

A Complete System for Candidate Polyps Detection in Virtual Colonoscopy

Marcelo Fiori, Pablo Musé, and Guillermo Sapiro

Abstract—Computer tomographic colonography, combined with computer-aided detection, is a promising emerging technique for colonic polyp analysis. We present a complete pipeline for polyp detection, starting with a simple colon segmentation technique that enhances polyps, followed by an adaptive-scale candidate polyp delineation and classification based on new texture and geometric features that consider both the information in the candidate polyp location and its immediate surrounding area. The proposed system is tested with ground truth data, including flat and small polyps which are hard to detect even with optical colonoscopy. For polyps larger than $6mm$ in size we achieve 100% sensitivity with just 0.9 false positives per case, and for polyps larger than $3mm$ in size we achieve 93% sensitivity with 2.8 false positives per case.

Index Terms—Computed Tomographic colonography, computer-aided detection, colonic polyp detection, colon segmentation, curvature motion, differential features.

I. INTRODUCTION

Colorectal cancer is nowadays the second leading cause of cancer-related deaths in the United States (only surpassed by lung cancer), and the third cause worldwide [1]. The early detection of polyps is fundamental, allowing to reduce mortality rates up to 90% [2]. Polyps can be classified into sessile, pedunculated, and flat, according to their morphology. Pedunculated polyps are attached to the colon wall by a stalk, sessile polyps grow directly from the wall, and flat polyps have less than $3mm$ of elevation above the colonic mucosa [3]. Nowadays, optical colonoscopy (OC) is the most used detection method due in part to its high performance. However, this technique is invasive and expensive, making it hard to use in large screening campaigns.

Computed Tomographic Colonography (CTC), or Virtual Colonoscopy (VC), is a promising alternative technique that emerged in the 90's [4]. It uses volumetric Computed Tomographic data of the cleansed and air-distended colon. Distention is carried out by placing a thin tube into the rectum, and performing the colon insufflation with room air or carbon dioxide. It is less invasive than optical colonoscopy, and much more suitable for screening campaigns once its performance is demonstrated. However, VC is less popular than OC not only because it is a relatively new technique, but also because, contrarily to OC, it is not yet reimbursed by insurance companies. On the other hand, in OC, on the average only around 70% – 80% of the colon can be explored [5]. Incomplete studies due to obstructing colorectal lesions,

colon twists, or anatomical variations are not rare (5% to 15% of OC examinations) [6], and there is an additional important risk of colon perforation. In a large study by Kim *et al.* [7], where about 3000 patients went through OC and another 3000 through VC, seven perforations occurred in OC while none was reported in VC.

Nevertheless, it takes more than 15 minutes for a trained radiologist to complete a CTC study, and the performance of the overall optical colonoscopy is still considered better. In this regard, Computer-Aided Detection (CAD) algorithms can play a key role, assisting the expert to both reduce the procedure time and improve its accuracy [8], [9], [10], [11], [12], [13].

Flat polyps are of special interest because these are an important source of false negatives in CTC, and although there are different opinions, many authors claim that flat polyps are around 10 times more likely to contain high-grade epithelial dysplasia¹ [14], [15], [16].

There are numerous discussions regarding the potential risks of the polyps according to their size. Even though some authors consider that “small” polyps may not represent risk, some gastroenterologists disagree [17]. Summers [18] claims that one of the major challenges in the field is in increasing sensitivity for smaller polyps, and Church [19] states that small adenomas can still be clinically significant and should not be ignored. At the same time, Bond [20] declares that the major disadvantage of VC is its current low performance for flat polyps.

The goal of the work presented in this paper is to exploit VC precisely to automatically flag (mark for attention of the expert) colon regions with high probability of being polyps, with special attention to results in challenging small and flat polyps. Towards this aim, we propose a complete pipeline that starts with a novel and simple segmentation step (segmentation of the colon surface/lumen). We then introduce geometrical and textural features that take into account not only the candidate polyp region, but the surrounding area at multiple scales as well. This way, our proposed CAD algorithm is able to accurately detect candidate polyps by measuring local variations of these features. The whole algorithm is completely automatic and produces state-of-the-art results. This paper extends our previous conference publication [21].

The rest of this paper is organized as follows. In Section I-A we briefly review prior related work and in Section I-B we present an overview of the whole proposed pipeline. We address the colon segmentation problem in Section II and the

M. Fiori and P. Musé are with the Instituto de Ingeniería Eléctrica, Universidad de la República, Herrera y Reissig 565, 11300, Uruguay.

G. Sapiro is with the Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina, 27708, USA.

¹An abnormality of development in cells that may become *cancer in situ* or *invasive cancer*.

feature extraction and classification in Section III. In Section IV we describe the classification and evaluation method, and in Section V we present numerical results. The discussion is presented in Section VI and we conclude in Section VII.

A. Virtual Colonoscopy CAD Review

Automatic polyp detection is a very challenging problem, not only because the polyps can have different shapes and sizes, but also because they can be located in very different surroundings. Most of the previous work on CAD of colonic polyps is based on geometric features, some using additional CT image density information, but none of them takes into account the (geometric and texture) information of the tissues *surrounding* the polyp. This local and adaptive differential analysis is part of the contributions of this work.

Early work on CAD methods by Vining *et al.* [22] is based on the detection of abnormal wall thickness. Since then, several different approaches were proposed. Most of them have a segmentation step first, and then the classification step itself. We briefly discuss the main ideas in each stage.

Segmentation

The most common segmentation techniques are based on region growing and thresholding methods. Among others, Yoshida *et al.* [23] use a Gaussian smoothing and thresholding. Summers *et al.* [24] use a region growing scheme and obtain the final segmentation by thresholding. Paik *et al.* [25] use thresholding as well, followed by morphological dilation. Chen *et al.* [26] use a modified adaptive vector quantization followed by region growing. Sundaram *et al.* [8] start from manually selected seeds and segment the region by thresholding with marching cubes. Franaszek *et al.* [27] use a modified region growing, fuzzy connectedness, and Laplacian level set segmentation to obtain a smooth surface. Chen *et al.* [28] also use a level set approach, whose velocity is determined by a Bayesian classification of the pixels.

Assuming that the value of each voxel is a mixture of different tissue types, Wang *et al.* [29] use the statistical expectation-maximization (EM) algorithm to estimate the parameters of these tissue types (they use air, soft tissue, muscle, and bone/tagged material). The voxels are classified as air, mixture of air with tissue, mixture of air with tagged materials, or mixture of tissue with tagged materials, and then dilation/erosion operations are performed.

Yamamoto *et al.* [30] use a vertical motion filter at the fluid level. Serlie *et al.* [31] use a sophisticated method to address the T-junction artifact with promising results.

Region growing and level sets seems to be the current chosen techniques for colon segmentation. However, not much work has been done in comparing the smoothing techniques (or the regularization term in the level set method) to see which one is more adapted to polyp detection. The segmentation technique here proposed is geared toward the subsequent step of polyp detection, and simultaneously segments and prepares the obtained surface for this task.

Classification

The main variations in this topic are both in the features used and in the classification method. Summers *et al.* [32]

detect polyps larger than $10mm$ by computing mean curvatures and sphericity ratio, and present results over a large screening patient population. Yoshida *et al.* [23] use the *shape index* (defined later in this paper) and *curvedness* as geometric features, applying fuzzy clustering and then using directional gradient concentration to reduce false positives. Paik *et al.* [25] also use geometrical features, computing the Surface Normal Overlap (SNO) instead of calculating curvatures. Wang *et al.* [33] compute a global curvature, extract an ellipsoid, and analyze morphological and texture features on this ellipsoid. They reach 100% sensitivity with a relative low false positives (FP) rate, using heuristic thresholds and texture features. Hong *et al.* [9] map the 3D surface to a rectangle, use 2D clustering, and reduce false positives with shape and texture features. Sundaram *et al.* [8] compute curvatures via the Smoothed Shape Operators method, and use principal curvatures and Gaussian curvatures to detect polyps. All these described techniques based on local geometric computations suffer from a high dependence on the regularity of the polyp shape itself, ignoring how pronounced it is with respect to the surrounding area. Using geometry alone is also very sensitive to the accuracy of the colon segmentation.

