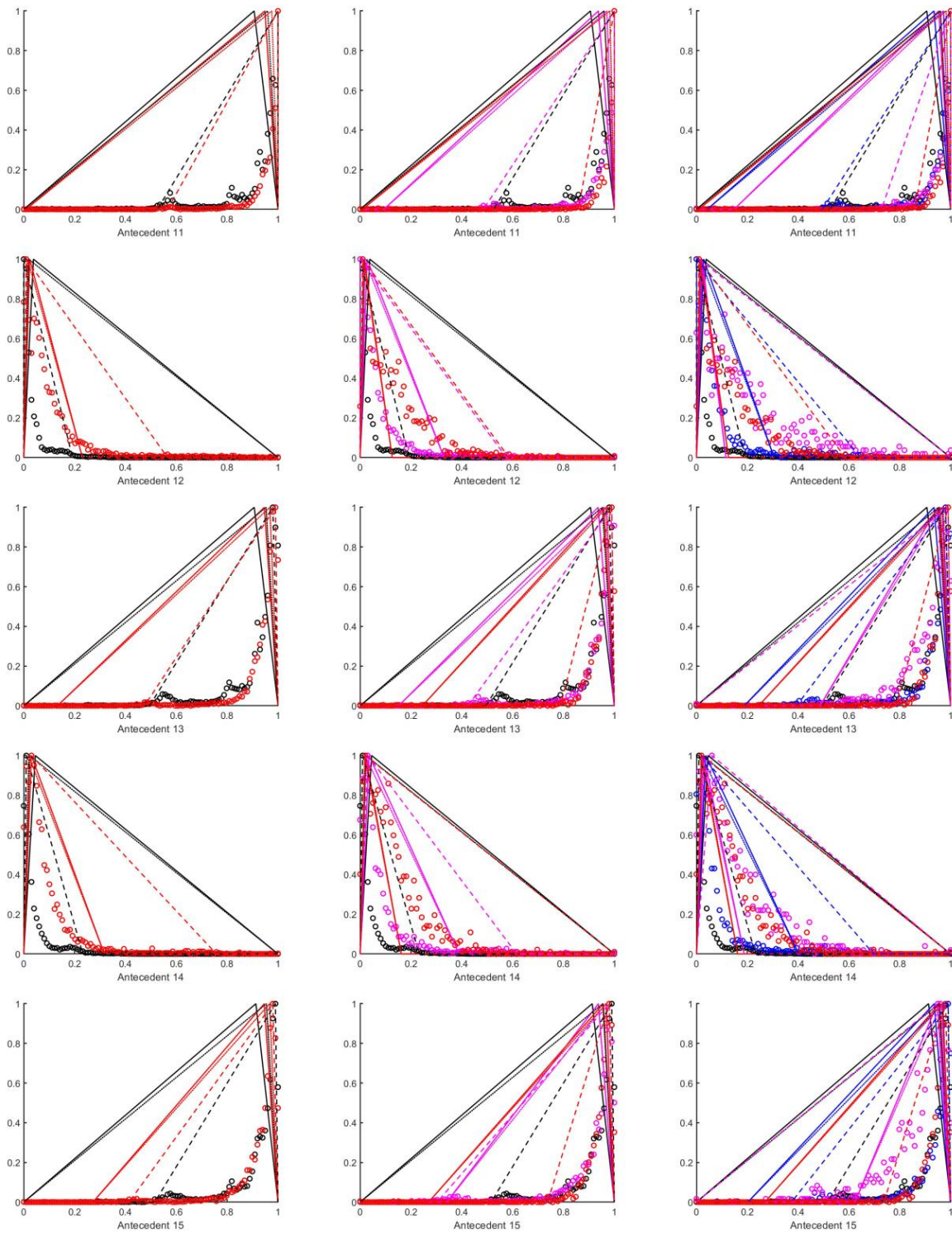
## Appendix A

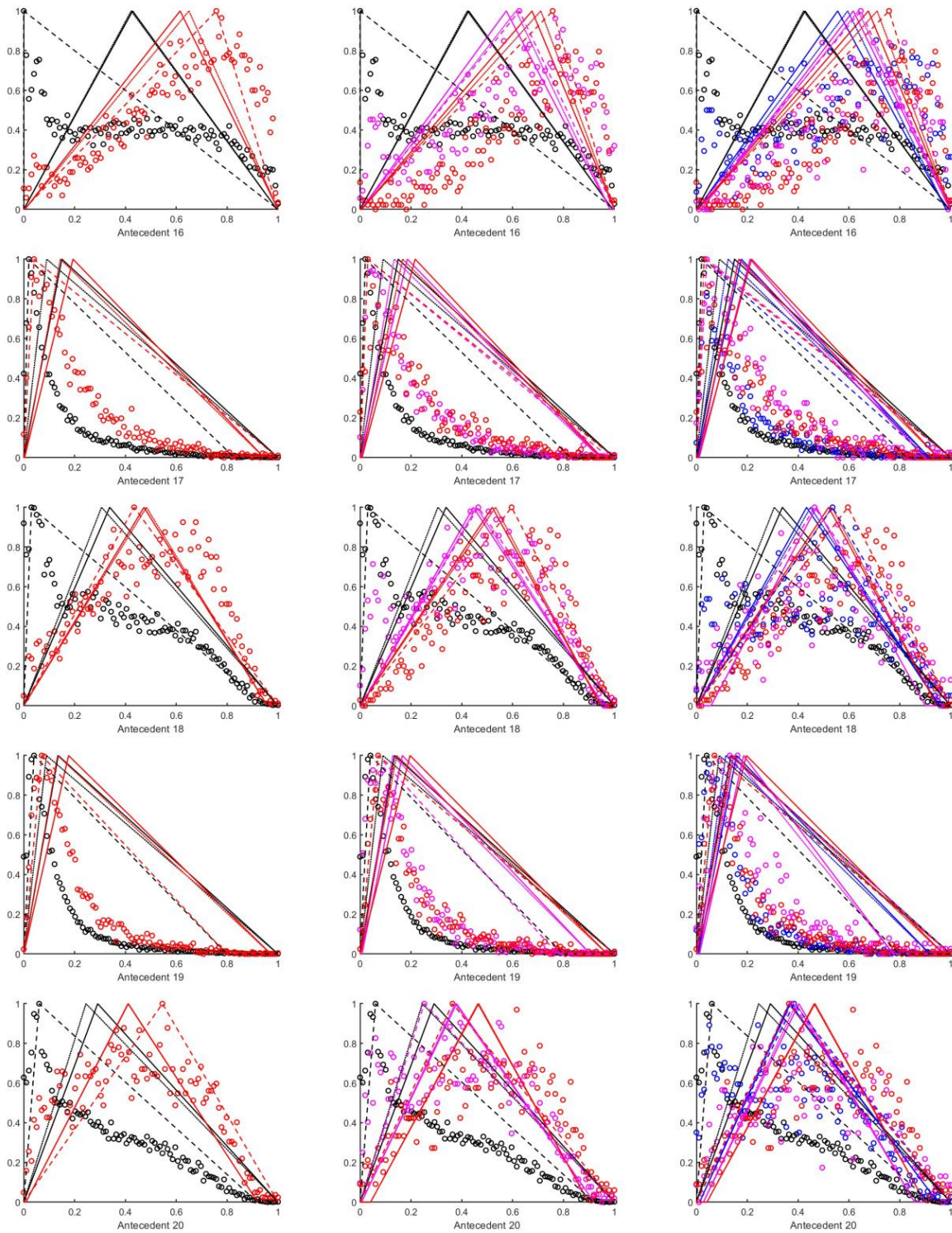
The histograms of all 99 antecedents using RGB color space, together with the mean-, median-centered, and histogram-fitted 1% triangular rules for the 2, 3, and 4 consequent classes. The consequent classes are plotted with the following colors. Red: yes polyp (with the highest threshold in the area percentage). Magenta: yes polyp, with 50% or lower percentage. Blue: Yes polyp with 20% or lower percentage in the area of the tile. Black: no polyp. The line types mean the following: No line, circle marker: the histogram. Continuous line: mean-centered rules. Dotted line: median-centered rules. Dashed lines: histogram-fitted rule.