More recently, van Wijk *et al.* [34] proposed a Partial Differential Equation (PDE) motion that flattens only the polyp-like shapes, and then they consider the difference between the original and the processed images. The main drawback of this approach is that the PDE motion is not capable of flattening polyps which are already flat. Consequently, for flat polyps, the difference between the original surface and its smoothed version is too small to be detected by the algorithm. Ong *et al.* [35] use the neighborhood to compute curvatures; instead of using only a 1-ring neighborhood to approximate second derivatives and compute the curvatures, they use a larger region to reduce the effect of noise. Konukoglu *et al.* [36] propose a heat diffusion field method to characterize polyps. Proprietary algorithms [37], [38] have been reported as well, but with no better results than the methods mentioned above.

Complementary techniques

In addition to these works that tackle either the segmentation, classification or both as a complete CAD system, some other complementary techniques were proposed. These algorithms are located in between the segmentation and the classification steps, or after classification as a false positive reduction technique. For example, Götkürk *et al.* [39] proposed a technique to reduce the false positives based on features calculated from three random orthogonal sections, and then classifying with SVM. Konukoglu *et al.* [40] proposed to introduce a polyp enhancement stage, in between the segmentation and classification stages of the pipeline in [25]. The polyp enhancement is based on the heat equation, with the conductivity term adapted to the local geometric properties of the surface. Suzuki *et al.* [41] use artificial neural networks to reduce the false positives of the algorithm in [23] described above. The results are very promising, achieving 96.4% sensitivity (over 28 polyps) with 1.1 FP per case. However, this evaluation does not take into account small lesions and polyps submerged in fluid.

General comments

The list of references included in this section is by no means exhaustive. We do not intend here to present a full review of existing techniques, but to discuss a set of works that are representative of the diversity of the proposed approaches for polyp detection. Direct quantitative comparison of the classification results obtained by these methods is strictly speaking meaningless, since they were obtained using different databases.

Table I summarizes the results reported by the methods discussed above. Notice that these results were obtained using databases containing polyps larger than $6mm$ in size (or larger than $10mm$ in some databases).

To the best of our knowledge, no algorithm reported in the literature can detect small polyps properly. On the other hand, for polyps larger than $6mm$ in size, no algorithm can achieve 100% sensitivity with less than one false positive per case. It is crucial to keep improving these rates, this is the reason why VC is still an active research field.

Technique	# Polyps	Size	Sensit.	FPS
Summers <i>et al.</i> [32]	≈ 30	$> 10mm$	89.3%	2.1
Summers <i>et al.</i>	≈ 120	$> 6mm$	60%	8
Yoshida <i>et al.</i> [23]	12	$> 5mm$	100%	2
Paik <i>et al.</i> [25]	7	$> 10mm$	100%	7
Wang <i>et al.</i> [33]	61	$> 4mm$	100%	2.68
Hong <i>et al.</i> [9]	≈ 120	$> 5mm$	100%	3
Sundaram <i>et al.</i> [8]	20	$> 10mm$	100%	18
Sundaram <i>et al.</i>	122	$> 2mm$	80%	24
van Wijk <i>et al.</i> [34]	57	$> 6mm$	95%	5
van Wijk <i>et al.</i>	32	$> 10mm$	95%	4
Suzuki <i>et al.</i> [41]	28	$> 5mm$	100%	1.1
Bogoni <i>et al.</i> [37]	21	$> 5mm$	90%	> 3
Taylor <i>et al.</i> [38]	32	$> 6mm$	81%	13
Näppi <i>et al.</i> [11]	12	$> 5mm$	100%	2.4

TABLE I

NUMERICAL COMPARISON OF THE REVIEWED METHODS, INDICATING THE NUMBER AND SIZE OF THE POLYPS IN THE DATABASE, AND THE ACHIEVED SENSITIVITY WITH THE FALSE POSITIVES (FPS) PER STUDY.

B. Overview of the Proposed System

The main goal we are addressing in this work is to highlight/flag all the candidate polyps, so the radiologist can quickly check them. It is crucial to minimize the false negatives, keeping a reasonable false positives number. We achieve this by a four-steps process that is completely automatic and constitutes the entire end-to-end algorithm, from data to candidate polyps flagging.

The proposed system with its four steps is illustrated in Figure 1. The first step is colon segmentation, which takes as input the CT volume data, and produces a 3D mesh representing the colon surface. In the second step, from the segmented mesh we perform an adaptive-scale search of candidates in order to capture the appropriate polyp size, obtaining a set of candidate patches. In the third step, both the CT volume data and the segmented colon surface are used to compute geometrical and textural features for each candidate patch that was identified in the previous stage. The final step consists of a machine learning algorithm that uses the computed features to classify

patches as polyps or normal tissue. In the following sections we describe each of these steps in detail.

C. Paper contributions

In this work we introduce several novelties at every stage of the polyps detection system. These novelties result from the consideration of several alternatives. As it will be clear from the following, the solutions that were kept outperform other alternatives that were considered in this work, and significantly contribute to the quality of the results presented in Section V. These contributions are:

- We propose a simple segmentation approach, specifically designed to reduce tagged fluid artifacts;
- The smoothing PDE algorithm outperforms the classical curvature motions in terms of polyp enhancement and classification;
- Our proposed *adaptive-scale* candidate search allows to precisely delineate the polyp region;
- We propose to characterize the texture in the tissue by *Haralick* features, which have proved successful in other applications in the image processing and computer vision literature;
- Instead of only measuring texture or geometric features within a candidate region, we consider measures that explicitly take into account the tissue properties in the surrounding area. We call these *differential features*, and we show that their use improves the classification performance.
- We explicitly deal with the class imbalance problem (polyp vs non-polyp) in the learning step. We consider three machine learning techniques that were specifically designed to deal with this problem. The best classification results were obtained with *cost sensitive* learning.

All these contributions will be detailed in the next sections.

II. COLON SEGMENTATION

The segmentation of the colon surface, which is critical in particular to compute geometric features, is divided into two parts: a pre-processing stage for dealing with the air-fluid composition of the colon volume, and a second stage that consists on smoothing the pre-processed image and obtaining the final colon surface by thresholding the smoothed volume. The overall procedure here presented is very simple and computationally efficient, leading to the state-of-the-art classification results later reported.

A. Classifying CT regions

All the cases from the used database have the same preparation, which includes solid-stool tagging and opacification of luminal fluid, and can be noticed as the white liquid in Figure 2. One of the strongest difficulties concerning the segmentation of the colon from abdominal volumes in CT is the presence of this tagged fluid and its interfaces with air and tissue. Figure 2 shows a CT slice and its pixel values over the highlighted vertical profile. At first sight there are three clearly distinguishable classes: the lowest gray levels correspond to

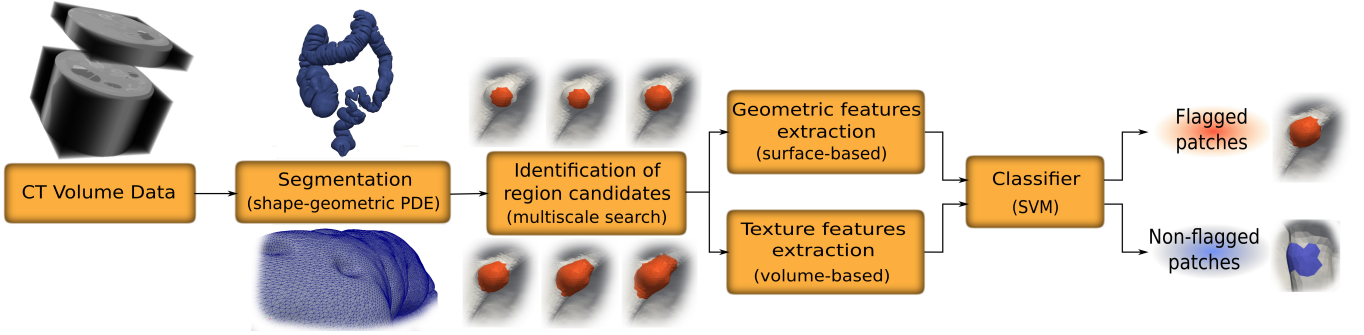


Fig. 1. Basic pipeline of the proposed polyp flagging system.

air, the highest levels correspond to fluid, and the middle gray values correspond to tissue. Nevertheless, there are around 6 interface voxels between air and fluid whose gray values, due to continuity, lie within the normal tissue range. Therefore, a naïve approach based on gray values only, ignoring the physical nature of the tissue and its environment, is not suitable for proper tissue classification and segmentation.

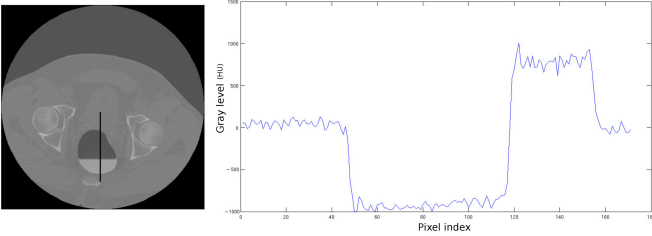


Fig. 2. CT slice and its different gray values for air, fluid and normal tissue, along the vertical profile.

The goal of this first pre-processing stage is to obtain an initial segmentation, as close as possible to the actual colon boundary (border). This border is diffuse due to the CT resolution, and if we adopt a binary segmentation approach (interior/exterior) the border will be necessarily bumpy. We prefer not to lose information about the gray decay throughout the border, so we compute a volume u_0 intended to have homogeneous values in the colon interior and exterior, and a smooth transition between them.

In order to do that, it makes sense to assign to each voxel the likelihood of being air, fluid or air-fluid interface. The air and fluid likelihoods are estimated as follows. We pick any CT study and manually segment the regions corresponding to air (class w_1) and tagged fluid (class w_2). Then, their gray level distributions are learned using standard kernel density estimation techniques [42]. This step is done just once. Then, at runtime, for each new study, the gray level histogram is computed. These histograms are trimodal, with peaks corresponding to air, tissue, and tagged fluid. We assume that the distribution of the gray values, both for air and tagged fluid, are shifted versions of the previously learned ones (this assumption was experimentally validated). Then, the air and tagged fluid distributions for the considered study, are determined by localizing their peaks. These shifted functions

are used to assign the air and tagged fluid likelihood values to the voxels.

Note that this assignment fails on the air-fluid and air-fluid-tissue interfaces. The basic idea for assigning a value to these voxels is presented next.

Here we take advantage of the physics of the problem, and in particular of the gravity and the position of the patient: the subject is laid horizontally so the interface between the fluid and the air is a plane parallel to the floor. The voxels situated on the interface then have a large gradient in the vertical direction. However, the transition is about 6 voxels wide for the standard data resolution used in this work, so the computations should be done taking this into account. Additionally, if a given voxel belongs to the interface layer, it is expected that at least half of the neighbor voxels at the same horizontal plane also belong to the interface layer.

The implementation of these criteria is as follows. A cubic neighborhood around each voxel \mathbf{x} is considered, and for each one of the “columns” that result of fixing the x and y coordinates, the air-likelihoods of the upper voxels and the fluid-likelihoods of the lower voxels are accumulated. If the tested voxel belongs indeed to the interface layer, then all these air and fluid likelihoods will be high. The value $IC(\mathbf{x})$ that represents the confidence level of \mathbf{x} being an interface voxel is then an increasing function of this accumulated measures. Algorithm 1 provides a pseudo-code that represents this procedure.

In order to guarantee that a high value is associated to every notoriously interior voxel, we assign to the initial segmentation u_0 the maximum of these three values, namely, the air and fluid likelihoods and the interface confidence level: $u_0(\mathbf{x}) = \max(p(x|w_1), p(x|w_2), IC(\mathbf{x}))$. This way, internal voxels (air, fluid and interface) will have values close to one, external voxels will have values close to zero, and the transition will be smooth across the actual colonic wall.

After the computation of the initial segmentation u_0 , some spurious (isolated) voxels may have high values (bones for example, or simply noise), so we clean the initial segmentation by keeping the connected components² containing some chosen voxels used as seeds. It is important for the system to be automatic, so no human intervention should be needed to

²Actually, since the initial segmentation u_0 is not binary, a (conservative) threshold of 0.6 is considered to separate the connected components.

```

for each voxel (x,y,z) do
  sum=0;
  for i=-1 to 1 do
    for j=-1 to 1 do
      for k=1 to 2 do
        sum += p((x+i,y+j,z+k)|w1);
        sum += p((x+i,y+j,z-k)|w2);
      end
    end
  end
  IC(x, y, z) = sum/18;
end

```

Algorithm 1: Computation of interface confidence level.

choose these seeds. In order to do that, the seeds are automatically detected by choosing the voxels with largest values of $IC(\mathbf{x})$, since these high values only occur at the interface between air and fluid. Voxels with the largest values of the air likelihood can be used too. After the first seed has been used and a connected component has been obtained, we look for a new seed outside this component, apply the filter with this seed, and append the new corresponding connected component to the previous ones. This iterative procedure continues until no seed outside the extracted connected components is found or a maximum number of connected components is reached. Usually, only two iterations are sufficient. This way, the segmentation step is able to handle the lumen discontinuities problem and to obtain the multiple pieces of the colon that might be disconnected.

Figure 3 shows a slice of the original volume data and the same slice of the initial segmentation u_0 .

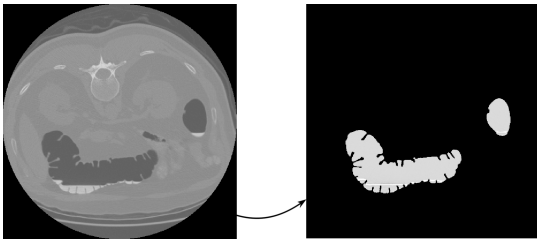


Fig. 3. CT slice and its corresponding slice on the initial segmentation u_0 . Patient in prone position.

The air-fluid-tissue joint may create artifacts in the segmentation. This is a critical point, not only because of the quality of the segmentation, but mainly for the potential of yielding several false positives in the polyps detection step. It is not rare that segmentation algorithms result in “gutter-like” shapes along this interface. This improvement is crucial, since if small oscillations occur along the “gutter” (which is expectable due to the resolution of the CT image), artifacts with polyp-like shape are produced, and this, of course, degrades the performance of the whole CAD system. We paid particular attention to this issue while designing the segmentation algorithm: the IC computation allows to avoid these artifacts. Figure 4 illustrates the performance of

our segmentation method compared to another version that presents some problems along this interface. The aim of the comparison is not to discuss which segmentation is better, but to show that our algorithm presents a smooth surface along the place where the gutter is expected to be.

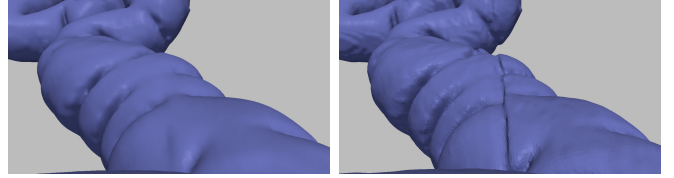


Fig. 4. Comparison of reduced artifacts in our segmentation (left) with a previously tested more standard version.

B. Smoothing and colon surface computation

In order to eliminate noise and to obtain a smoother colon surface after the segmentation stage, we proceed to smooth the initial segmentation u_0 . In this section we derive a PDE-driven smoothing technique that preserves the shape of the polyps, while obtaining a smooth enough surface to reliably compute local geometric features such as curvatures. Of course, the ultimate goal of the method is to simplify and to improve the polyp/non-polyp classification system. The effectiveness of the proposed approach will be assessed with experiments both qualitatively and quantitatively in Section V, where ROC curves obtained with the proposed PDE and other filtering alternatives will be compared.

We concentrate here on a family of smoothing PDEs of the form

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = \beta |\nabla u| \quad , \quad u(\mathbf{x}, 0) = u_0(\mathbf{x}) \quad , \quad (1)$$

where the initial volume u_0 results from the preprocessing technique described in the previous section. After a few iterations of this PDE evolution, the inner colonic wall will be extracted as a suitable iso-level surface of the resulting 3D image $u(\mathbf{x}, T)$. The choice of the number of iterations and the iso-level are not arbitrary and will be discussed in detail at the end of this section.

We recall that the Level Set Method [43] states that if $u(\mathbf{x}, t)$ evolves according to (1), then its iso-levels (level sets) satisfy

$$\frac{\partial \mathcal{S}}{\partial t} = \beta \vec{\mathcal{N}} \quad , \quad (2)$$

where \mathcal{S} is any iso-level surface and $\vec{\mathcal{N}}$ its unit normal. This geometric view enables to design β to fulfill a set of requirements we will impose to the surface evolution. In particular, we are interested in motions driven by the principal curvatures κ_{max} and κ_{min} .