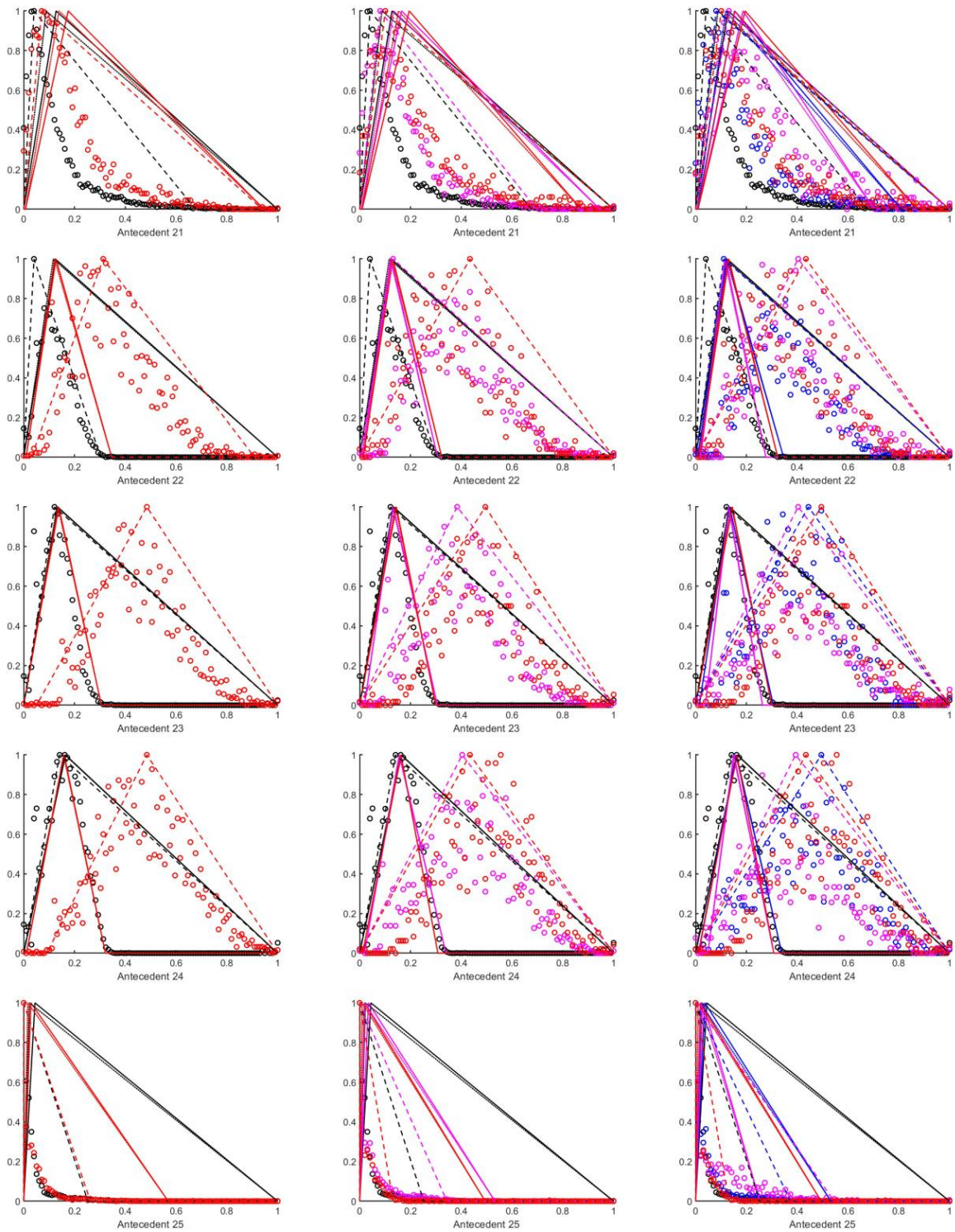


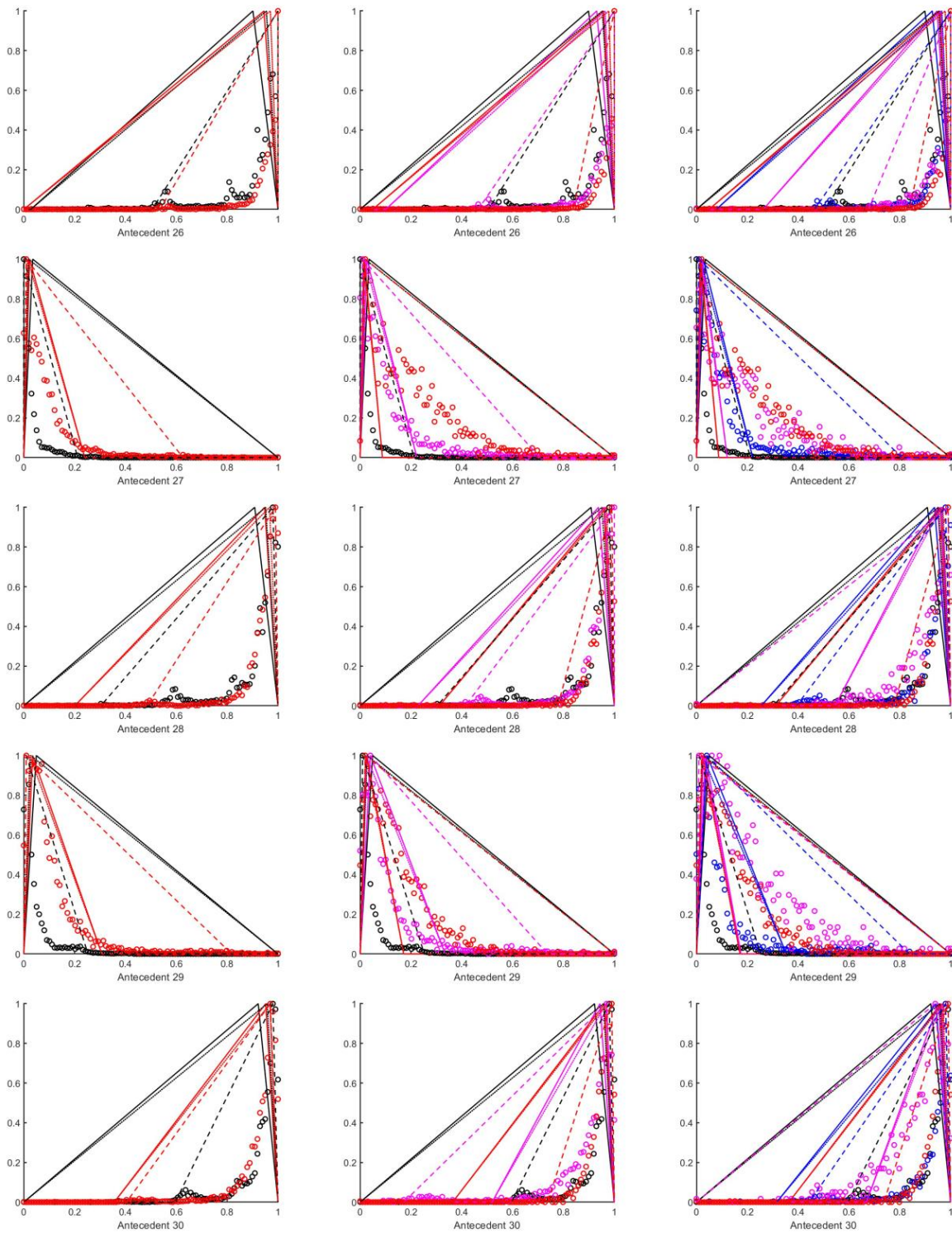




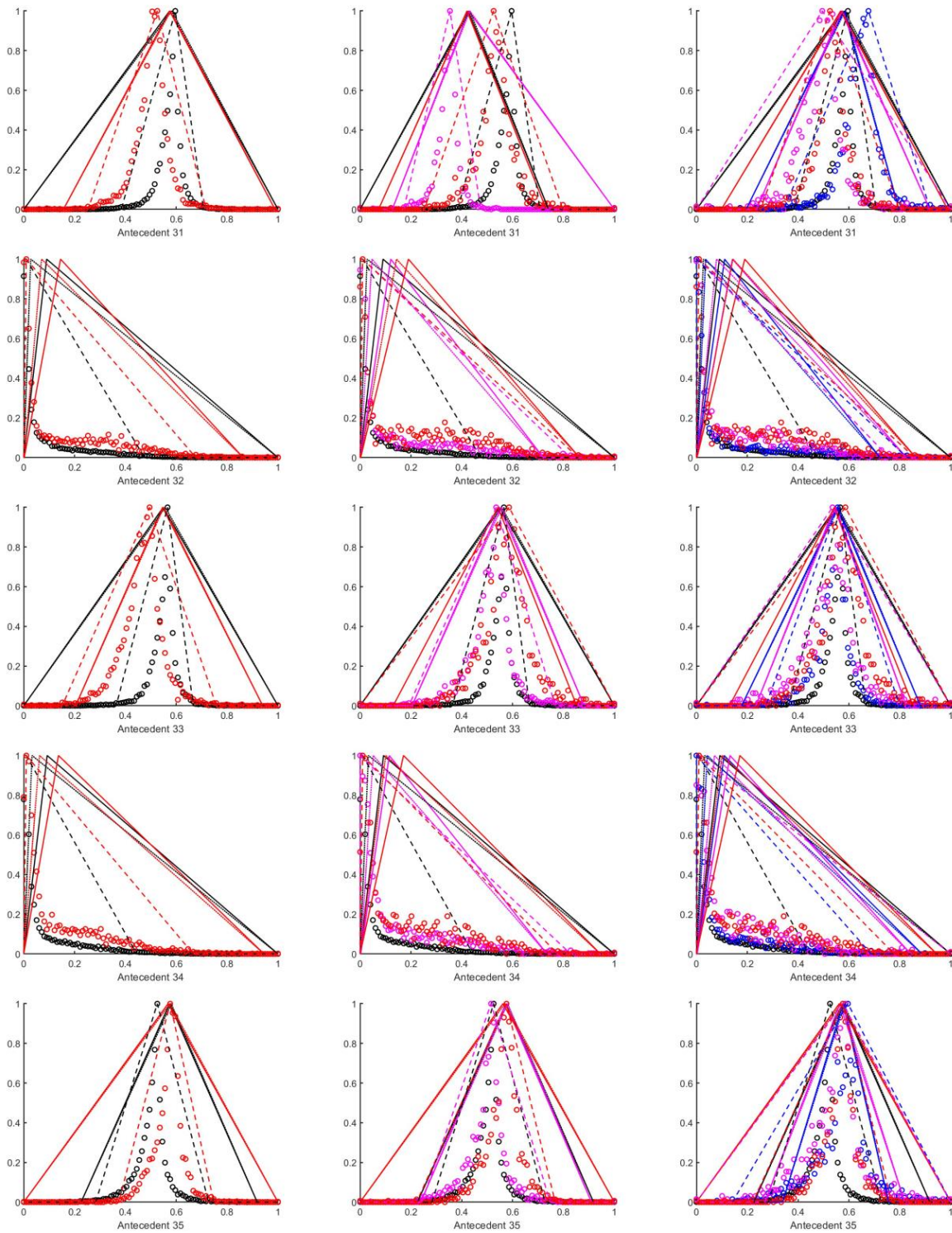


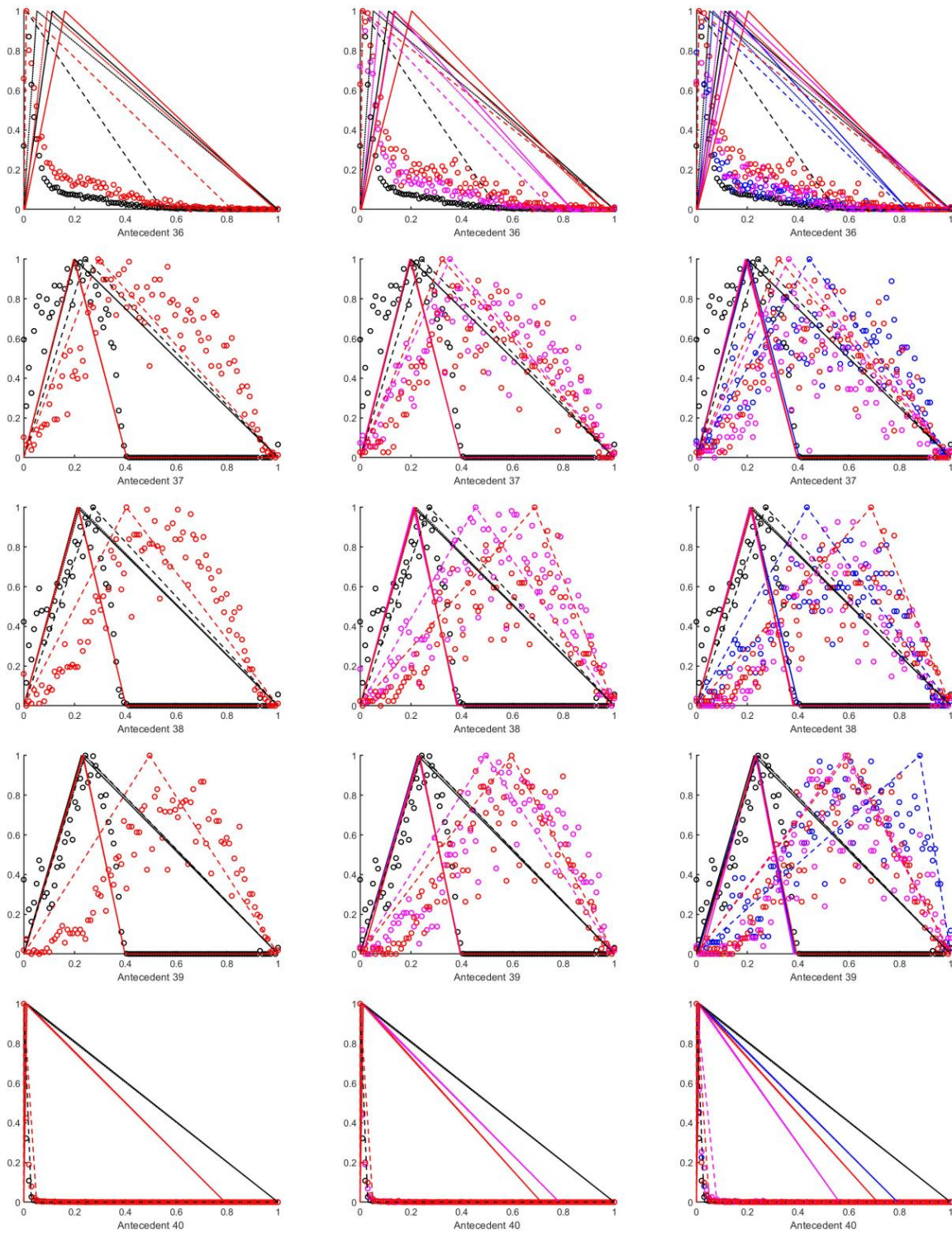




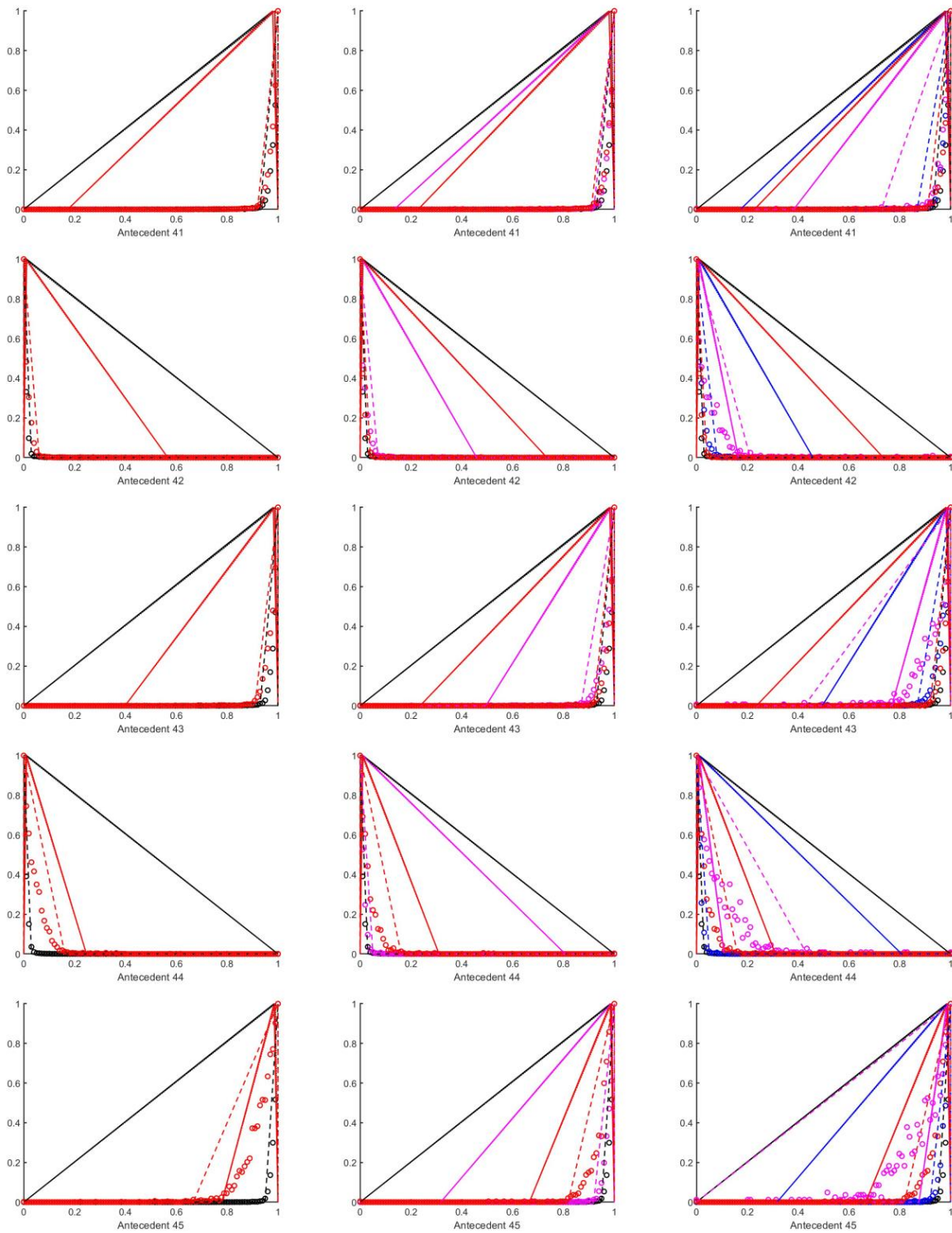


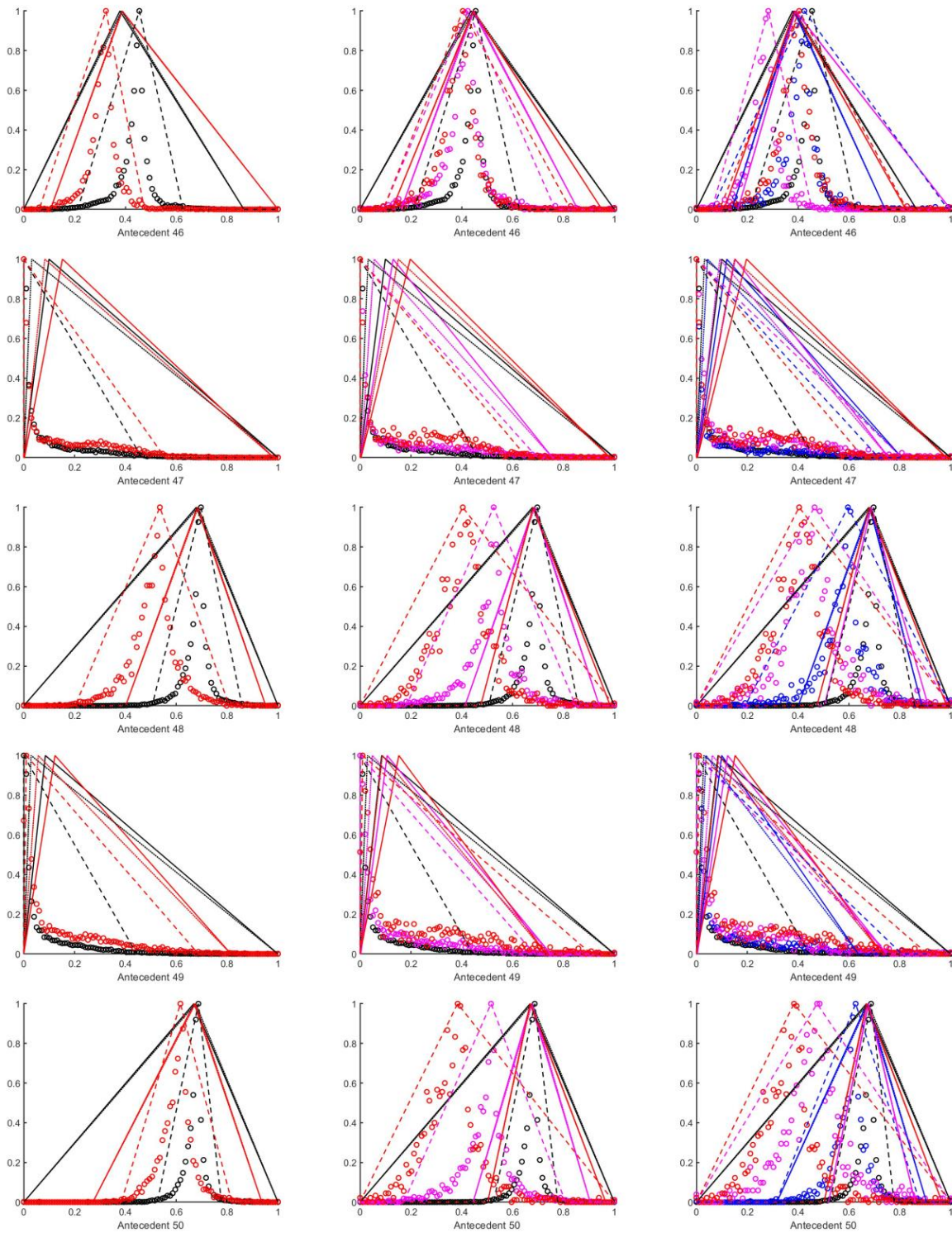




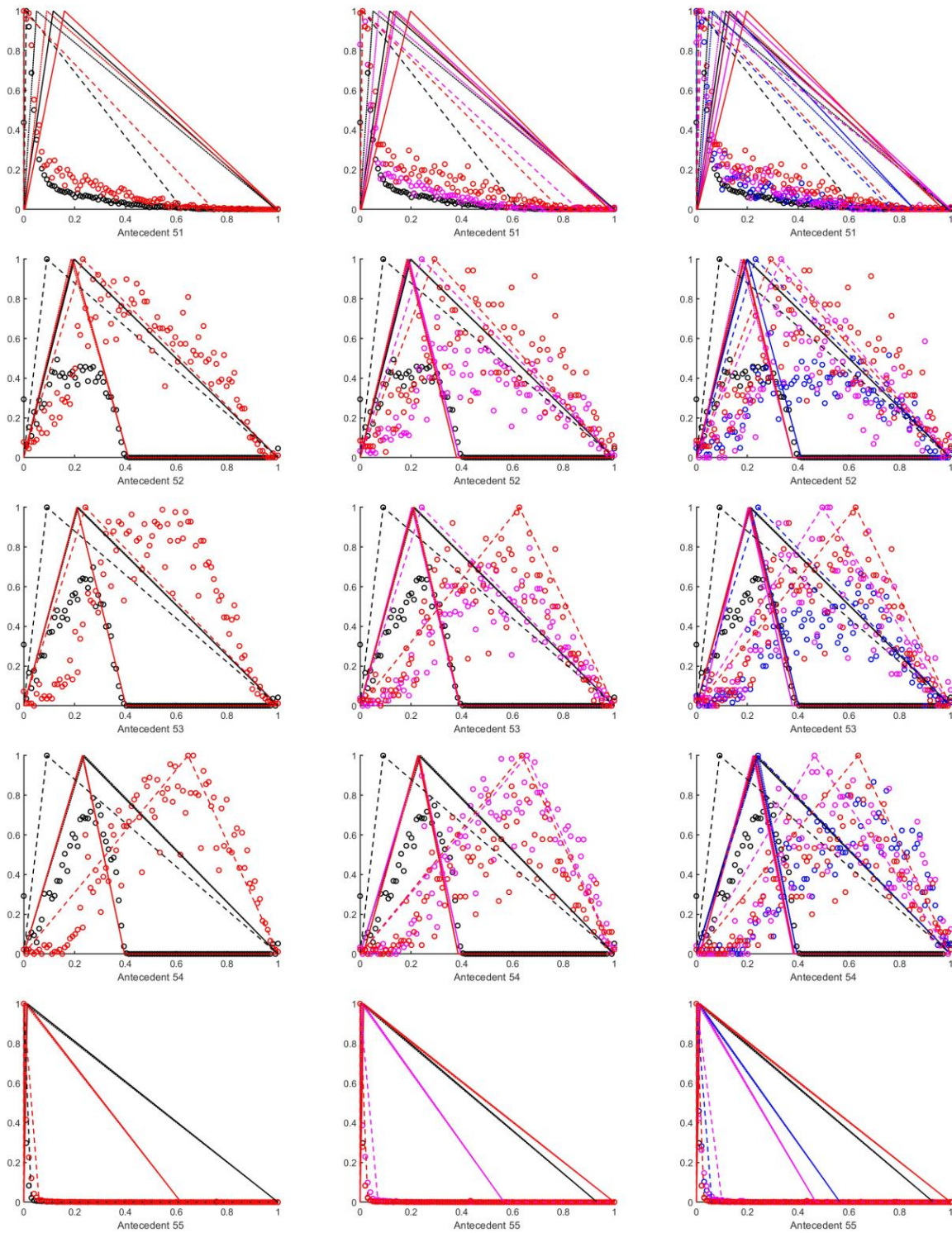


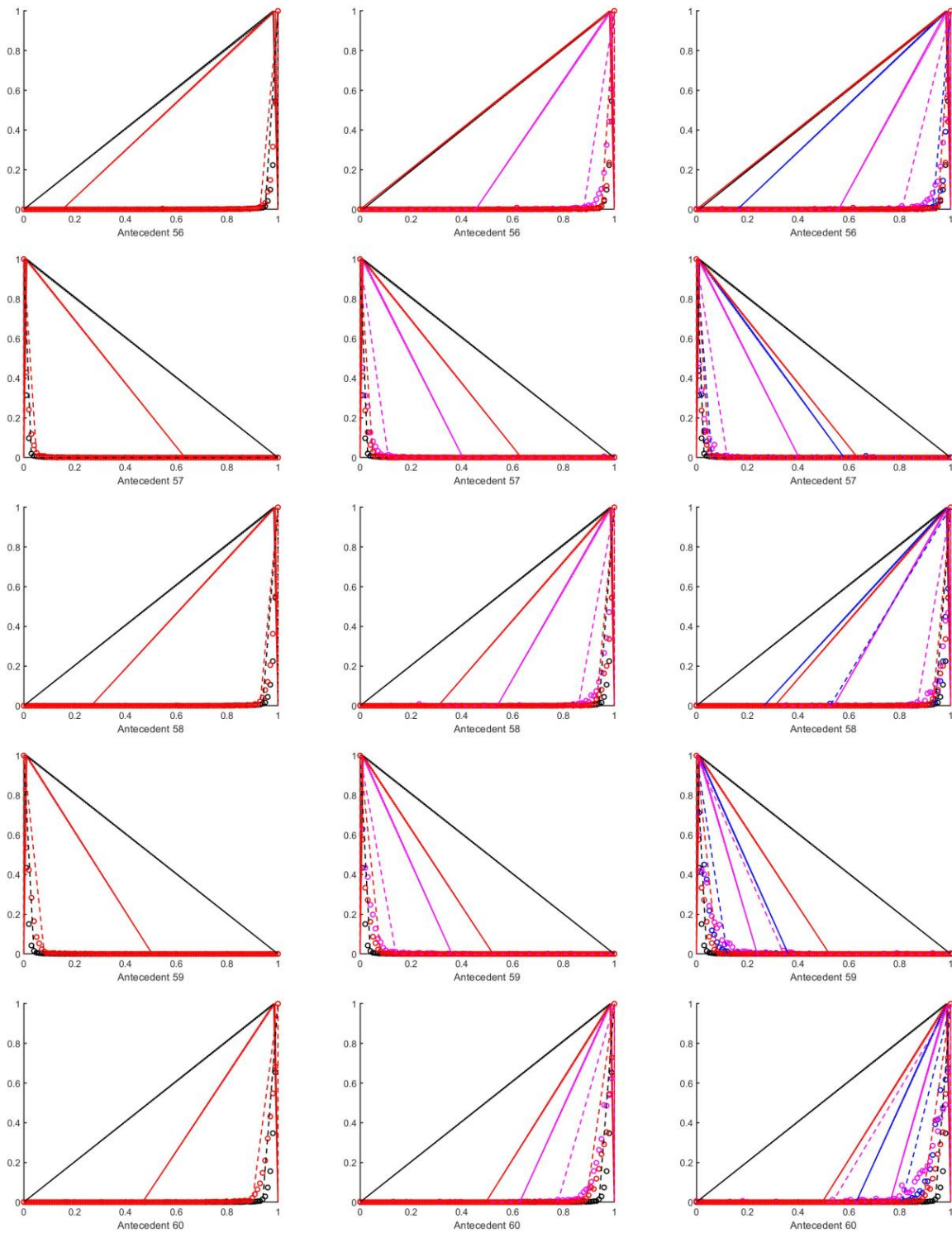




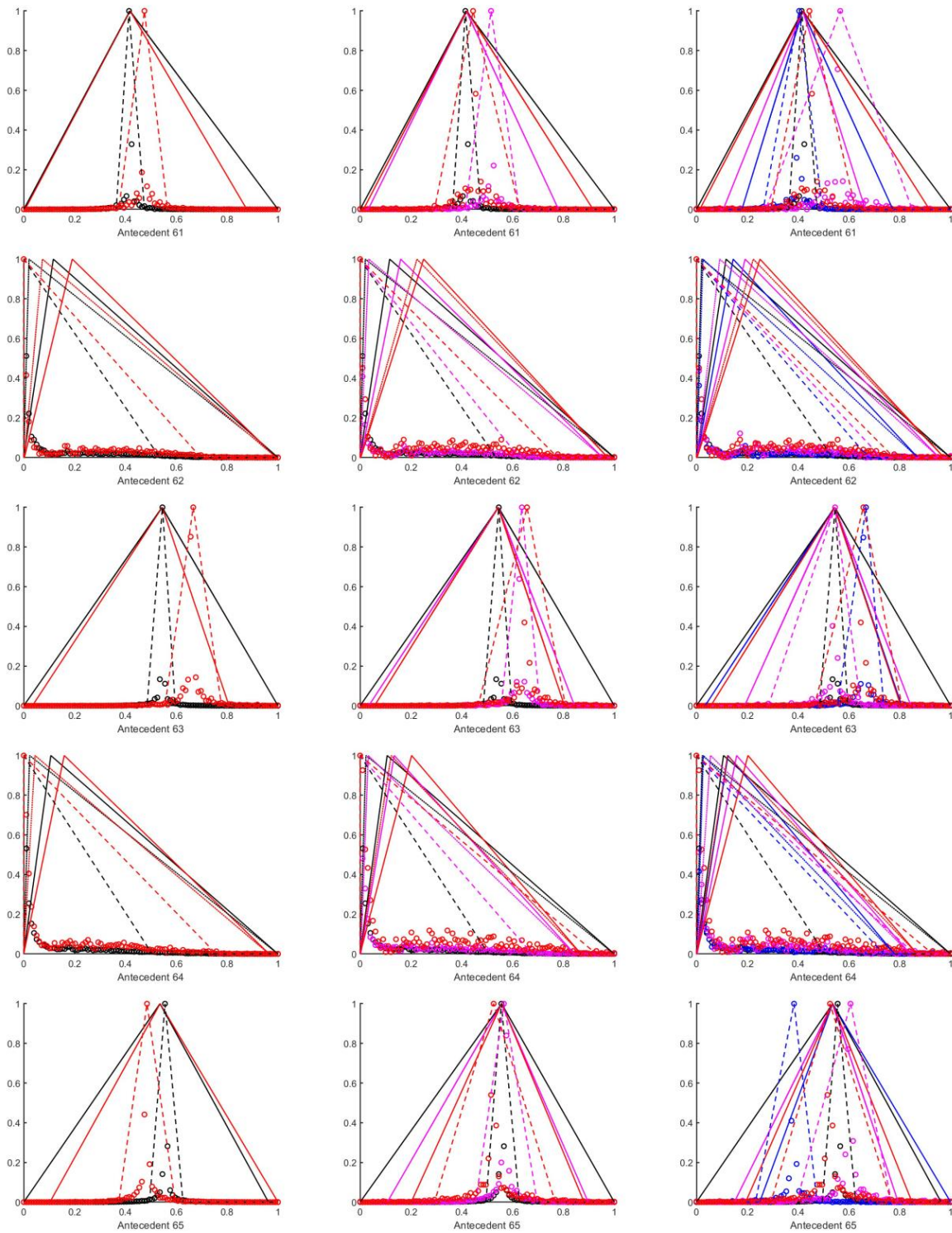


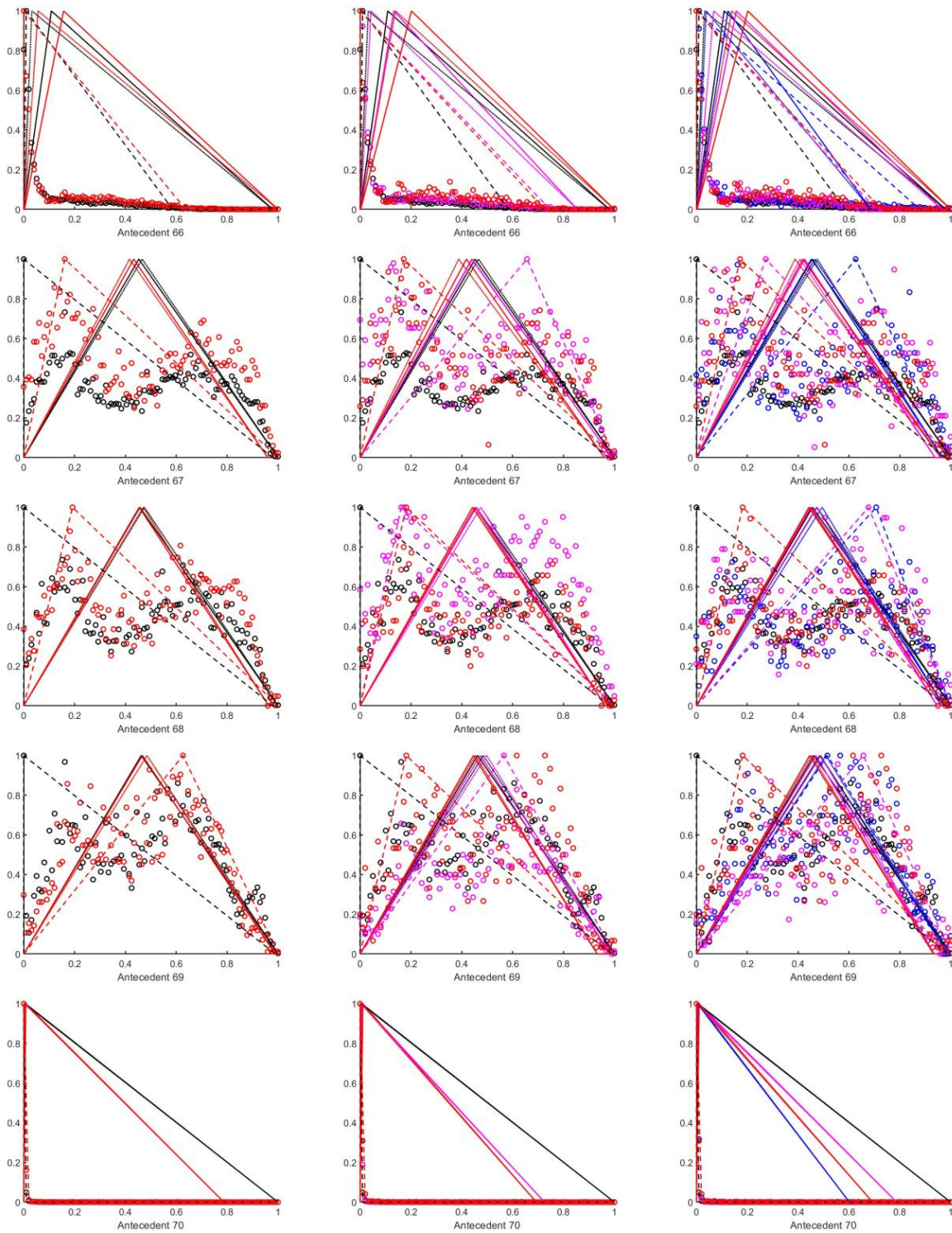


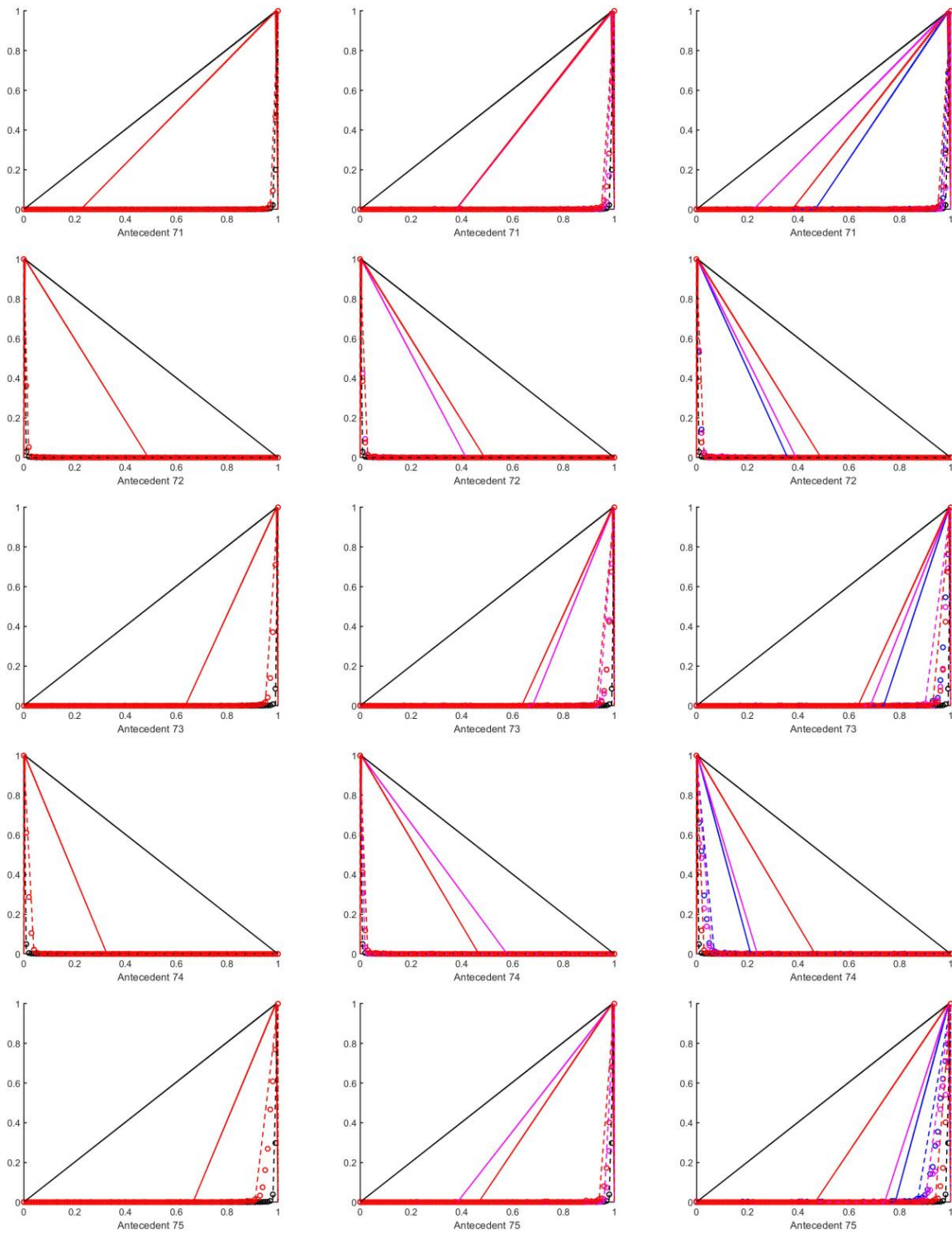




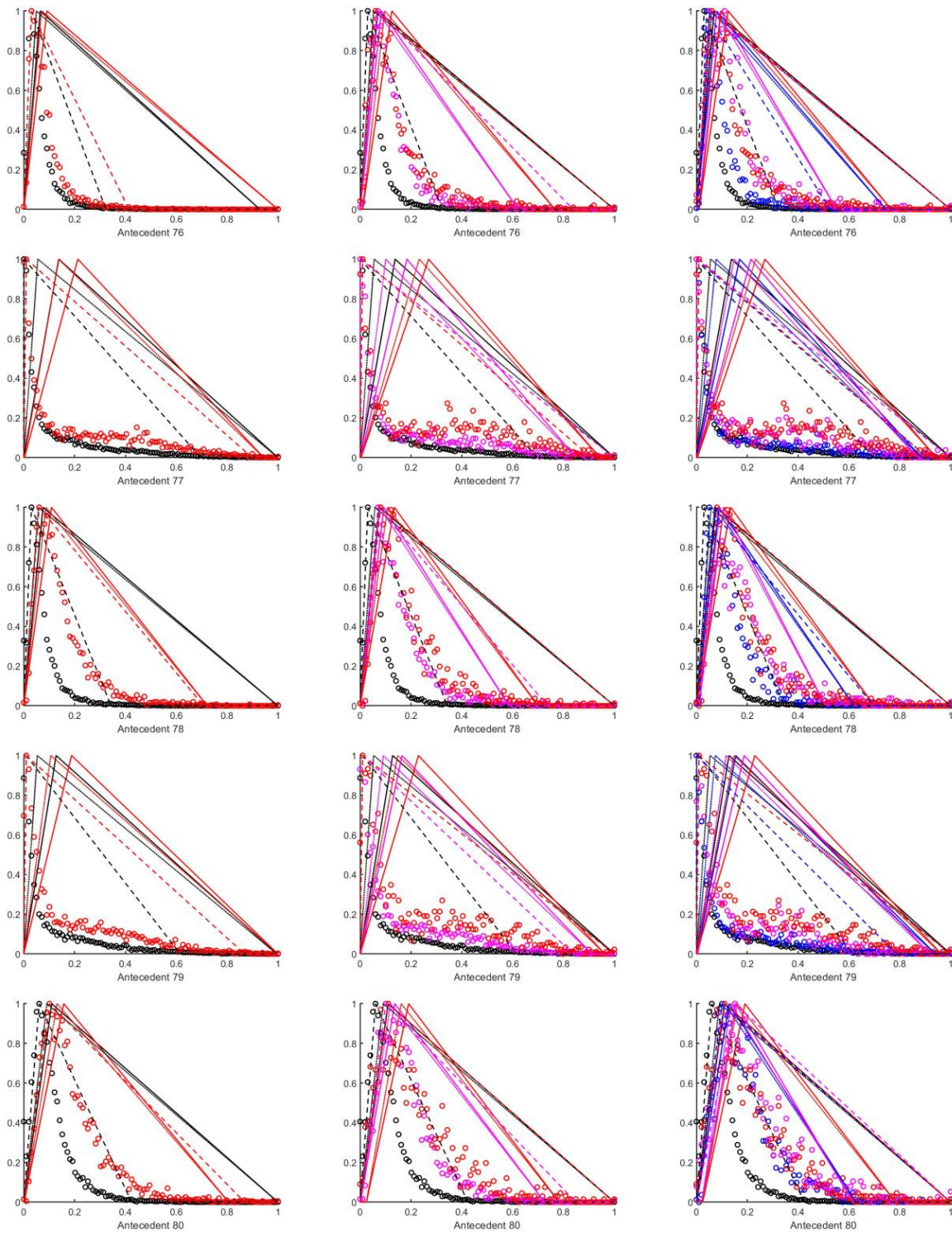


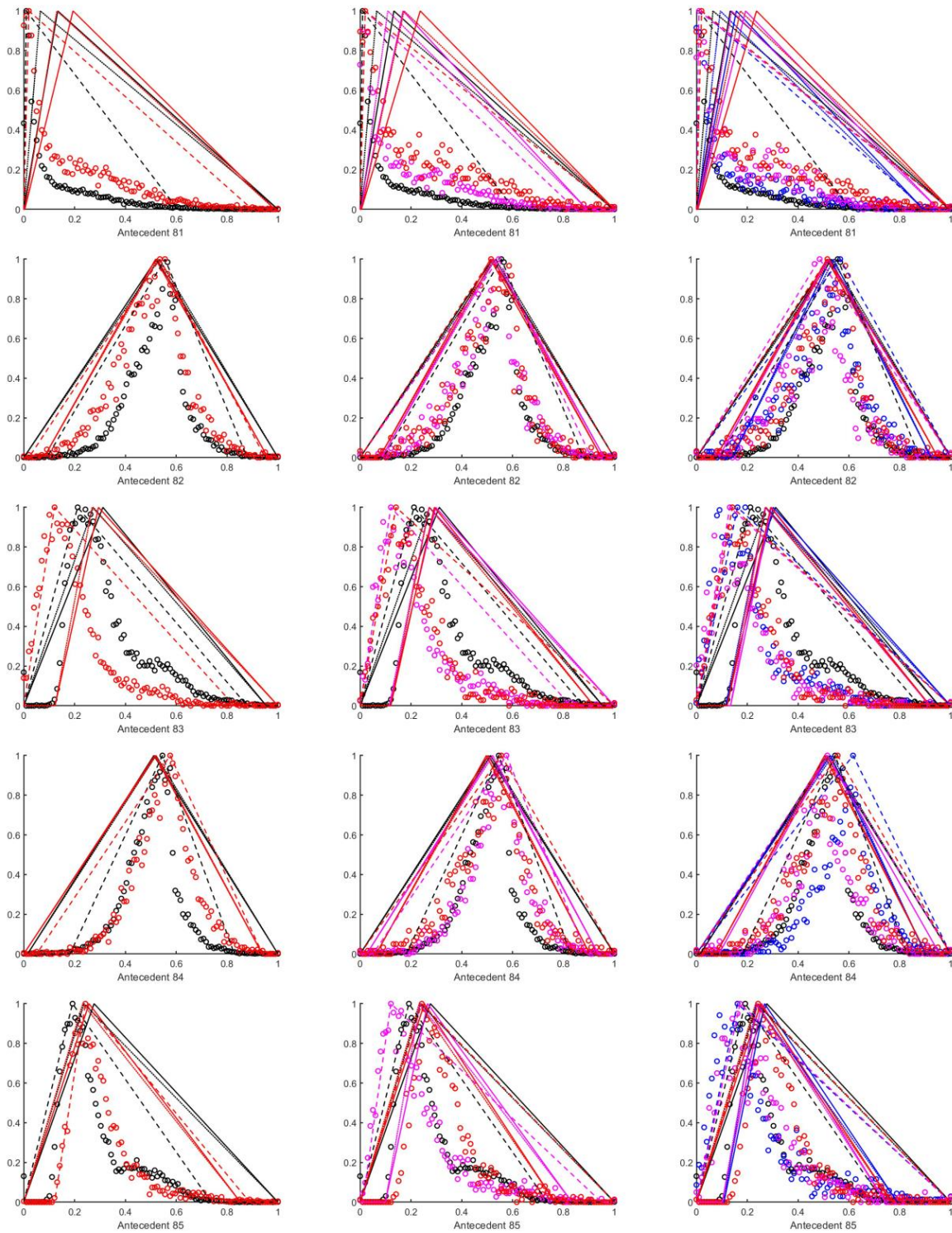




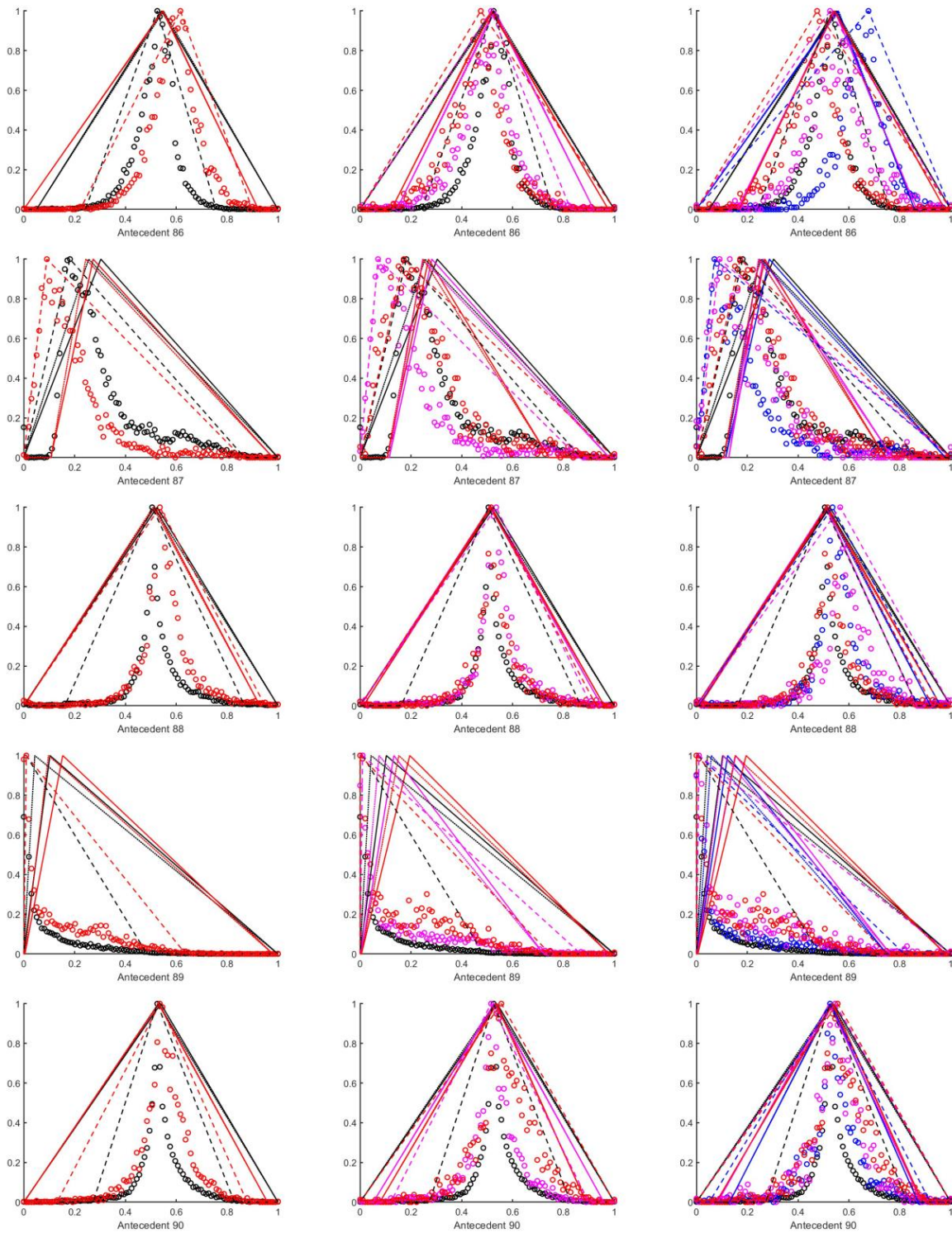


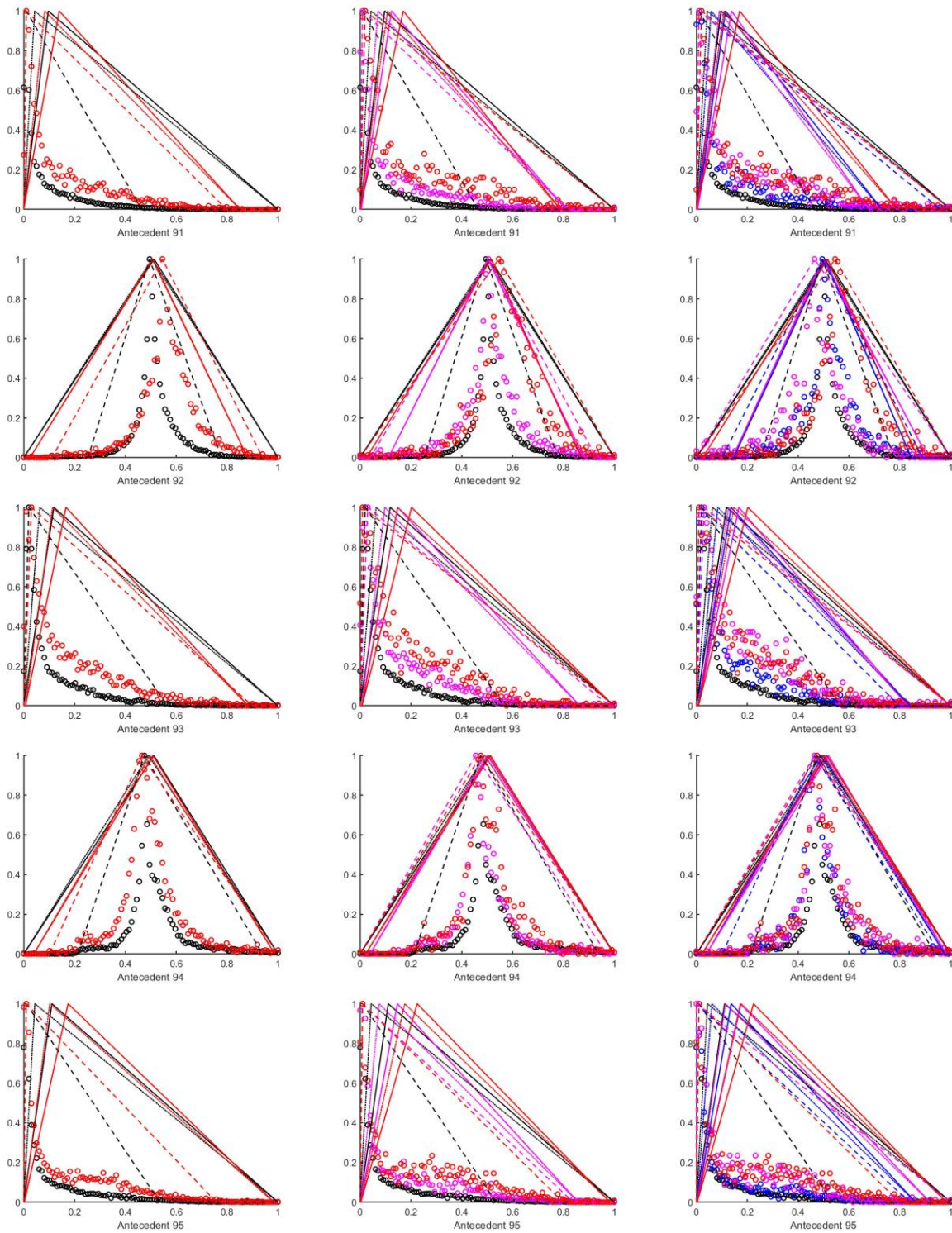




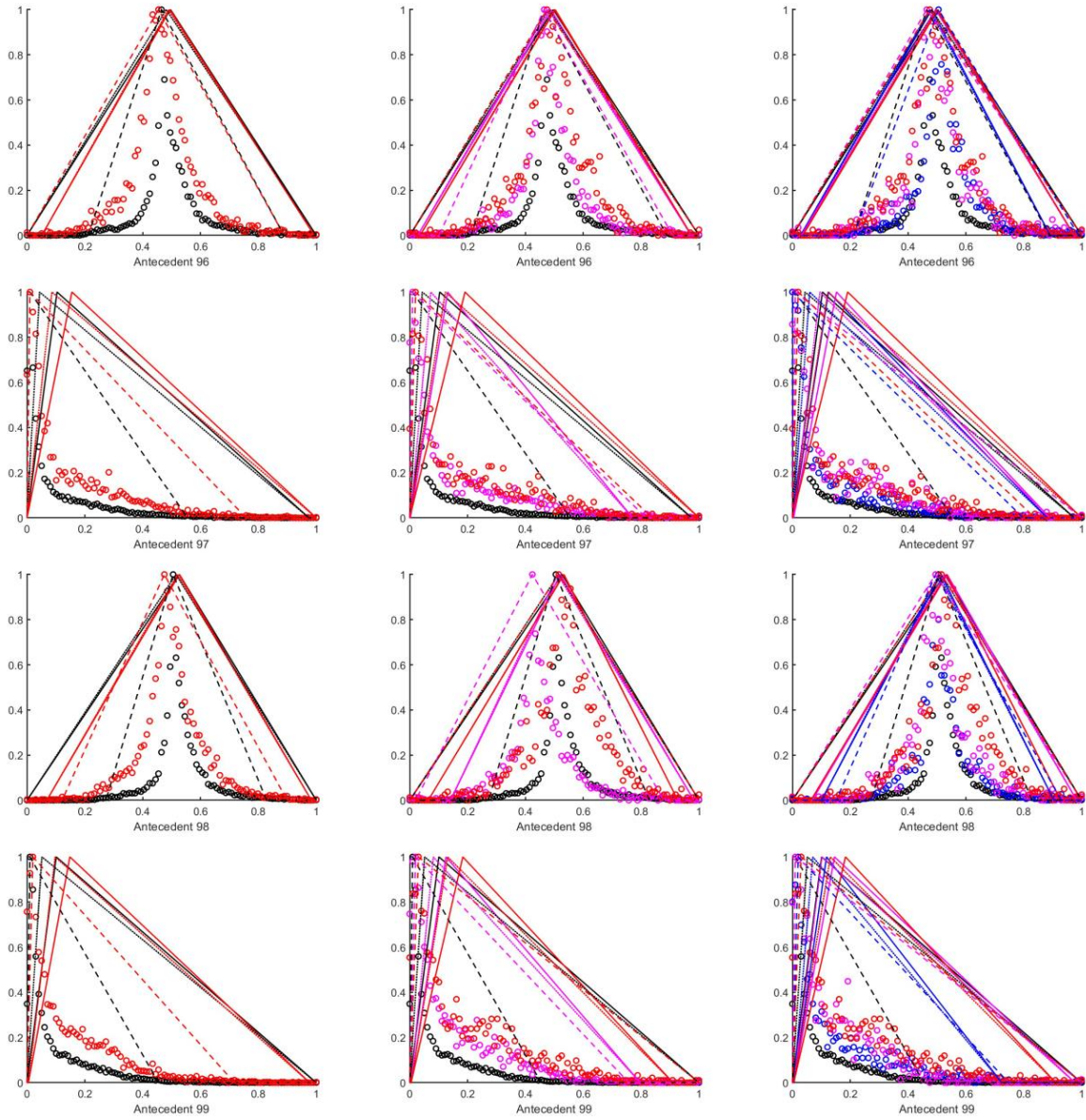








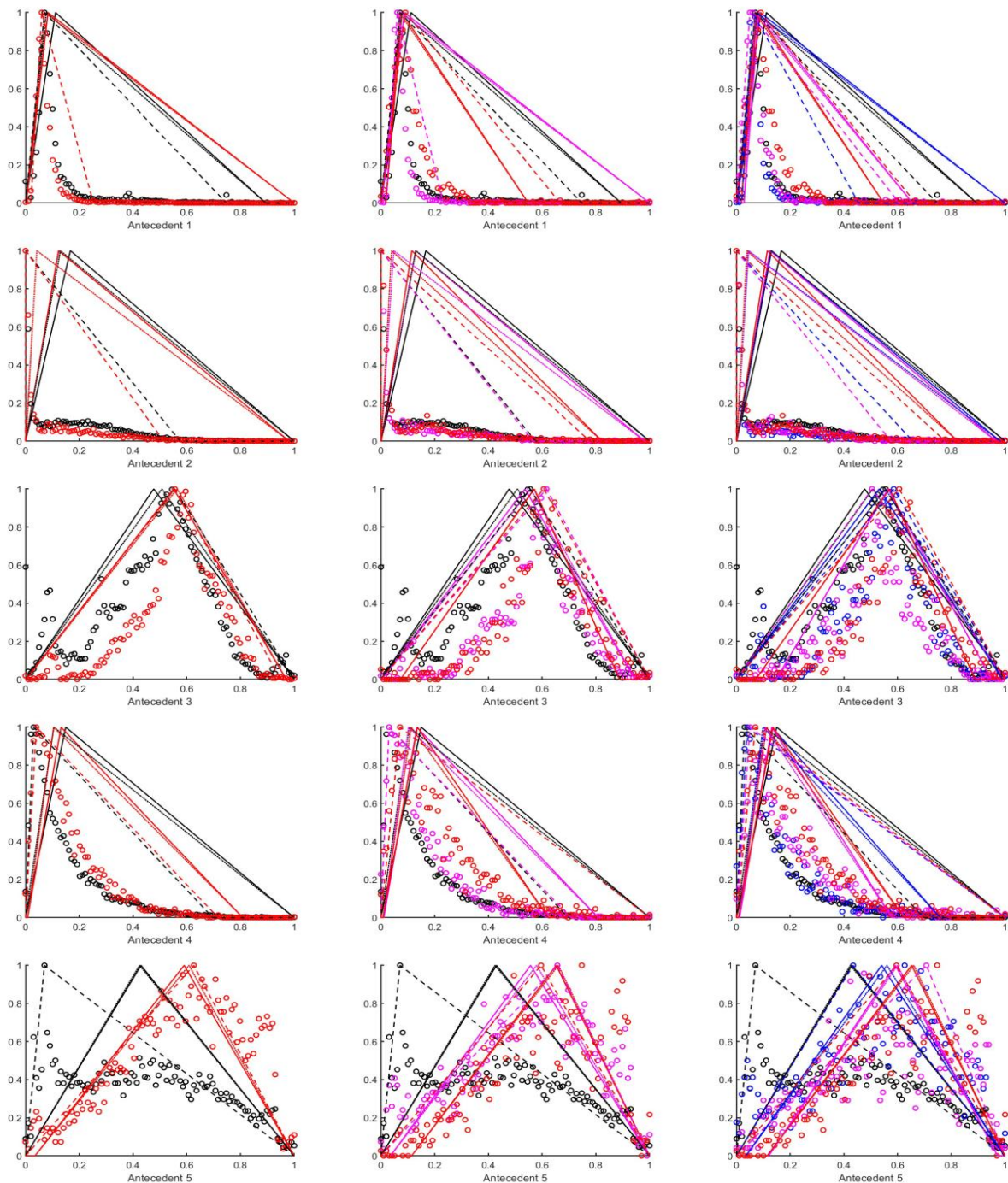