With the mean curvature motion ($\beta = \mathcal{H}$), the polyps are flattened too fast, as shown in in Figure 5. A suitable variation of the motion by Gaussian curvature,³ namely the affine motion

$$\frac{\partial \mathcal{S}}{\partial t} = (K^+)^{1/4} \vec{\mathcal{N}} \quad \text{where} \quad K^+ = \max(K, 0) \quad ,$$

³The motion by Gaussian curvature $\beta = K$ has several problems with surfaces containing non-convex parts [44].

has a better behavior in general, but the results regarding polyp flattening are comparable with the mean curvature.

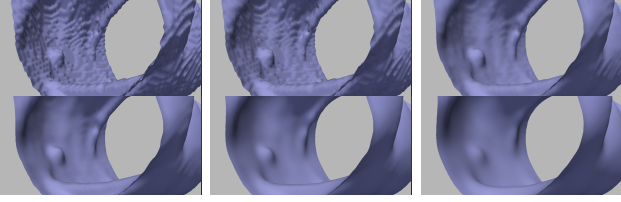


Fig. 5. Mean Curvature Motion: original surface and the result after 2, 8, 15, 30 and 50 iterations. Note how both the surface (as desired) and potential polyps (undesired) are smoothed and flattened.

A classical motion that appears to be well suited for our problem is the motion by minimal curvature [45]. Indeed, polyps have a curve of inflection points all around it, separating its upper and lower sections (see Figure 6). Along this curve, the minimal curvature is $\kappa_{min} = 0$, and therefore this section of the polyp does not move (or moves very slowly), so intuitively under this motion the polyps should persist longer. In our application, this evolution yields very good results in terms of both surface smoothing and polyp enhancement.

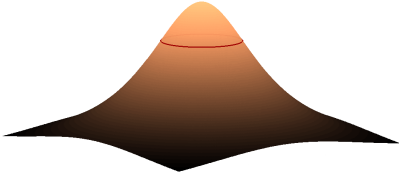


Fig. 6. Polyp with a curve of inflection points (in red), where $\kappa_{min} = 0$.

This PDE can be modified to obtain better results in terms of polyp detection. We now propose a set of modifications that lead us to the proposed smoothing evolution equation, and we show qualitative results to support this claim. The improvement in terms of polyp detection performance is discussed in Section V.

The first modification is inspired by the exponent $1/4$ of the affine motions in dimension 3, and yields to the following curvature motion equation:

$$\frac{\partial \mathcal{S}}{\partial t} = \kappa_{min}^{1/4} \vec{N}.$$

Figure 7 shows the result after a few iterations; comparing to Figure 5, it can be readily seen that this motion achieves a better trade-off in terms of noise reduction and polyp preservation. Figure 8 evidences the difference with a comparative image: the result of the motions by κ_{min} and $\kappa_{min}^{1/4}$ are shown in gray and in orange, respectively. On the polyp protrusion, the orange surface is above the gray surface, while the opposite is observed in the surrounding area. This shows that the evolution by $\kappa_{min}^{1/4}$ leads to better polyp enhancement.

The second modification that we introduce is based on the idea of preserving the polyps qualities that we later use to identify them. Towards this aim, we take into account a surface property that will be used in the feature extraction stage as

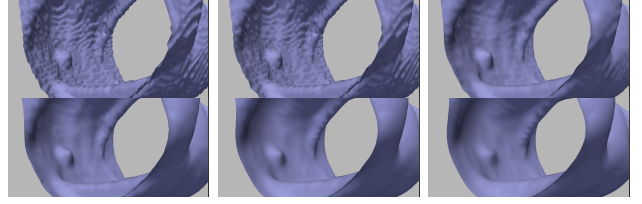


Fig. 7. Evolution by $\kappa_{min}^{1/4}$: original surface and the result after 2, 8, 15, 30 and 50 iterations.

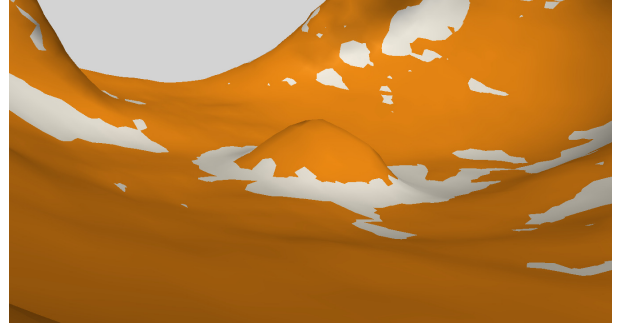


Fig. 8. Comparison between evolutions. Motion by k_{min} in light gray vs. motion by $k_{min}^{1/4}$ in dark gray. Both surfaces are overlaid, so sections that are not visible are hidden below the other surface.

well. A measure of the local shape of a surface is the so-called *shape index* [46],

$$SI := -\frac{2}{\pi} \arctan \left(\frac{\kappa_{max} + \kappa_{min}}{\kappa_{max} - \kappa_{min}} \right).$$

A complementary measure called *curvedness* C , is defined as

$$R := \sqrt{\frac{\kappa_{max}^2 + \kappa_{min}^2}{2}}, \quad C := \frac{2}{\pi} \ln R.$$

The $(\kappa_{max}, \kappa_{min})$ plane is then transformed into the (SI, C) plane. While the value of SI is scale-invariant and measures the local shape of the surface, the value of C indicates how pronounced it is. Figure 9 shows different shapes and their corresponding shape index. Due to the chosen orientation, shape index values close to -1 (protrusions) are of special interest for polyp detection.

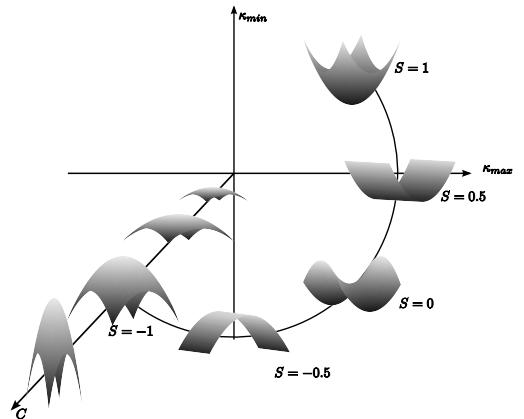


Fig. 9. Some shapes and their corresponding shape index values.

Special care must be taken when computing the shape index. Principal curvatures are usually derived from \mathcal{H} and K , however, as shown by [47], this may lead to inaccuracies that can be strongly amplified when computing the shape index. A much more stable estimation can be obtained by computing κ_{min} and κ_{max} directly from u (see [47]).

Back to the PDE motion, the next step is to include this information concerning the shape of the surface in order to make potential polyps evolve differently than the rest of the colon surface. More precisely, we modify the best motion so far ($\beta = \kappa_{min}^{1/4}$), in such a way that the resulting motion further enhances the potential polyps. In order to achieve this, we first need to characterize the potentially polyp points, and then modify the deformation function accordingly.

We define a function that acts as a multiplying factor to the term $\kappa_{min}^{1/4}$, making the surface evolve slower at the interest points. One option is to choose this function to depend on the shape index only, assigning low values to shape index near -1 , and values close to unity to other points. A smooth function $g(SI)$ verifying these constraints is shown in Figure 10.

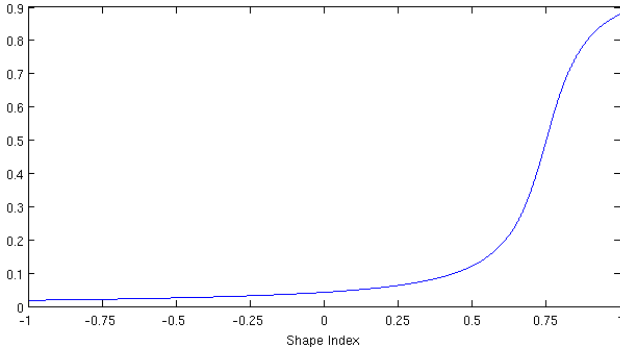


Fig. 10. Function $g(SI) = \frac{1}{\pi} \arctan((SI - 0.75) \cdot 10) + \frac{1}{2}$, multiplying factor for PDE curvature evolution.

The final motion then becomes

$$\frac{\partial \mathcal{S}}{\partial t} = g(SI) \kappa_{min}^{1/4} \vec{N}.$$

This proposed evolution keeps all the advantages of the motion by $\kappa_{min}^{1/4}$ and in addition, the polyps are flattened slower, so at the end the obtained surface is smooth and the polyps are still outstanding.

As discussed in Section I-A, the aim of the work by [40] is to enhance polyps, although they present their work as a preprocessing stage and not a complete CAD system. However, there are several differences with our approach. We use a function of both minimum curvature and shape index, while they use (modified) mean curvature; we use the same evolution for the whole surface, while they use a different function for each polyp candidate; our algorithm smooths the whole surface (needed in order to compute curvatures properly), where their approach is thought to modify the polyp regions only. As our evolution algorithm is part of the segmentation step, while theirs is a stage between segmentation and classification, it is not clear how to quantitatively compare the performances. In any case, as our proposed PDE is part of the segmentation, these two techniques are complementary.

The number of iterations can be set by choosing the value that maximizes the overall performance of the system, measured in terms of the free-response ROC curve (FROC), so the number of iterations is set to obtain the best FROC curve. This can be done by trying with several values and keeping the one which maximize this performance, using cross validation over the training dataset. Alternatively, we can consider a sphere of the size of the CT resolution and compute analytically the number of iterations that are needed to make it vanish (see Appendix A). The idea behind this procedure is to smooth the surface up to the resolution limit. These two approaches led to the same result, namely 15 iterations, and therefore this is the chosen value for the experiments in this paper.

At this point, after choosing the appropriate diffusion and the number of iterations, we have a smoothed volume $u(\mathbf{x}, T)$ indicating the volume inside of the colon. We then extract the surface of the colon, using the marching cubes algorithm [48], obtaining the iso-value surface of level $\alpha \in [0, 1]$. The choice of the value α can be made by maximizing some criteria, in order to obtain the most contrasted surface in a given sense [49]. This optimization-oriented method was tested, and we observed that in our particular application all the consistent surfaces are very close to each other (see Figure 11), and all of them are reasonable segmentations of the colon. Therefore, the computational effort is not justified and we simply kept the iso-level surface $\alpha = 0.7$. Note that this choice can be safely made once for all the data. The result of this stage is then a triangulated surface \mathcal{S} representing the colon wall, Figure 12.

Due to the lack of ground truth, we simply evaluated the segmentation visually and as a component of the polyp detection pipeline (for which we do have ground truth). Visual inspection indicated that the segmentation only missed two very small parts in the entire dataset, and no extra regions (like small bowel) were included. We estimate that only $< 1\%$ of all the database colon surfaces was missed. We have also observed that earlier versions of our segmentation algorithm, e.g., with different velocity functions, do negatively affect the polyps detection. We did not test our algorithm with another bowel preparation due to the lack of such data, and this is an important subject of future research.

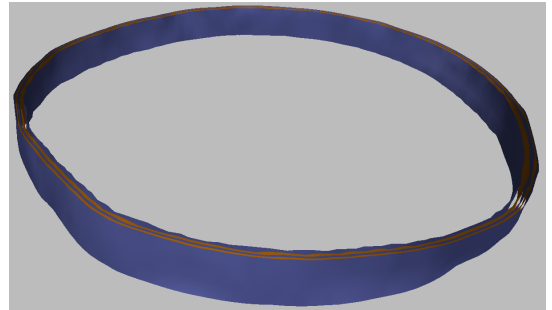


Fig. 11. Different iso-level surfaces (a thin section of the colon wall): α values 0.5, 0.6, 0.7, 0.8, and 0.9. Recall that the whole variation range for α is $(0, 1)$. Note that, although the α values are very different, the obtained iso-level surfaces are very close to each other.

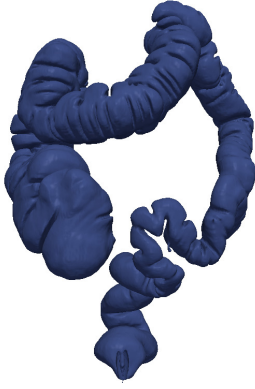


Fig. 12. Example of a segmented colon.

III. POLYP DELINEATION AND FEATURE EXTRACTION

For each case study, we consider as input both the raw CT images and the segmentation of the colon volume as obtained following the procedure described in the previous section. The first stage of the proposed polyps flagging algorithm consists in detecting surface patches that are candidates of being polyps. The complete set of connected points that constitutes the candidate patch is found by progressively growing the patch and keeping the one that maximizes the geometric dissimilarity with the surrounding area, in the sense of the features presented below (the starting point for this growth is also detailed below). The detection system is therefore based on differential (non-absolute) features, thereby better adapting to intrinsic variabilities both of the colon region and of the potential polyps, as further explained next.

All the polyp detection methods reported in the literature try to detect or classify the polyps from properties defined only within the candidate region, without considering the data surrounding the region (of course, the fact that the polyp is raised respect to the surrounding area is always used, but no features measured explicitly taking into account context information were proposed yet). However, it is important to analyze the context in which the candidate patch is located, not only because different sections of the colon present different characteristics, but also because polyps can be situated over different structures such as folds or plain colonic wall. A good feature including the shape of the neighborhood for example, can help in the discrimination between irregular folds and polyps over folds. In addition, looking for significant differences in the gray level imitates the human-based inspection, which highlights zones that contrast with their vicinity.

In this regard, most of the features described in this section take into account the local information of the area surrounding the candidate patch. Polyps (actually all the candidate patches) are then characterized not only by their intrinsic geometry and structure, but also by their relationship with the surrounding area. This makes the features more robust to the particular local phenomena, in a context where the natural variations of the properties of the colon tissue impact the measures and make absolute thresholds or decision rules impractical. The normal tissue of different cases may vary (due to different

biological properties of the subjects or to different conditions of the studies), so absolute thresholds in texture features lack meaning; while texture patterns differ from study to study, what does not vary is the fact that polyps have different properties than normal tissue.

A. Candidate detection and geometrical features

The starting point for the geometric features described in this section is the segmented surface \mathcal{S} . Let us consider the shape index as a function $SI : \mathcal{S} \rightarrow [-1, 1]$, and recall that the polyps have shape index values close to -1 . Therefore, it is expected that a region (patch) of the surface that corresponds to a polyp contains at least one local minimum of the shape index function. The detection of the candidate patches starts from this observation, and follows an adaptive-scale search. For each local minimum $x_0 \in \mathcal{S}$ of the function SI , several level sets of SI ($\mathcal{P}_1 \dots \mathcal{P}_n$) around x_0 are tested, and the level set \mathcal{P}_i that maximizes the distances between the histograms described below, is the considered candidate patch, which we simply denote by \mathcal{P} , Figure 13. A total of $n = 7$ level sets are tested, corresponding to the shape index values from -0.8 to -0.5 with a 0.05 step. The following description is given for the final chosen patch \mathcal{P} , but the ring and histogram computations are made for all the level sets \mathcal{P}_i in order to select the most appropriate of them (the most appropriate scale).

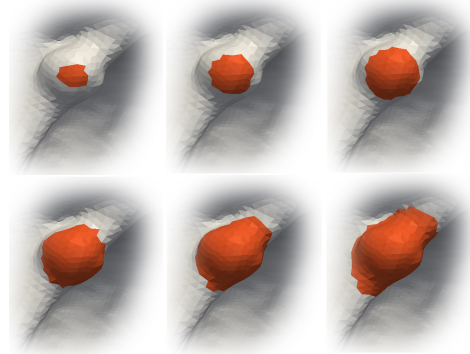


Fig. 13. Sets $\mathcal{P}_1 \dots \mathcal{P}_n$: different sizes are tested in order to select the most appropriate patch. $\mathcal{P}_1 \dots \mathcal{P}_n$.

Given a candidate patch \mathcal{P} , a ring \mathcal{R} around \mathcal{P} is computed, in order to consider geometrical measurements with respect to the area surrounding the patch. The ring is calculated by dilating the patch \mathcal{P} a certain geodesic distance, such that the areas of \mathcal{P} and \mathcal{R} are equal. The geodesic distance computation is performed using the algorithm by [50]. Figure 14 shows a candidate patch (actually a true polyp), and its corresponding ring.

Histograms of the shape index values are then computed for the patch \mathcal{P} and the ring \mathcal{R} , and two different distances between them are computed: the L_1 distance (defined as $\sum_i |x_i - y_i|$, where x_i and y_i are the histogram values) and the symmetric Kullback-Leibler divergence [51]. If the patch corresponds to a polyp-like shape then the values of the histogram \mathcal{P} will be concentrated around the -1 extrema, on the other hand, the histogram \mathcal{R} will be inclined to the other

extreme in case of a polyp on a normal colon wall (concave), or with tendency to values near -0.5 if the polyp is on a fold. These two features give a measure of the geometric local variation of the candidate patch \mathcal{P} . We assume that there are no other polyps in \mathcal{R} or that they do not significantly affect the statistics on the ring.