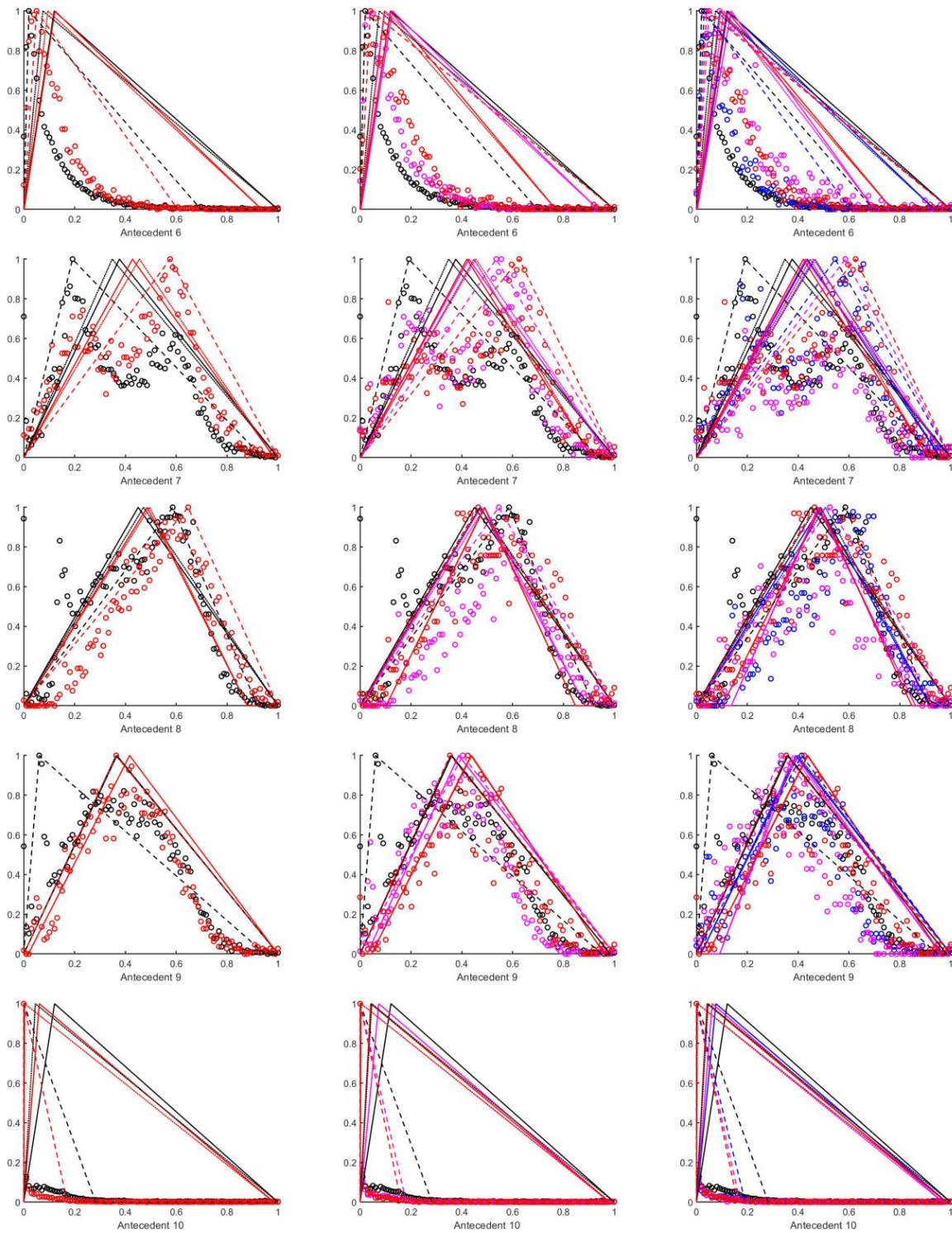




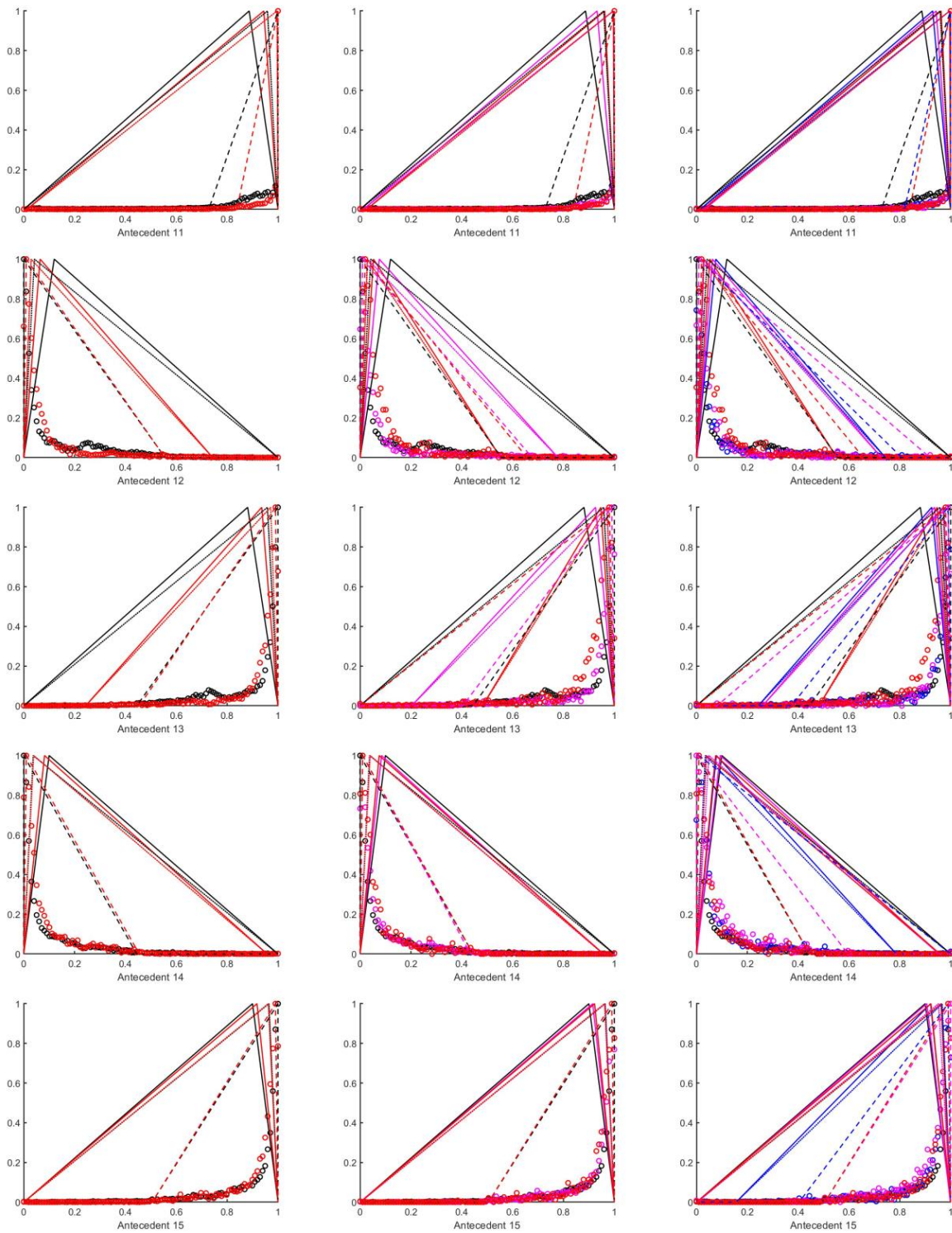
## Appendix B

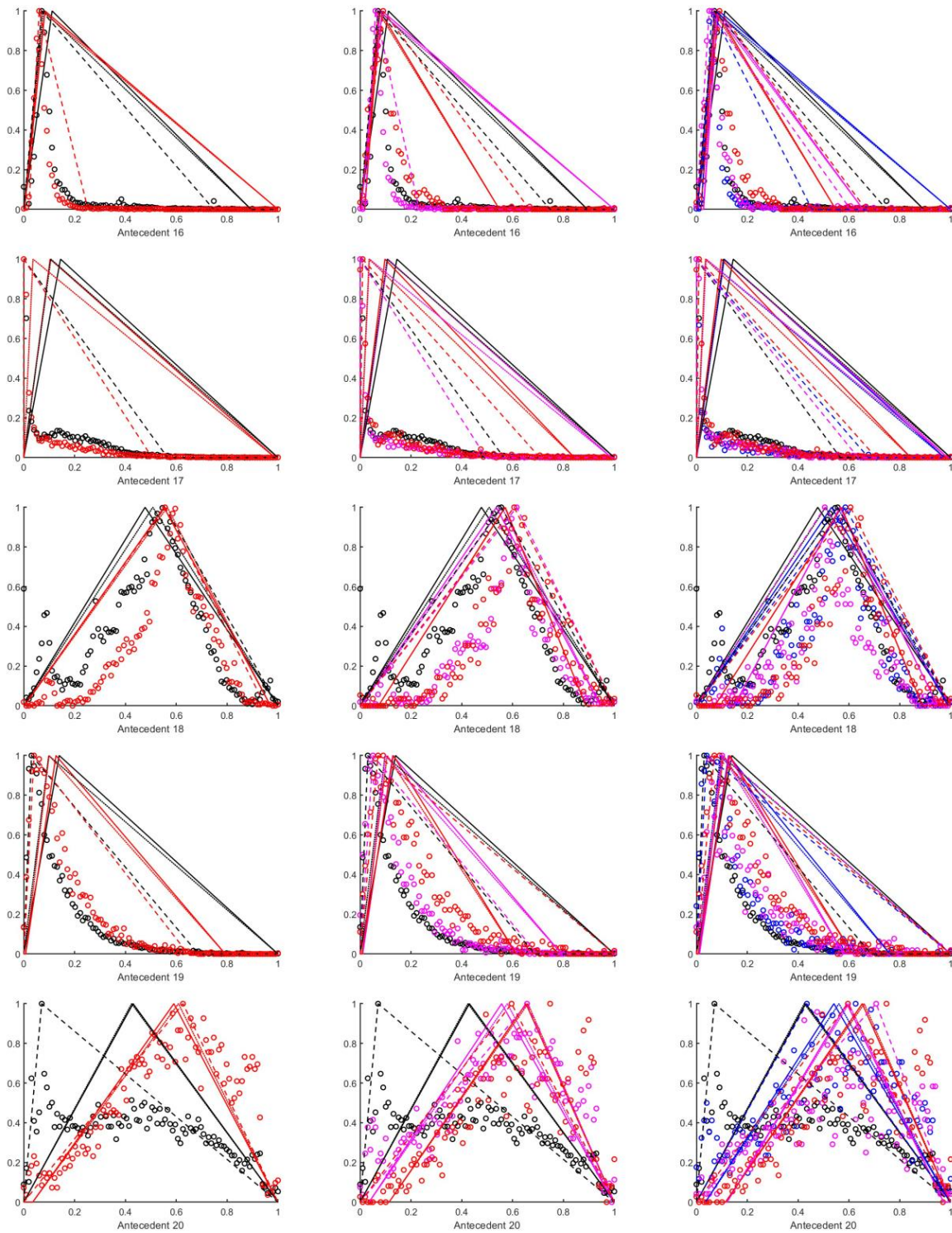
The histograms of all 99 antecedents using HSV color space, together with the mean-, median-centered, and histogram-fitted 1% triangular rules for the 2, 3, and 4 consequent classes. The consequent classes are plotted with the following colors. Red: yes polyp (with the highest threshold in the area percentage). Magenta: yes polyp, with 50% or lower percentage. Blue: Yes polyp with 20% or lower percentage in the area of the tile. Black: no polyp. The line types mean the following: No line, circle marker: the histogram. Continuous line: mean-centered rules. Dotted line: median-centered rules. Dashed lines: histogram-fitted rule.