Fig. 14. Ring (in blue) surrounding a candidate polyp (in orange).

Although these two distances are the most discriminative features⁴, we also consider the following additional ones since they still help to discriminate some typical false positives:

- The mean value of the shape index over the patch \mathcal{P} , which describes the shape of the selected patch.
- The area of the patch, since we want to detect polyps in a certain range of size.
- The growth rate of the areas at the adaptive-size stage, meaning the ratio between the area of the chosen patch $\mathcal{P} = \mathcal{P}_i$ and the area of the immediately smaller tested patch \mathcal{P}_{i-1} ; this feature measures how fast the shape of the patch is changing, in a context where it is difficult to quantize the variation of the shape.
- And finally the *shape factor*,

$$SF = \frac{4\pi \cdot Area}{Perimeter^2} ,$$

which measures the shape of the patch border, how efficiently the perimeter is used in order to gain area,⁵ and it favors circle-like patches (like the polyp patch in Figure 14), avoiding elongated patches (like the false positives in folds).

We then end-up with a total of 6 geometric features for detecting candidate polyps, namely: L_1 and Kullback-Leibler distance between shape index histograms of patches and corresponding rings, the mean shape index over \mathcal{P} , the area of the candidate patch, the growth rate of the areas and the shape factor.

B. Texture features

Due to the differences in biological activity of polyp cells, the gray-level of the CT image and its texture can be very helpful for detecting polyps. This is in particular useful for flat

or small polyps, where the geometric information is limited. Some work has been done on the inclusion of texture features (inside the candidate polyps only), in order to reduce false positives [33]. According to the reported results, there is a lot of room for improvement in texture features. [11] also propose a texture feature, but it relies on structures located at the interface between the polyp and the interior of the colon (tagged fluid or air). It does not make explicit use of the information of the normal tissue surrounding the polyp. We propose both the use of new texture features and the inclusion of the information on the candidate's surrounding area.

First, for each polyp candidate $\mathcal{P} \subset \mathcal{S}$, a volume V_1 is calculated, containing the patch \mathcal{P} and a portion of the inner tissue next to the patch. The volume V_1 is obtained by dilating (in 3D) the patch \mathcal{P} towards the inner colon tissue (we discard the air or fluid voxels). A second volume V_2 surrounding V_1 is calculated by dilating V_1 . Volume V_2 is intended to contain normal tissue in order to compare it with the polyp candidate tissue. In order to choose how much dilation is performed, we use a technique similar to the one in the previous section: several dilation distances are tested, and we keep the distance that makes the differential features most discriminative.

The features chosen are a subset of the classical Haralick texture features [53], namely, entropy, energy, contrast, sumMean, and homogeneity. Seven co-occurrence matrices (considering seven directions in \mathbb{R}^3 , $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(1, 0, 1)$, $(1, 1, 0)$, $(0, 1, 1)$ and $(1, 1, 1)$) are calculated with the voxels of V_1 , and all the five features are averaged over the seven directions. The analogous computation is made for V_2 , and the differences between the two volumes, for each texture feature, is considered. Additionally, the mean gray levels of the voxels in both volumes is computed, and their difference is considered as a feature. In this way, six texture features are considered. The rest of the Haralick features were discarded because the overall performance was not as good as the performance with this particular subset. We have confirmed this choice with several feature selection techniques, with the same procedure described above. This approach for computing the texture features, measuring differences with the surrounding area, leads to better discrimination than the features computed just for V_1 , as demonstrated next.

IV. CLASSIFICATION

After the candidates detection with the adaptive-scale approach, the number of true polyps was much lower than the number of non-polyps patches, a relation on the order of 500:1, which is a significant problem for the learning stage of the classifier. Three techniques were considered to overcome this shortcoming.

The MetaCost approach [54] consists of combining several instances of the classifier instead of stratification (modify the proportion of classes in the training data according to the costs). This method does not work with "stable" classifiers (those that produce similar models with slightly different training sets) like SVM or Naive Bayes.

The Cost Sensitive Learning approach, unlike the MetaCost, tries to balance the classes before the learning stage. The

⁴Several feature selection techniques confirmed this. For instance, the following methods available in Weka [52]: *Information Gain*, *Gain Ratio*, and *Relief-F*. Using the data described below, these techniques sort the features according to their discriminative power.

⁵The maximum value for the shape factor is 1 and it is achieved only by the circle.

implementation we used from *Weka* [52] simply takes as input the re-balance parameters (cost matrix) and replicates instances of the minority class. One of the advantages of this approach is that no assumptions are made about the behavior of the classifiers (unlike the previous method) nor the distribution of the data (unlike the next method).

Finally, the Synthetic Minority Over-sampling TEchnique (SMOTE) is a method to generate artificial instances of the minority class, in order to get a balanced data to learn from. The new artificial instances are created as a convex combination of the existing instances of the minority class. Therefore, there is an underlying assumption that the optimal partition in the feature space gives convex sets, which may not be the case in several applications.

We tested all these options and the best results were obtained using Cost Sensitive Learning, as expected from the comments above.

The numerical results listed below were obtained by classifying with SVM using Cost Sensitive Learning, after normalizing the data; Naïve Bayes performed similarly.

V. RESULTS

A total of 150 patients of the Walter Reed Army Medical Center (WRAMC) database [55] were used to test the proposed CAD algorithm⁶. Most of these patients have two sets of CT images, one for supine and one for prone position. Taking precautions not to train the classifier with, for instance, the prone images set and test it with the supine set (i.e. one cannot use a supine study of a given patient for training, and the corresponding prone study for testing, or vice versa), we can consider the 300 images sets as independent. From now on, we refer to each of these 300 images sets as a *case*. The database contains 134 polyps detected by optical colonoscopy, including 12 flat polyps. Among these 134 polyps, 86 are larger than $6mm$ in size, and the other 48 are between $3mm$ and $6mm$ in size. Figure 15 shows the distribution of polyps' sizes in the database. The size and shape classification of these polyps was taken from the WRAMC database description. These descriptions were provided by the physicians involved in the OC examinations. Taking these examinations as the ground truth, patches classified as polyp were considered TP if the distance to a ground truth polyp was less than $3mm$.

The evaluation was carried out by splitting the dataset into two halves, training and testing. Under this setting, we obtained the FROC in Figure 16, which shows the performance for different polyps sizes. About 40% of the polyps were covered by tagging, but the classification results do not vary depending on this fact, the performance is the same for covered and for non-covered polyps.

These values are comparable with the state-of-the-art results [33], [41], but our database includes very small polyps. A more precise comparison of results is not necessarily meaningful, since in general each work considers its own database.

The FROC curve in Figure 17 compares the performance of the system with the different smoothing methods discussed in

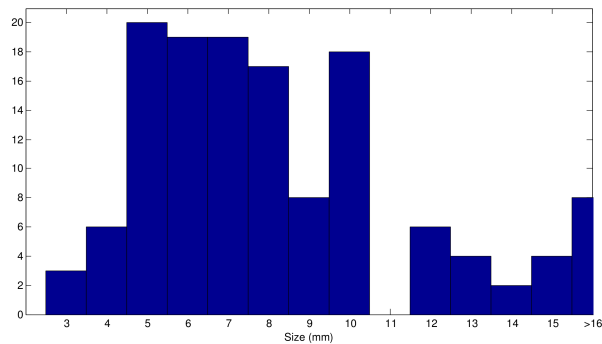


Fig. 15. Histogram of the polyps' sizes.

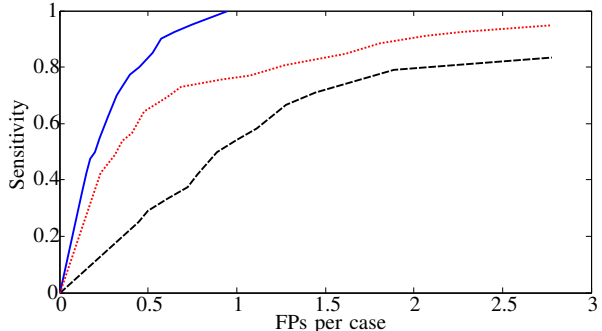


Fig. 16. FROC curve of the proposed system for different polyps sizes: larger than $6mm$ (solid), smaller than $6mm$ (dashed), and all polyps (dotted).

Section II. The proposed smoothing technique achieves better results than the other discussed methods.