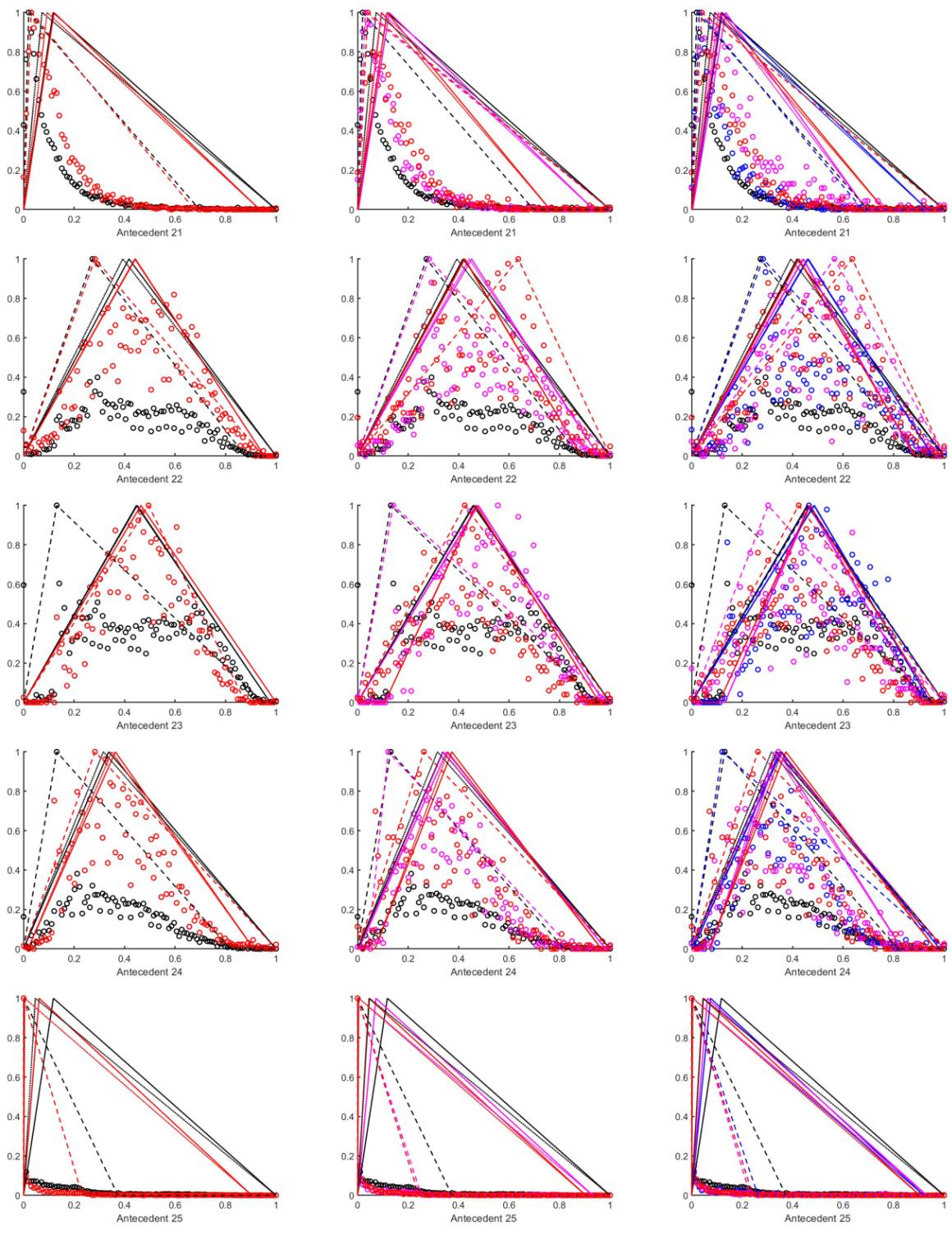


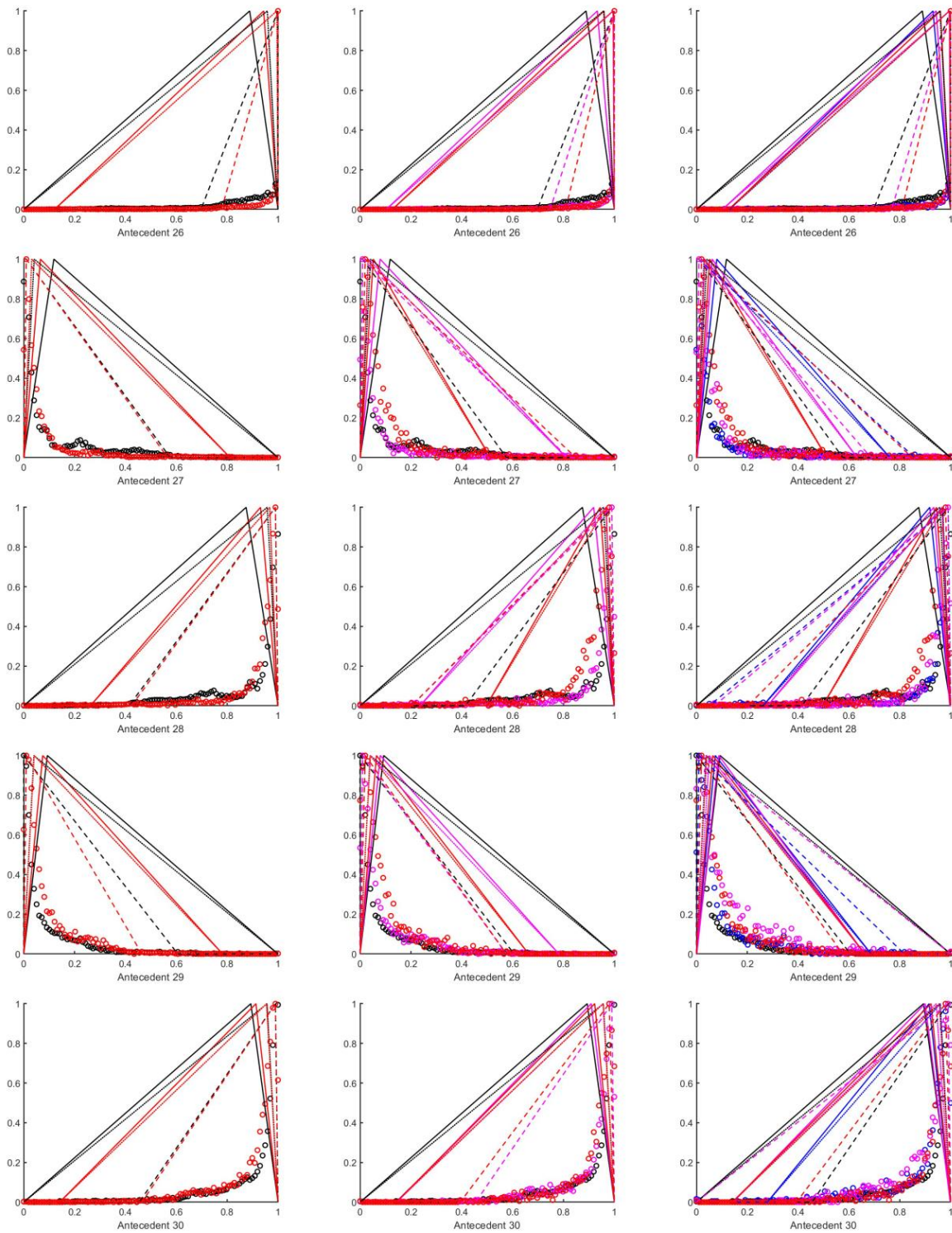




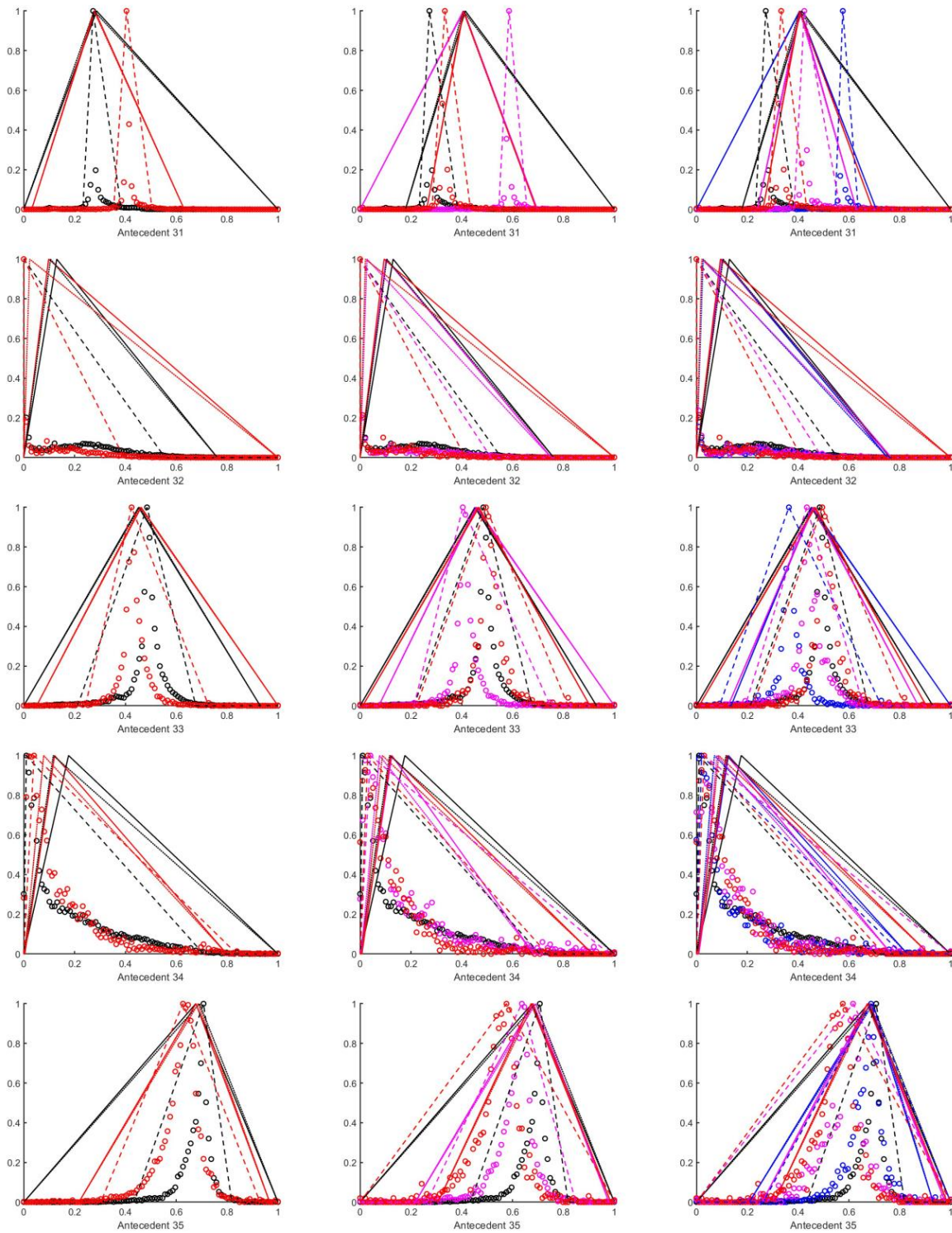




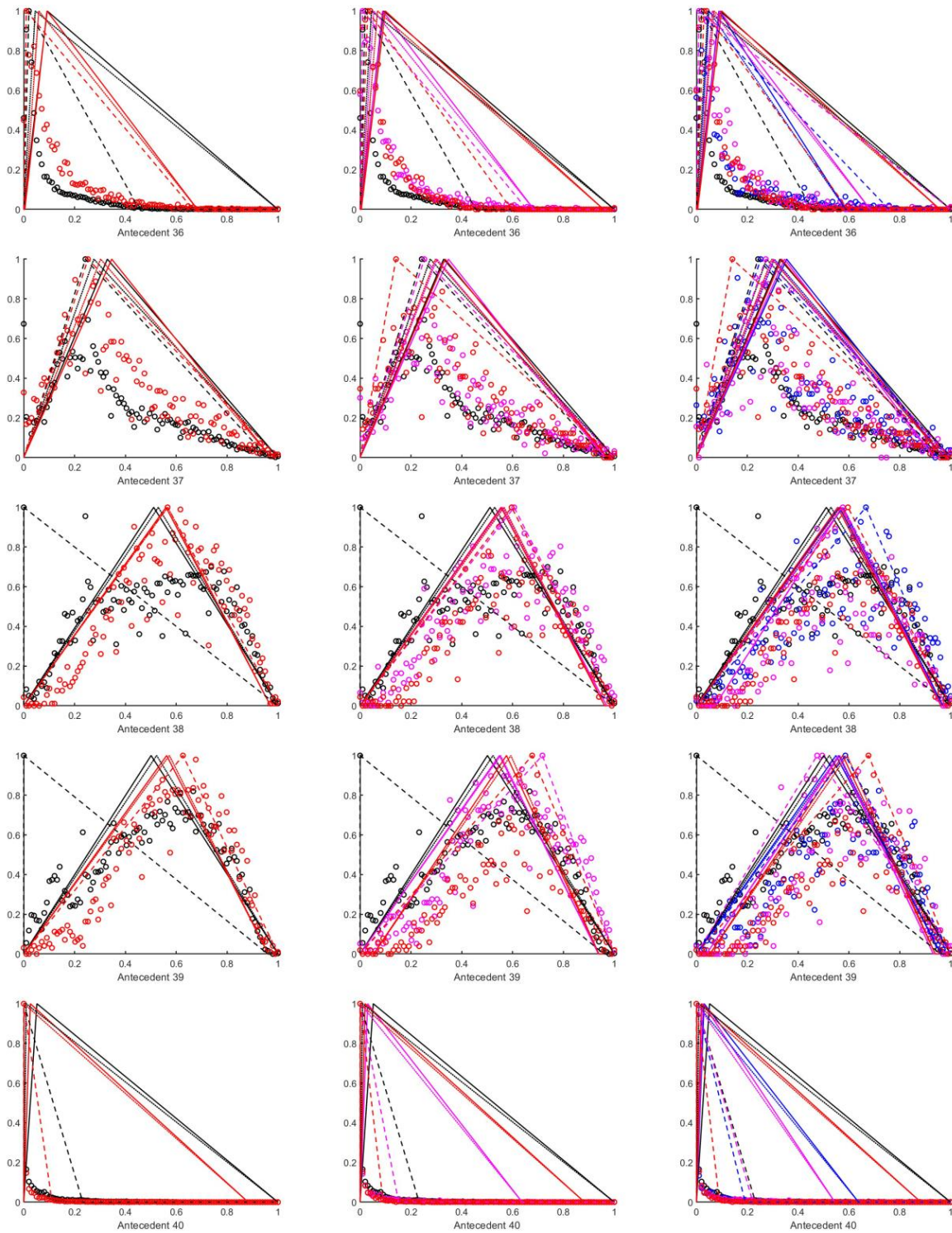


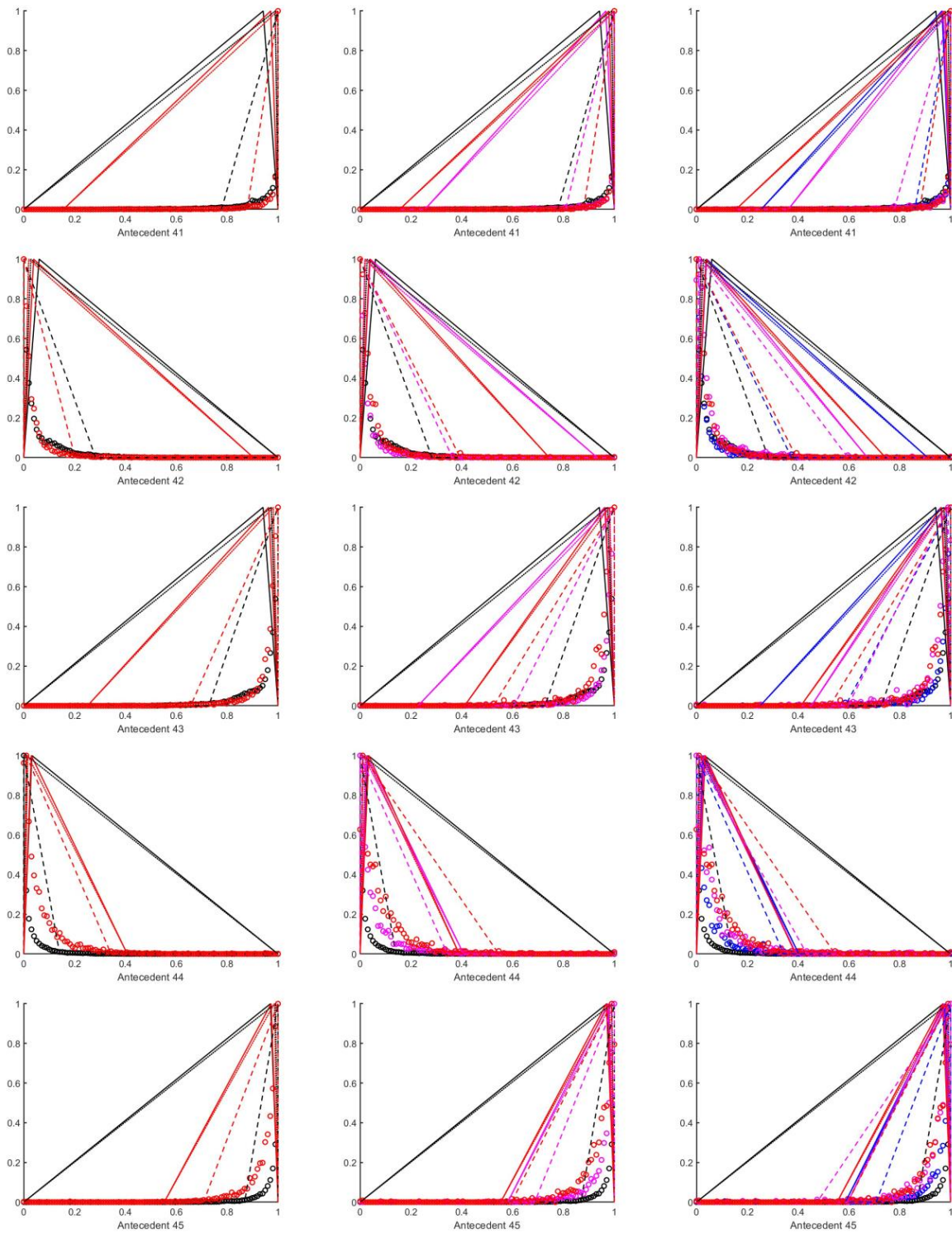




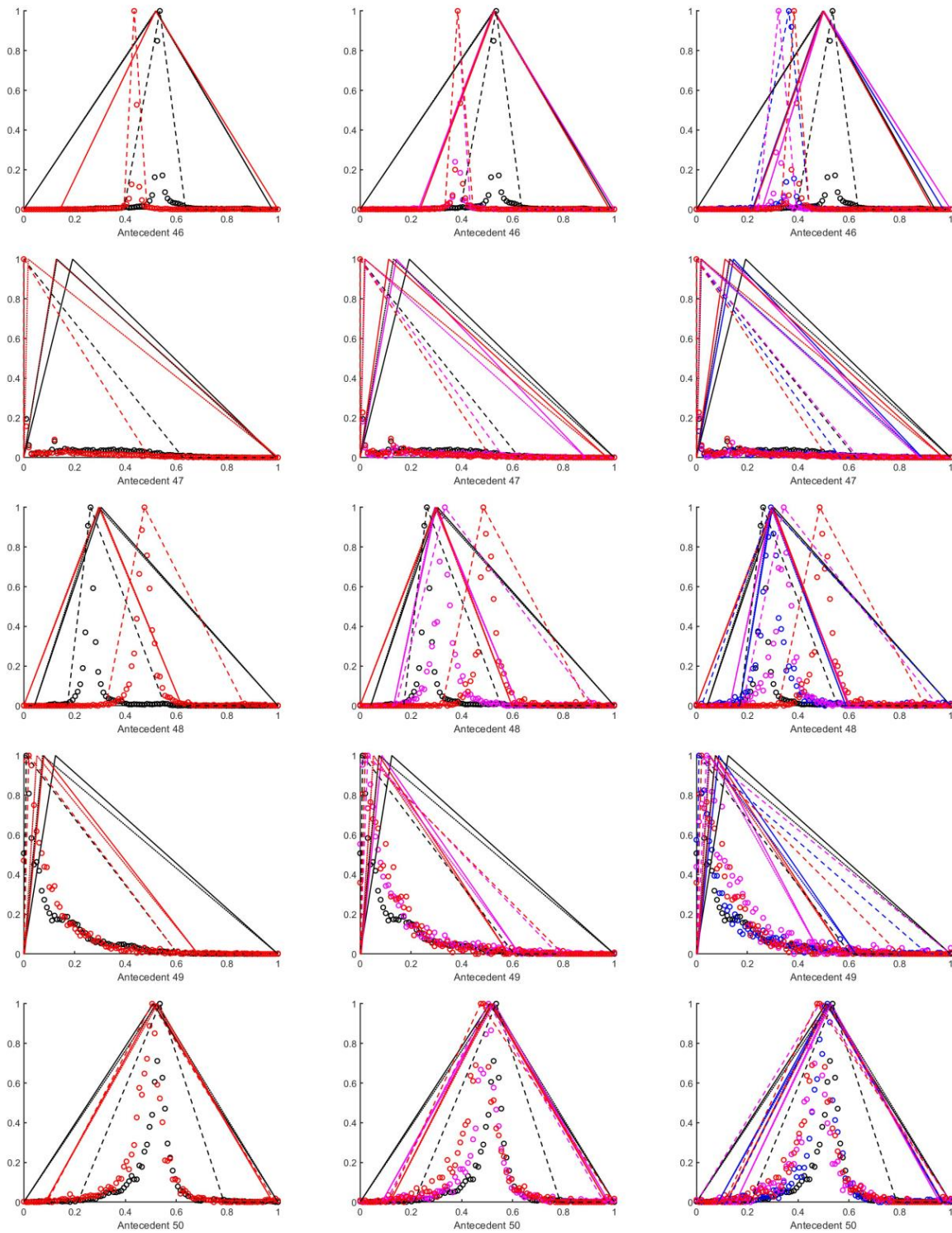


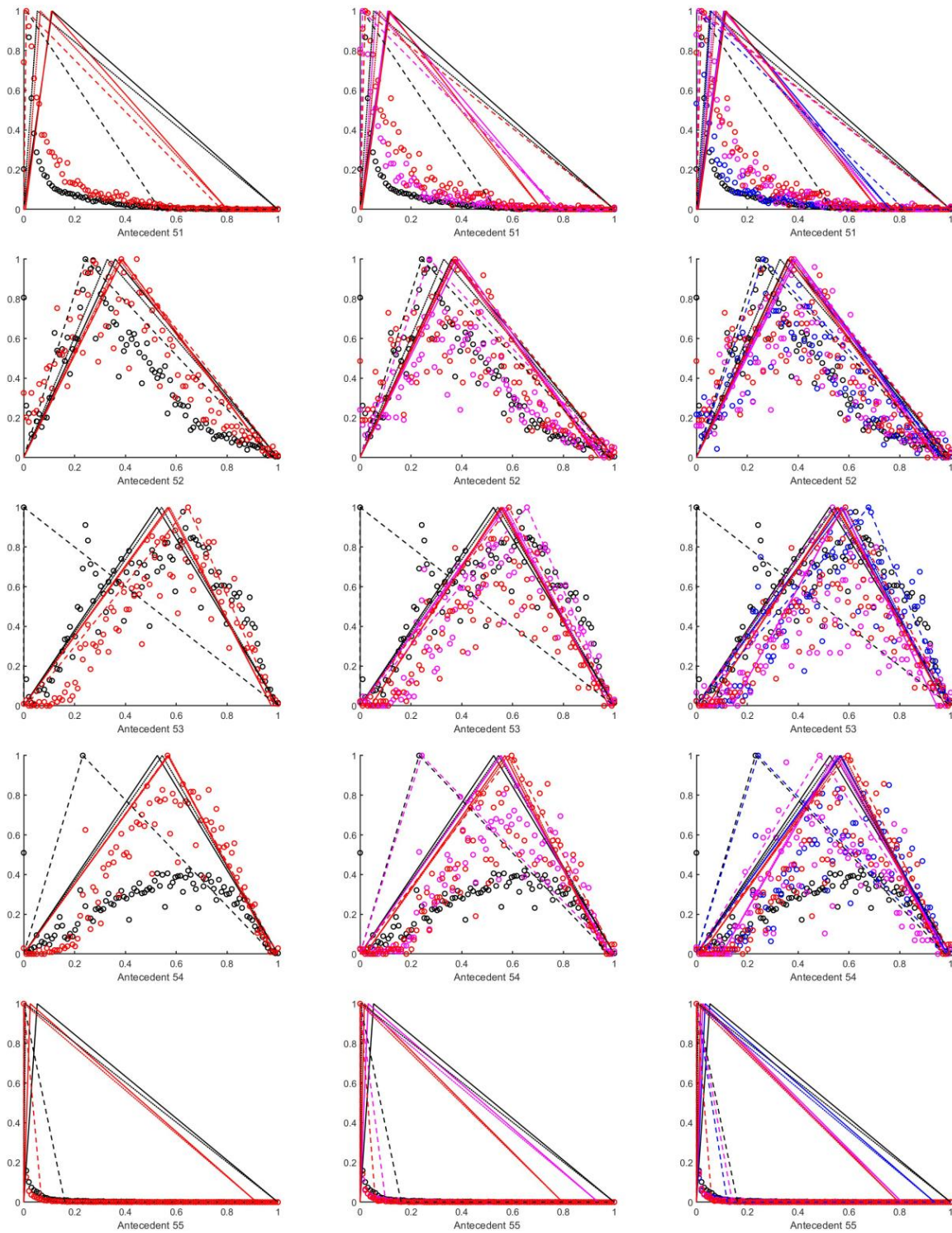


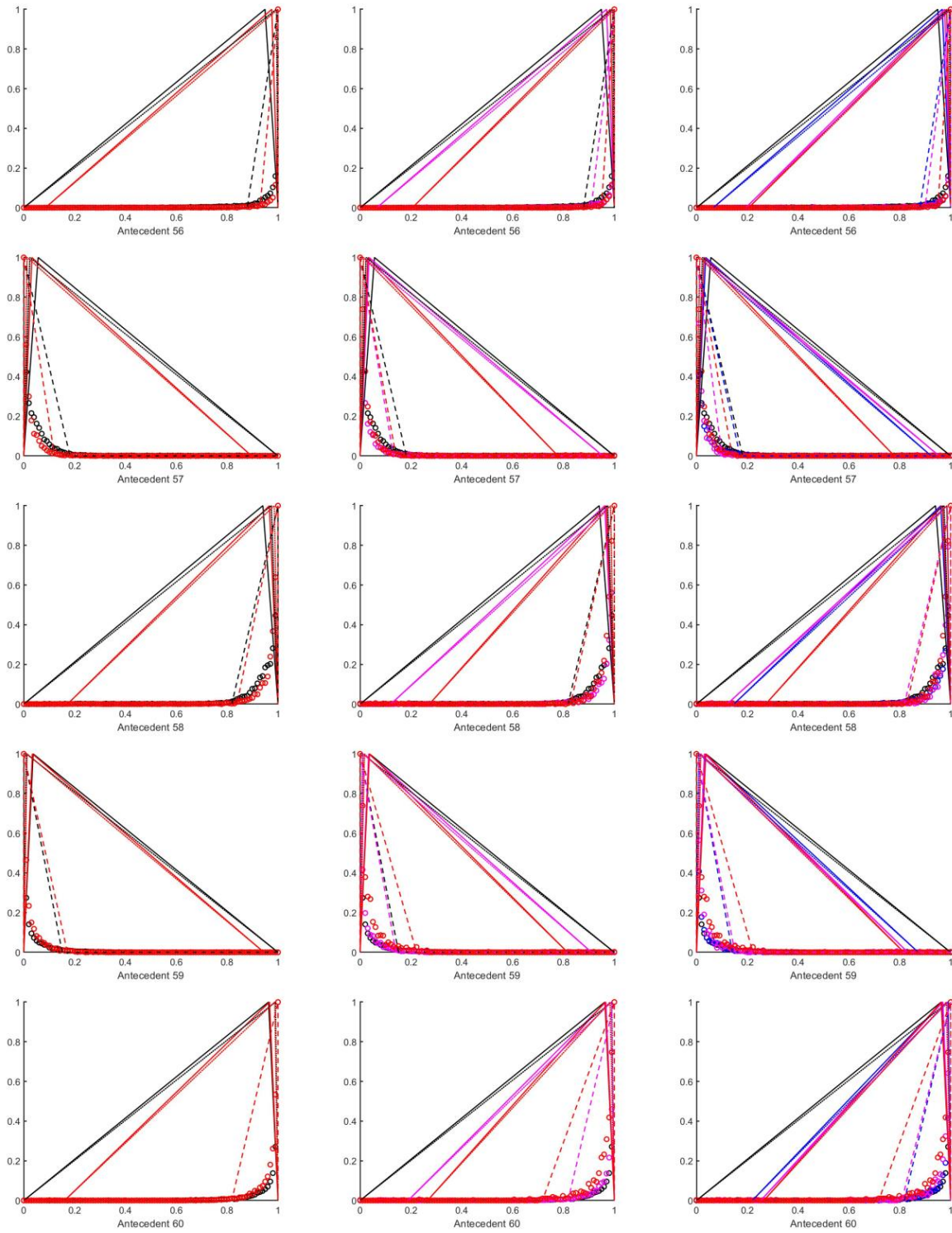




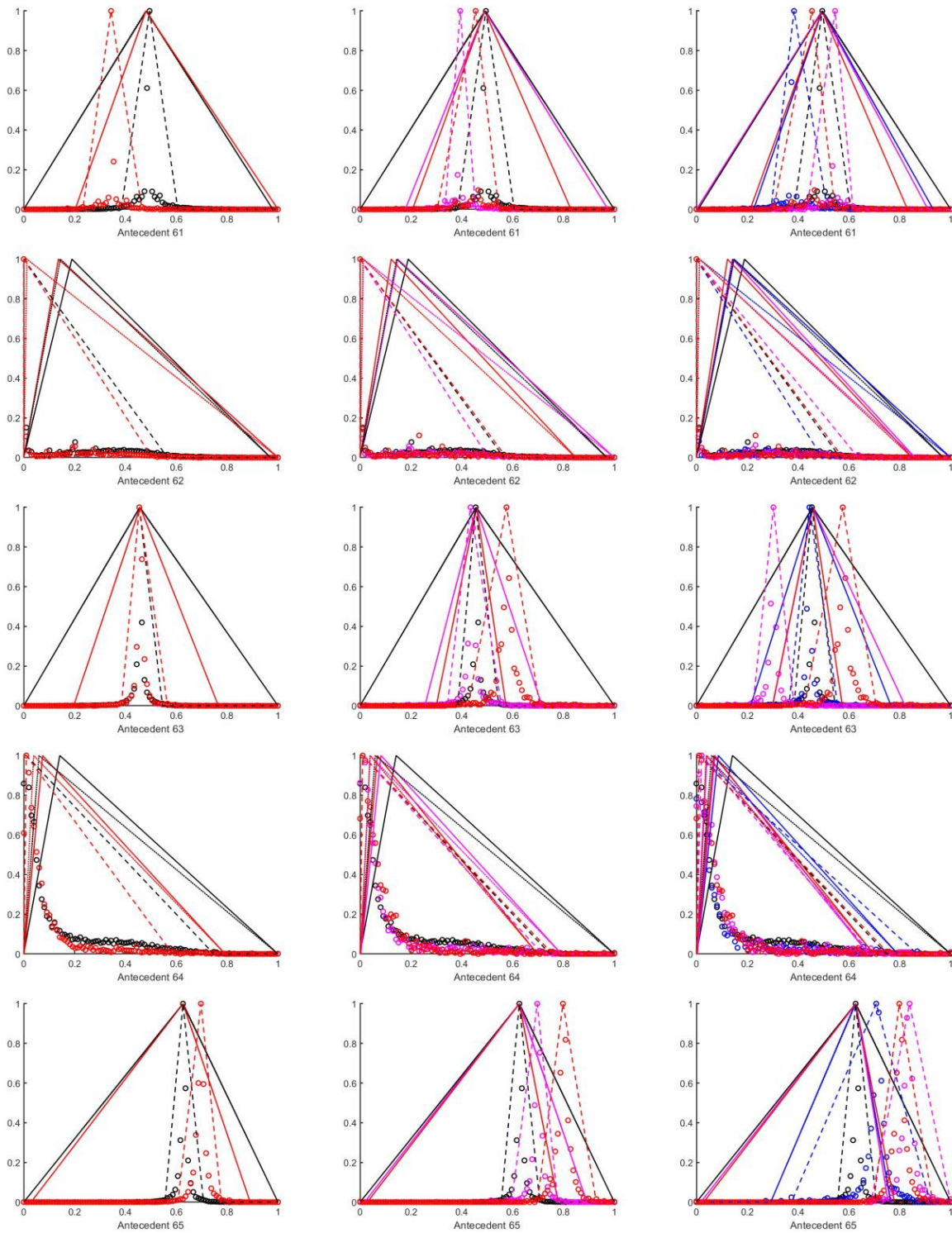


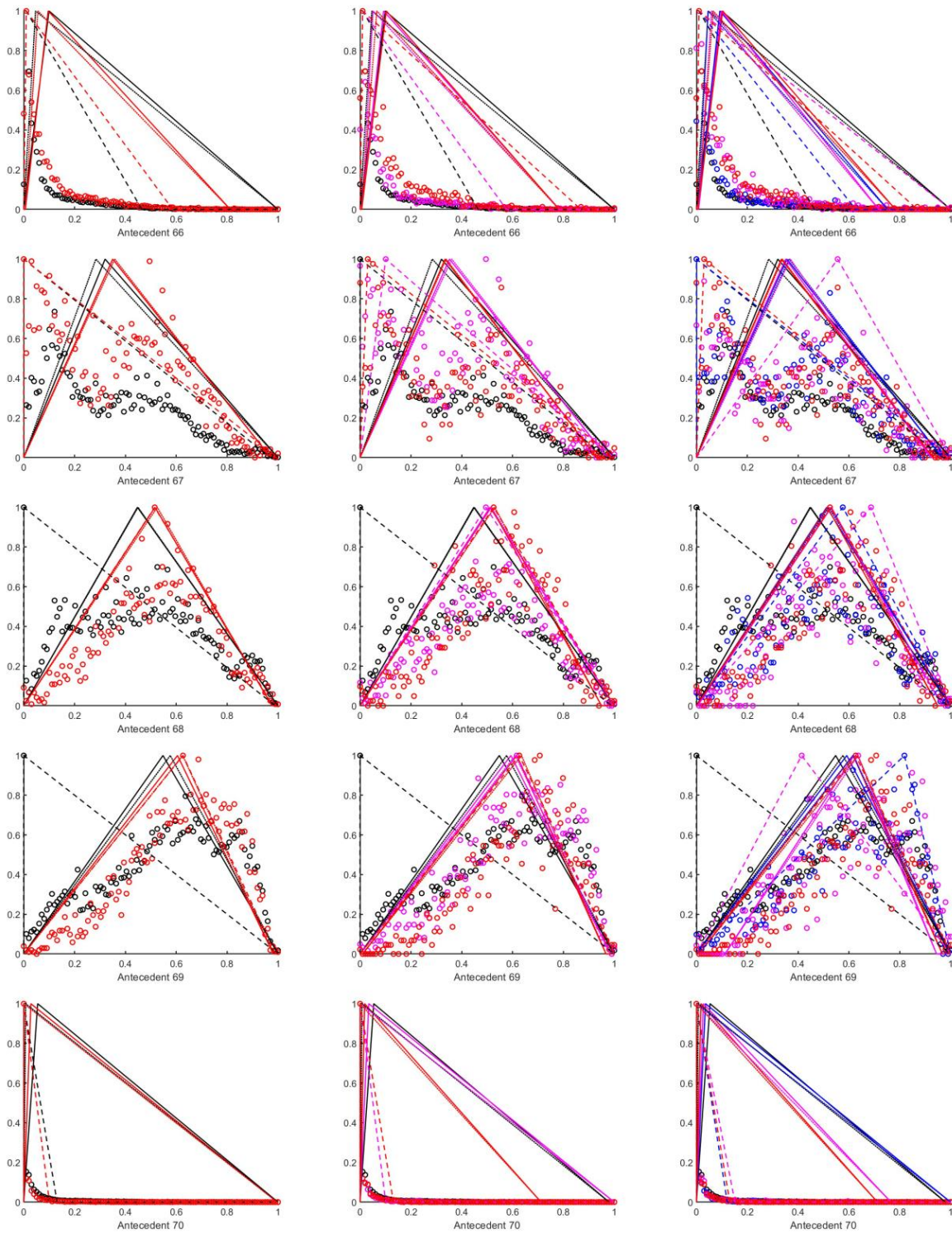




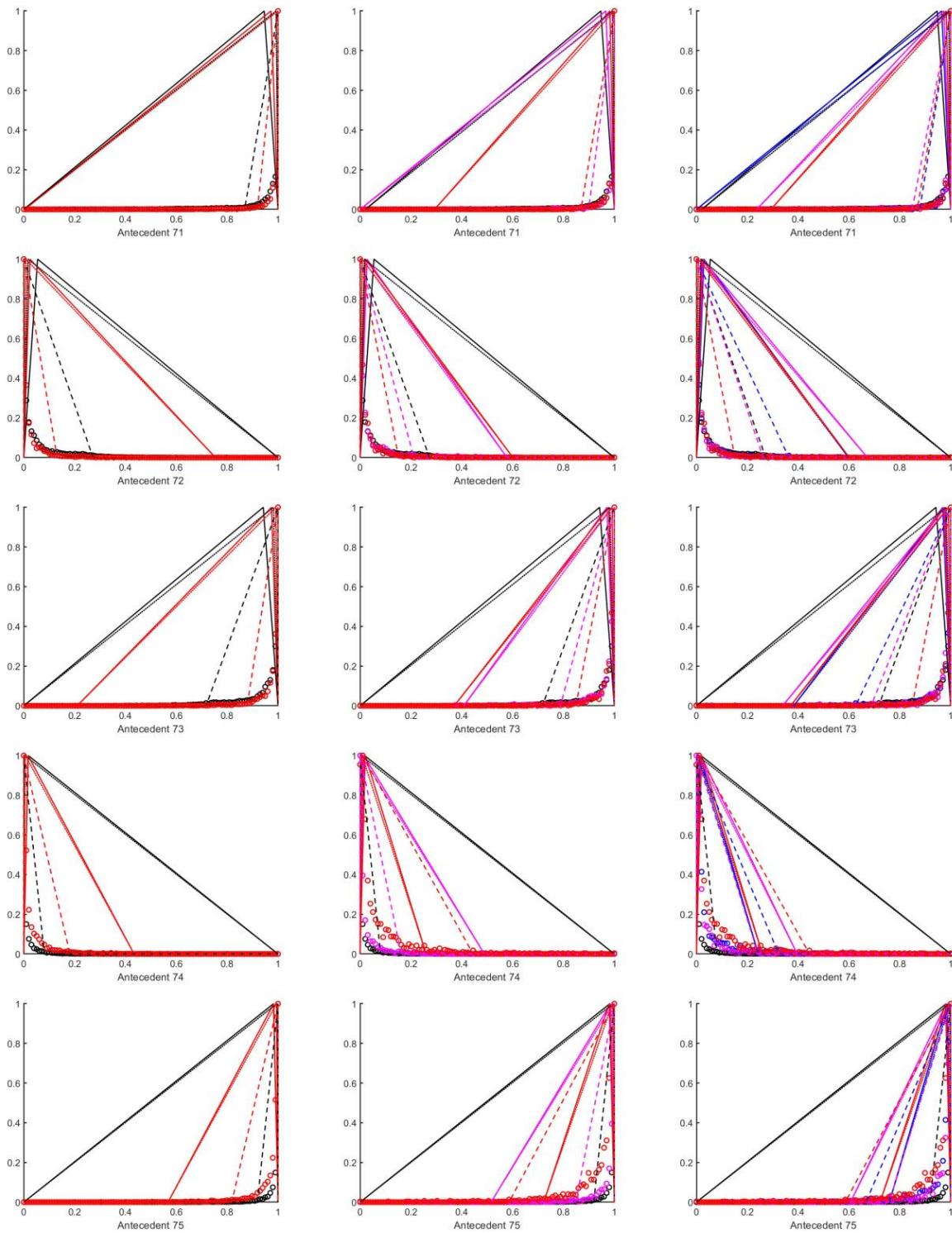


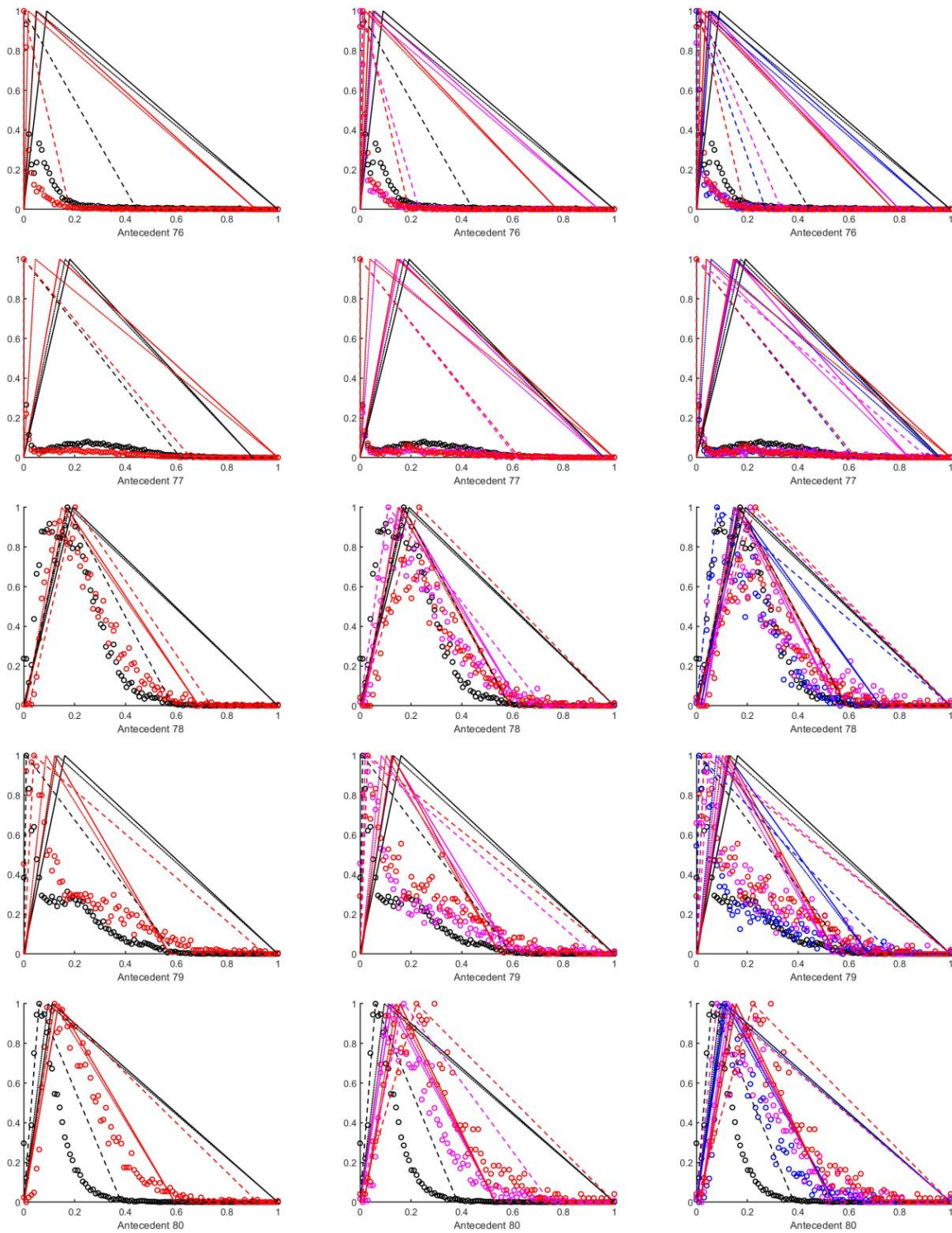




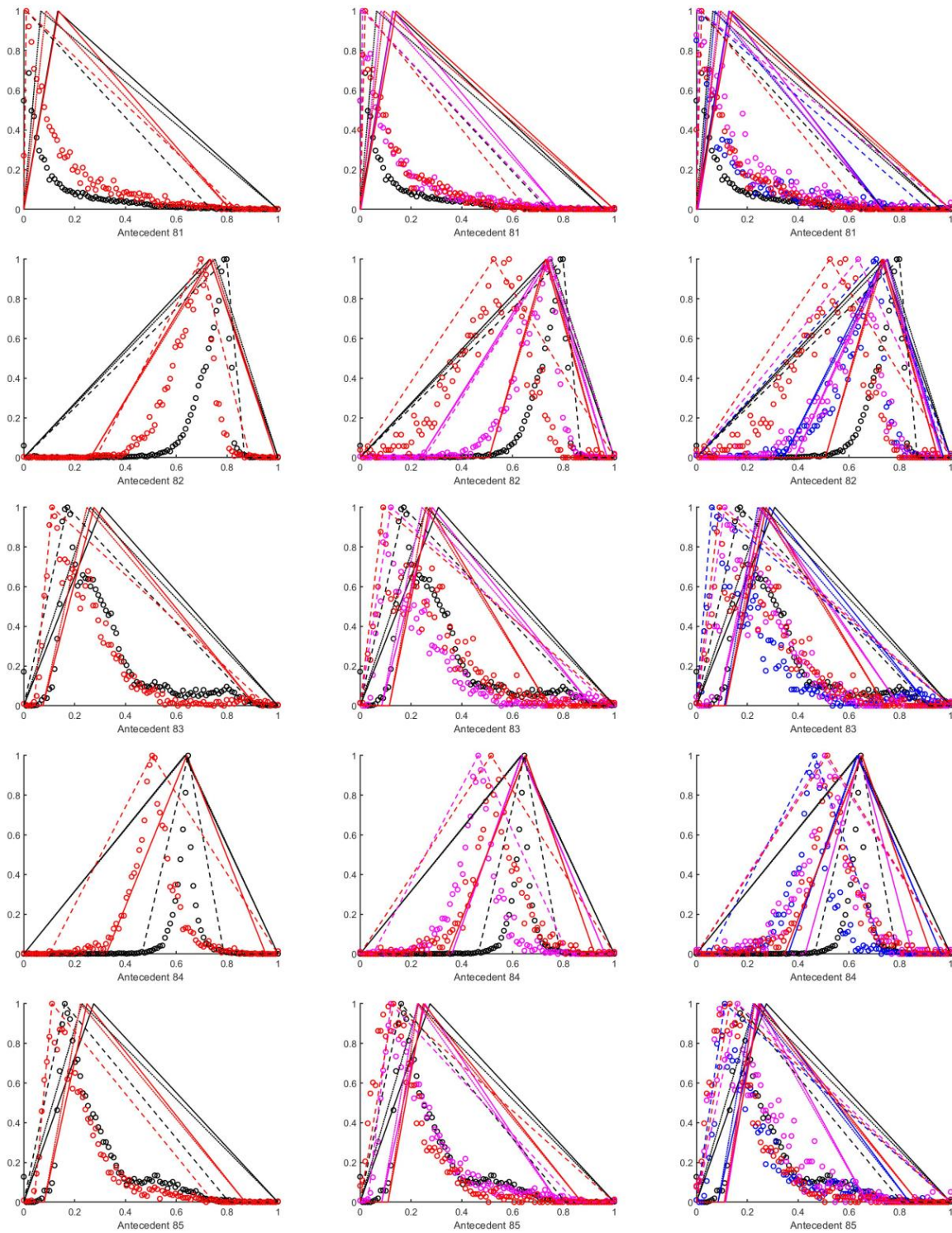


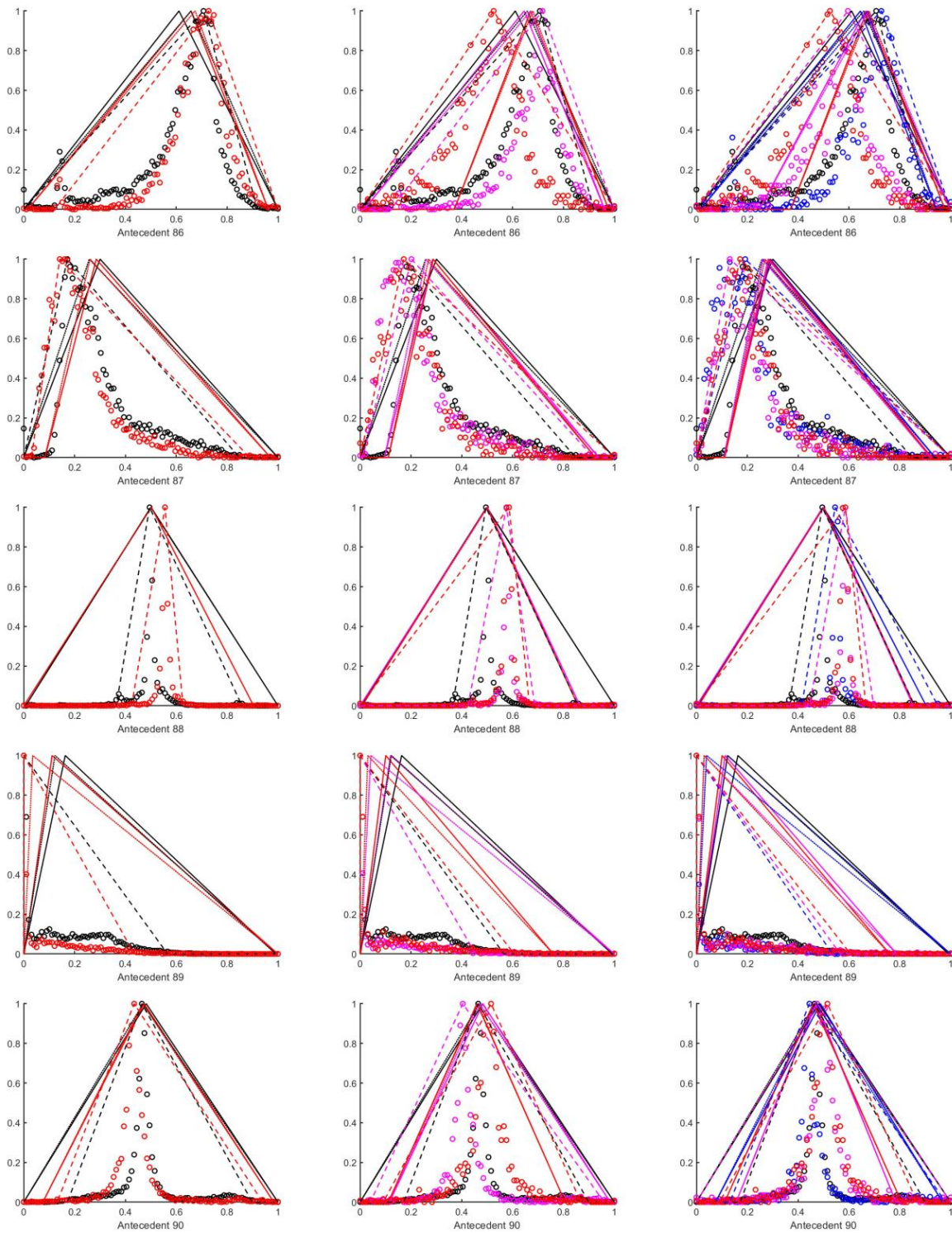




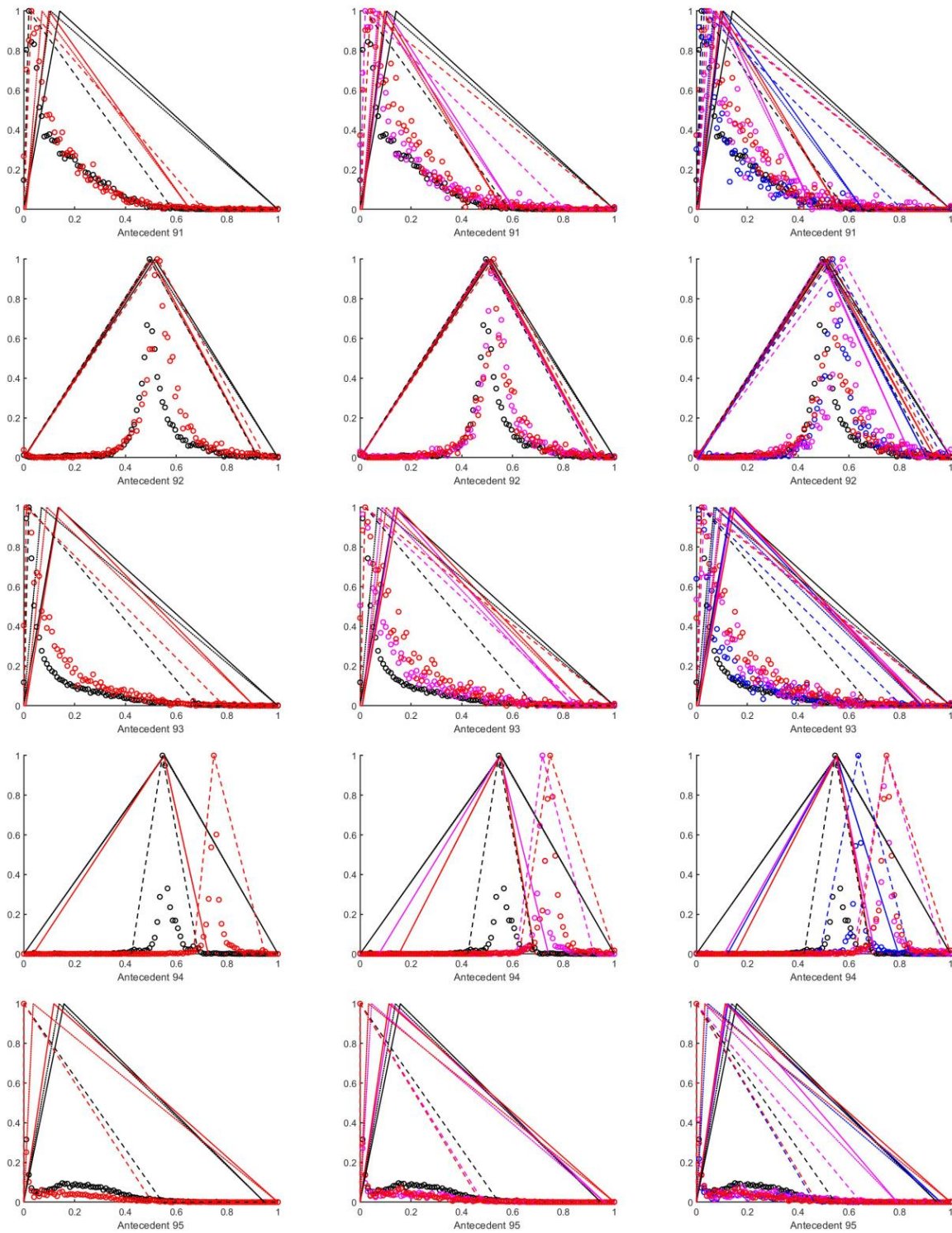


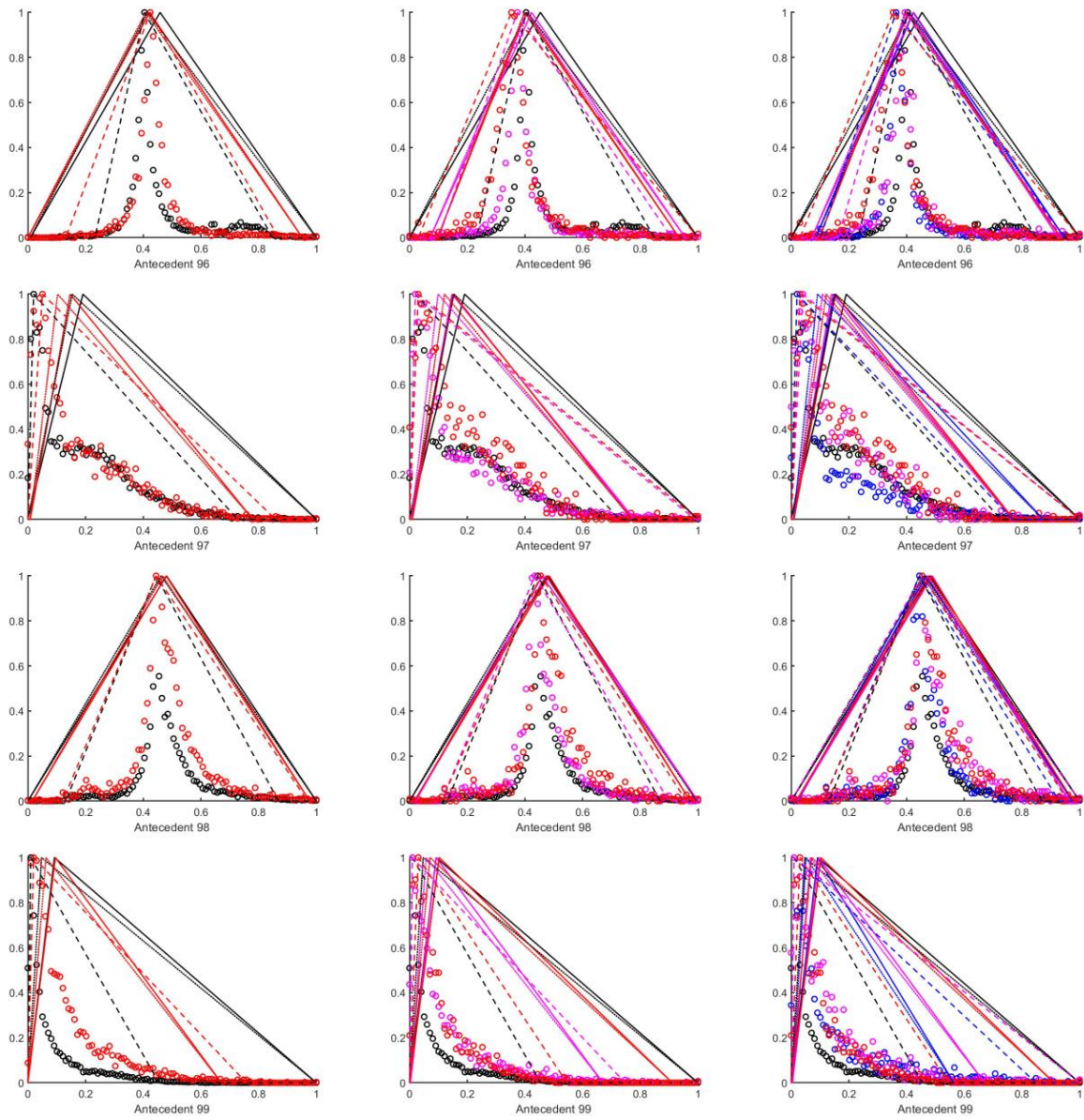














## Appendix C

All test samples and their masks which were used in Section 3.4 are given in this appendix.

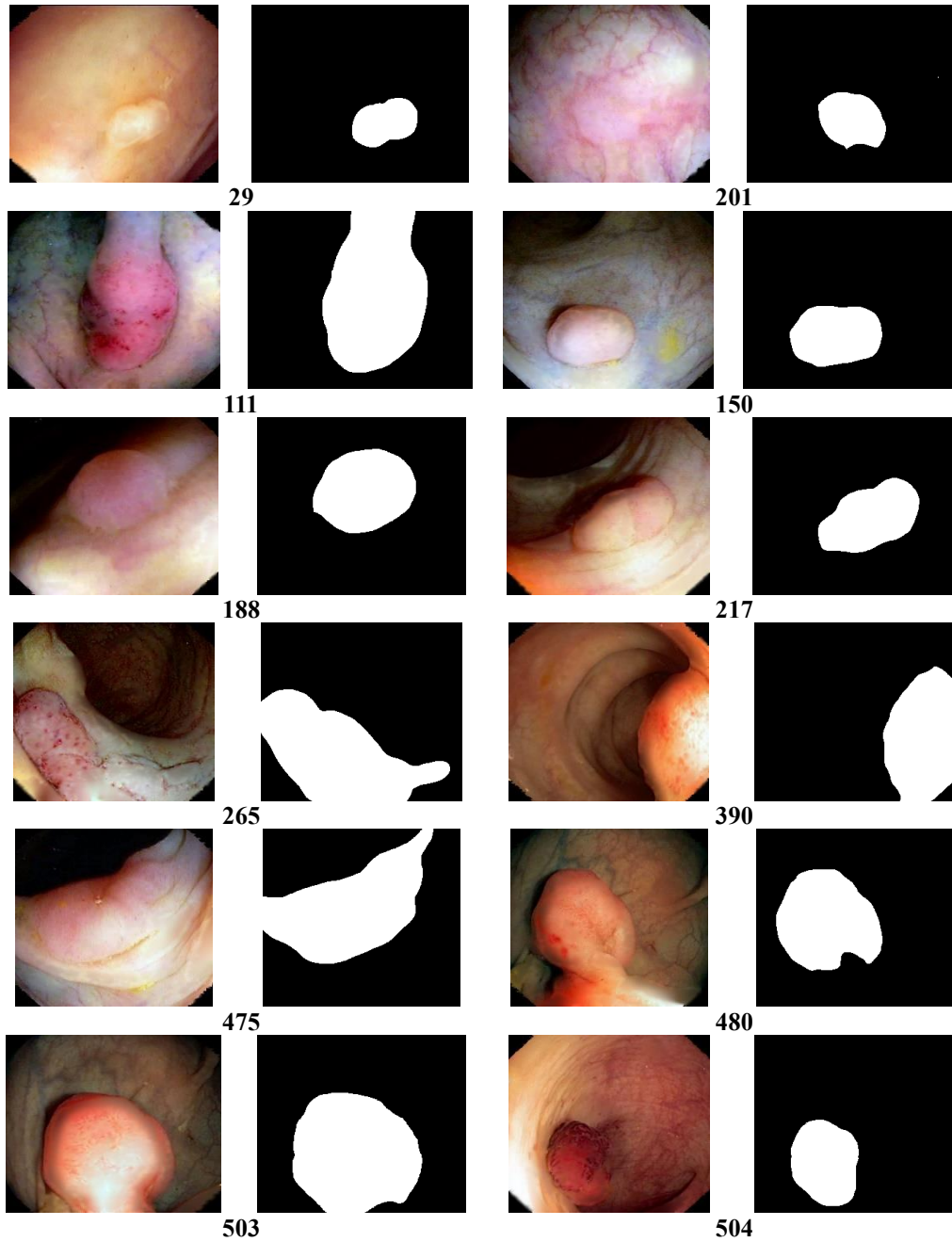


Figure C-1 The test samples for database CVC-Clinic and their ground truth masks