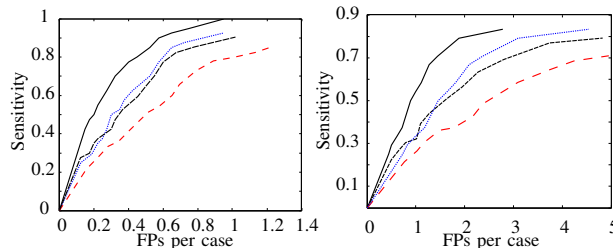


Fig. 17. FROC curve comparing the performances using the different smoothing methods, classifying large polyps (left) and small polyps (right). The curve for the proposed evolution is shown in solid line, the results for the evolution by \mathcal{H} and κ_{min} are shown in dotted and dashed lines respectively, and the lower curve is the result when no smoothing is performed.

	Texture features	
	Absolute	Differential
Sensitivity	96%	100%
FP per case	3.1	2.2

TABLE II

COMPARISON OF ABSOLUTE AND DIFFERENTIAL TEXTURE FEATURES, WITH POLYPS LARGER THAN $3mm$ IN SIZE.

Table II shows the comparison between absolute and differential texture features. The classification was performed using all the geometric features and either the absolute texture features (computed just for V_1), or the differential texture features, using the standard leave-one-patient-out strategy. The results show that, when combined with the differential geometric

⁶Data provided courtesy of Dr. Richard Choi, Virtual Colonoscopy Center, Walter Reed Army Medical Center.

	Polyp sizes	
	> 3mm	> 6mm
Sensitivity	93%	100%
FP per case	2.8	0.9

TABLE III
COMPARISON OF PERFORMANCE BY POLYP SIZE.

features, differential texture features are more discriminative than the absolute ones. The FROC curve in Figure 18 extends the results in Table II and compares the performance of the classifier when using differential or absolute features. Table III compares the classification results according to the polyps size. Again, the work with such small, as well as flat polyps, is unique to the framework here presented, as will be discussed in Section VI-A.

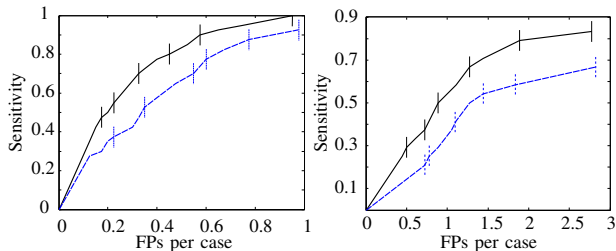


Fig. 18. FROC curve with 95% confidence intervals, comparing the performances with differential (solid) and absolute (dashed) texture features, classifying polyps larger than 6mm in size (left) and smaller than 6mm in size (right). Confidence intervals computed according to [56].

Finally the FROC curve in Figure 19 compares the results of the different classification approaches. Cost Sensitive, SMOTE, and MetaCost were used as a pre-processing stage for SVM, AdaBoost was used with C4.5 trees. Parameters in all classifiers were optimized via cross validation. Other tested classifiers were not included due to their lower performance.

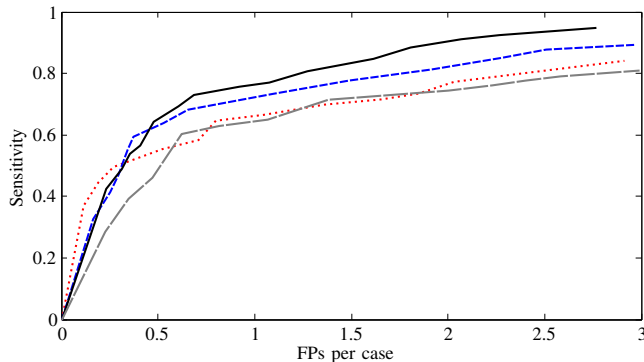


Fig. 19. FROC curve comparing the performances of different classification approaches. SVM with Cost Sensitive (solid), SVM with SMOTE (dashed), C4.5 trees with AdaBoost (dotted) and plain SVM (long-dashed).

VI. DISCUSSION

A. Small, big and flat polyps

It is clear that both the small and flat polyps are much more difficult to detect than the other polyps. What is not clear is if the same kind of algorithm and features are suitable for

detecting all the range of polyp types and sizes. We showed that the proposed combination of features, although it may not be optimal for every specific type of lesion, is able to correctly detect all of them.

All the stages in the pipeline, specially the segmentation and the features, contribute to the good classification results for the whole database. However, it would be interesting to study which pre-processing techniques and features are better for each type and size of polyp, and to possibly propose different CAD systems for each class of polyp. Nevertheless, the 93% sensitivity together with the 2.8 FP rate for polyps larger than 3mm in size is as remarkable as the 0.9 FP rate for polyps beyond 6mm in size with 100% detection.

B. Geometric and texture importance

Although geometrical features are the most discriminative ones (see Table IV), texture features still play a fundamental role in the classification. Adding the texture features to the geometrical ones, the sensitivity reaches 93%, and at the same time the false positives rate decreases by 30%.

	Features			
	All	Geometric	Texture	Excl. L_1 & KL
Sensitivity	93%	88%	68%	83%
FPs p/case	2.8	6.5	19	12.8

TABLE IV
COMPARISON OF PERFORMANCE USING ONLY GEOMETRIC VS ONLY TEXTURE FEATURES, AND EXCLUDING THE HISTOGRAM L_1 AND KL DISTANCES, FOR POLYPS LARGER THAN 3mm.

Figure 20 shows a detected polyp, where the geometry is crucial, because the gray-level does not present considerable local variations. This is specially true for polyps submerged in tagged material. On the other hand, in the flat polyp of Figure 21, the geometry is weakly discriminative (although the measure considering the ring enhances the detectability), and the texture features lead to a correct classification.

Texture information is very important also because it is more robust to segmentation errors, as the texture features are computed by integrating over the volumetric data. Moreover, the differential texture features (the differences between V_1 and V_2) outperform the absolute texture features (just computed in V_1), as shown in Table II.

C. Qualitative analysis of false positives

In addition to the number of false positives, it is very important to study how these FP patches look like, since some of them can be quickly ruled out by the expert and some can be avoided by improving some aspects of the segmentation step.

About half of the false positives are quite reasonable, in the sense that they are (usually small) sections of the colon that are polyp-like shaped, (see Figure 22), specially taking into account that we designed the system to also detect small and flat polyps. Among these false positives, 40% are very small patches that may be avoided by incorporating some new features to the classification, or by adding a size threshold

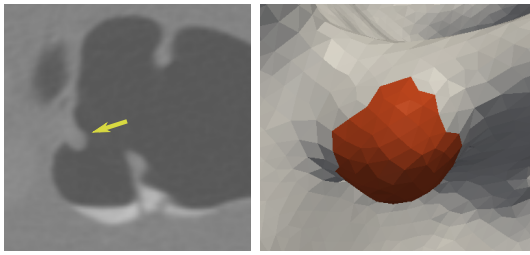


Fig. 20. Polyp with no texture information.

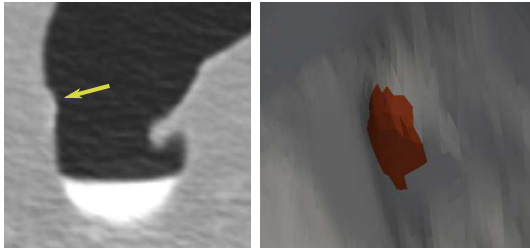


Fig. 21. Polyp with texture information, but weak geometric information.

if very small polyps are not considered of interest (which was not our case). On the other hand, about 10% of the FPs occurred in fold sections of the wall, Figure 22, and another 10% occurred in parts of the insufflation tube. All these patches (from folds and the insufflation tube), are easily ruled out by visual inspection. Another 20% of the FPs were caused by colon segmentation errors, this number having been significantly reduced thanks to the texture features as discussed above. About half of these false positives due to segmentation errors are caused by bad quality original CT slices in the region (generally due to the partial volume effect), like the bottom-left example in Figure 23. Another fraction of FP due to segmentation are protuberances caused by some fluid voxels near to the colon wall, where the gray value is away from standard values. This is the case of the bottom-right example in Figure 23. An additional representative set of false positives is shown in Figure 23.

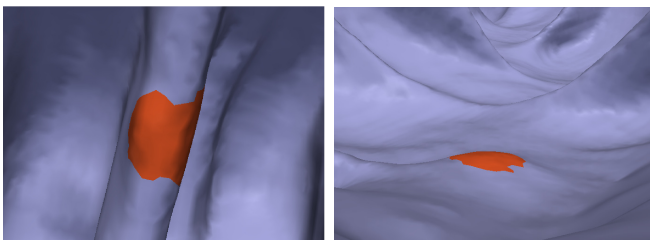


Fig. 22. False positives: fold and patch similar to polyp.