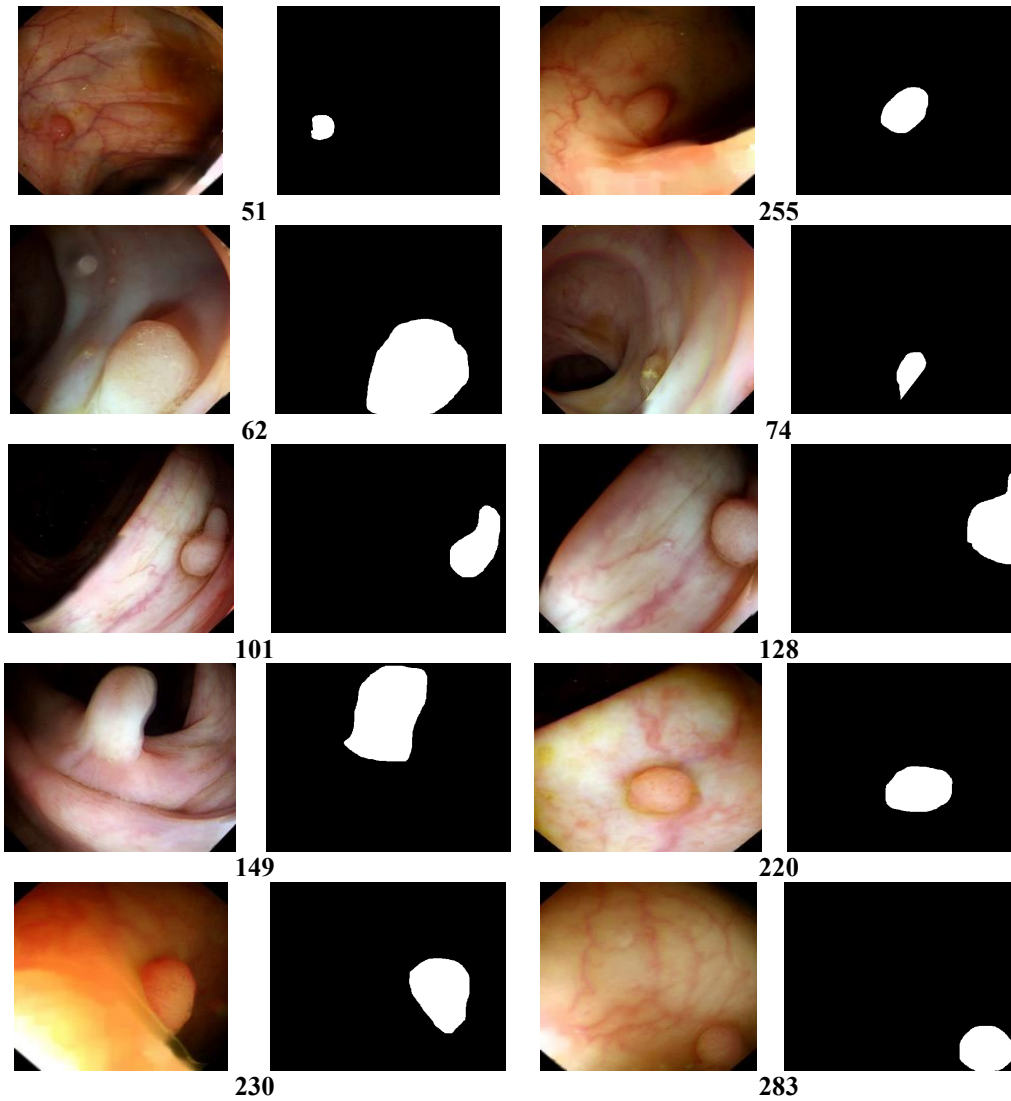
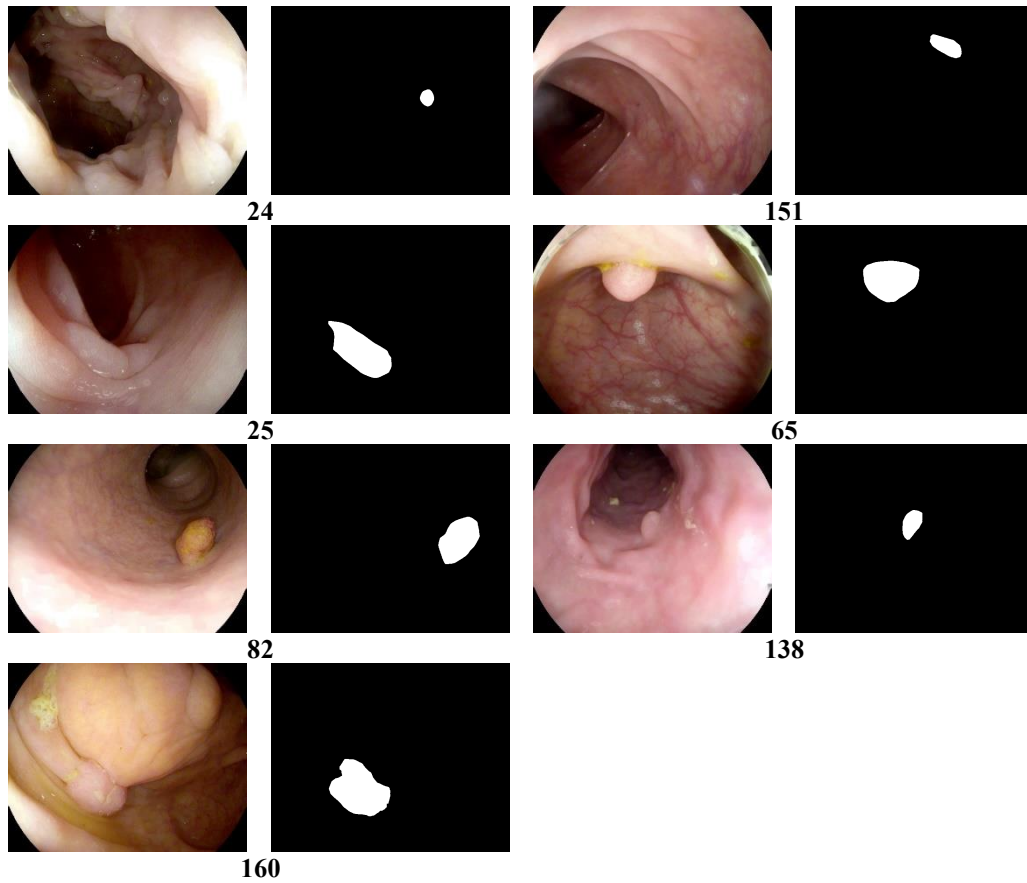


Figure C-2 The test samples for database CVC-Colon and their ground truth masks



*Figure C-3 The test samples for database ETIS-Larib and their ground truth masks*

## Appendix D

Abbreviations and nomenclature table for the proposed method of Chapter 3 is given in this appendix.

$R_{calc}$ (3.14)	A ratio to represent the quality of the edge detection method regarding the calculation efficiency of Hough transforms and polyp detection
$R_{edge}$ (3.15)	A ratio to measure the goodness of the edge pixels finding the ideal polyp contour
$\sigma$	$\sigma = 3,5,7,9,11,13$ and $15$ are the chosen values for the width of the voting membership function of the fuzzy Hough transforms
$P_p$	$P_p = 50\%, 60\%, 70\%, 80\%$ , and $90\%$ are the peak percentage values of the selected local maximum thresholds of the global maximum of the votes
$N_{total}$	The total number of the final resulting circles
$N_{ring}$	The number of final circles within the ring mask
$A_r$ (3.16)	A metric to measure the effectiveness of finding polyp-related final circles
$x_{min}, x_{max}$	The minimum and maximum coordinates of the ground truth mask points in $x$ direction
$y_{min}, y_{max}$	The minimum and maximum coordinates of the ground truth mask points in $y$ direction
$r_{avg}$ (3.17)	Average radius of the polyp mask
$(c_x, c_y)$ (3.18)	The $x$ and $y$ coordinates of the center of the polyp mask
$\Delta_{radial}$ (3.19)	The radial displacement of mask contour point $(m_x, m_y)$
$\delta_{roundness}$ (3.20)	The maximum of the roundness error
Original	The full Hough transforms for all the edge points
Wide	The restricted Hough transforms for the edge points with normalized gradient values within a wide threshold interval $[0.06, 0.3]$
Thin	The restricted Hough transforms for the edge points with normalized gradient values within a thin threshold interval $[0.08, 0.2]$