The region in Figure 24 was classified as polyp by our system. However it was not labeled as such in the database, so we have decided to count it as a false positive. We have first included this false positive in the “segmentation error” category, as suggested by visual inspection. Nevertheless, a more careful examination using the original CT slices shows that the structure is indeed a tissue protuberance, and it is not a segmentation artifact. It may be a part of the colon

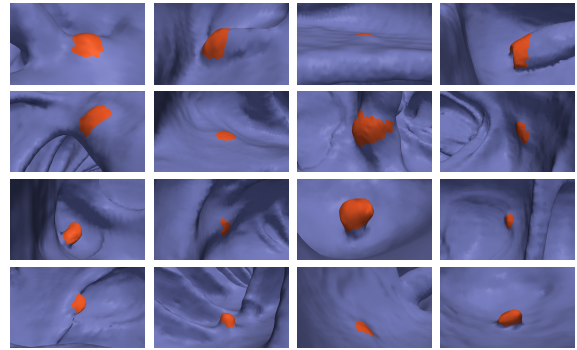


Fig. 23. Examples of false positives according to the available labeled data, with some segmentation errors, parts of the insufflation tube, and some patches with polyp-like shape.

not explored in the OC. Unfortunately we have no additional information that would allow us to clarify whether this region corresponds to a polyp or not.

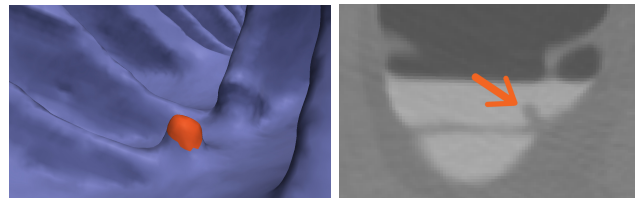


Fig. 24. Visual inspection of the CT slices seems to support that the geometry is well reconstructed by the segmentation procedure. Nevertheless, since this structure was not labeled as a polyp in the ground truth database, here we consider this detection to be a false positive.

D. Contribution of each novelty

In this work we have presented a complete system for polyp flagging, with several novelties. Each one of these novelties, presented in the introduction and detailed along the paper, contributes to different aspects and in different ways to reach the final system performance:

- Artifacts due to the tagged fluid are usually a very important source of false positives. We have minimized these false positives with our region pre-processing approach.
- The smoothing PDE algorithm helps to detect small polyps, and it has proven to outperform the classical curvature motions.
- The *adaptive-scale* candidate search provides a correct delineation of the polyp region, as illustrated in the figures presented along the paper.
- The new *differential features* improve the performance with respect to the features measured on the polyp region only, as shown by the corresponding FROC curves.
- The use of *Cost Sensitive* learning is crucial to overcome the class imbalance problem. It has proven successful, allowing to reach higher performance than other techniques also specifically designed to deal with the class imbalance problem.

VII. CONCLUSION

We introduced a complete pipeline for a Computer Aided Detection algorithm that flags candidate polyp regions. The segmentation stage is very simple and fast, and its main novelty is the smoothing PDE which enhances the polyps, leading to a better detection. In addition to the incorporation of the Haralick texture features, the main yet simple novelties of the proposed features and classification stages are twofold. First, the surrounding area of candidate polyps are explicitly taken into account. Indeed, the proposed (so-called differential) features are computed by comparing properties in the central and surrounding regions of the polyps. We show that differential features are more discriminative than the absolute ones, as they emphasize local deviations of the geometry and texture over the colon. The other novelty is a adaptive-scale strategy that test regions of different sizes and automatically selects the region that best delineates each candidate polyp. The obtained quantitative results are very promising, detecting 100% of the true-polyps, including small and flat ones, with a low false positives rate. Additional improvement of the segmentation and, in collaborations with radiologists, finding features that are tailored to polyp-like geometries, can further improve these results.

ACKNOWLEDGMENTS

Work supported in part by NSF, NGA, ARO, AFOSR and ANII-Uruguay. We thank Sergio Aguirre from Echopixel for important feedback regarding VC.

REFERENCES

- [1] World Health Organization. (2011) Cancer.
- [2] American Cancer Society. (2010) Colorectal cancer facts & figures 2008-2010. [Online]. Available: <http://www.cancer.org/Research/CancerFactsFigures>
- [3] M. Zalis, M. Barish, J. Choi, A. Dachman, H. Fenlon, J. Ferrucci, S. Glick, A. Laghi, M. Macari, E. McFarland, and Others, "CT Colonography reporting and data system: A consensus proposal," *Radiology*, vol. 236, no. 1, p. 3, 2005.
- [4] D. Vining, D. Gelfand, R. Bechtold, E. Scharling, E. Grishaw, and R. Shifrin, "Technical feasibility of colon imaging with helical CT and virtual reality," *American Journal of Roentgenology*, vol. 162, no. Suppl, 1994.
- [5] A. Kaufman and S. Lakare, "Fantastic voyage of the virtual colon," in *Scientific Visualization: The Visual Extraction of Knowledge from Data*, ser. Mathematics and Visualization, 2006, pp. 149–161.
- [6] A. Blachar, M. Graif, A. Kessler, and J. Sosna, "State-of-the-art CT colonography: Update on technique and performance," *Current Colorectal Cancer Reports*, vol. 3, no. 1, pp. 49–54, May 2007.
- [7] D. Kim, P. Pickhardt, A. Taylor, and W. Leung, "CT colonography versus colonoscopy for the detection of advanced neoplasia," *New England Journal of Medicine*, vol. 358, no. 1, p. 89; author reply 90, Jan. 2007.
- [8] P. Sundaram, A. Zomorodian, C. Beaulieu, and S. Napel, "Colon polyp detection using smoothed shape operators: preliminary results," *Medical Image Analysis*, vol. 12, no. 2, pp. 99–119, 2008.
- [9] W. Hong, F. Qiu, and A. Kaufman, "A pipeline for Computer Aided Polyp Detection," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 861–868, 2006.
- [10] V. F. van Ravesteijn, C. van Wijk, F. M. Vos, R. Truyen, J. F. Peters, J. Stoker, and L. J. van Vliet, "Computer-aided detection of polyps in CT colonography using logistic regression," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 120–31, Jan. 2010.
- [11] J. Näppi and H. Yoshida, "Automated detection of polyps with CT colonography: evaluation of volumetric features for reduction of false-positive findings," *Academic Radiology*, vol. 9, no. 4, pp. 386–97, Apr. 2002.
- [12] —, "Fully Automated Three-Dimensional Detection of Polyps in Fecal-Tagging CT Colonography1," *Academic Radiology*, vol. 14, no. 3, pp. 287–300, 2007.
- [13] H. Yoshida, Y. Masutani, P. MacEneaney, D. T. Rubin, and A. H. Dachman, "Computerized detection of colonic polyps at CT colonography on the basis of volumetric features: pilot study," *Radiology*, vol. 222, no. 2, pp. 327–36, Mar. 2002.
- [14] J. L. Fidler, C. D. Johnson, R. L. MacCarty, T. J. Welch, A. K. Hara, and W. S. Harmsen, "Detection of flat lesions in the colon with CT colonography," *Abdom Imaging*, vol. 27, no. 3, pp. 292–300, 2002.
- [15] J. Fidler and C. Johnson, "Flat polyps of the colon: accuracy of detection by CT colonography and histologic significance," *Abdom Imaging*, vol. 34, no. 2, pp. 157–71, 2009.
- [16] R. A. Wolber and D. A. Owen, "Flat adenomas of the colon," *Human Pathology*, vol. 22, no. 1, pp. 70–74, 1991.
- [17] A. J. Aschoff, A. S. Ernst, H.-J. Brambs, and M. S. Juchems, "CT colonography: an update," *European Radiology*, vol. 18, no. 3, pp. 429–37, Mar. 2008.
- [18] R. M. Summers, "CT colonography computer-aided detection: Effect on radiologist observers," in *MICCAI 2010*, 2010.
- [19] J. Church, "Clinical significance of small colorectal polyps," *Diseases of the Colon & Rectum*, vol. 47, pp. 481–485, 2004, 10.1007/s10350-003-0078-6.
- [20] J. H. Bond, "Colorectal cancer screening: the potential role of virtual colonoscopy," *Journal of Gastroenterology*, vol. 37 Suppl 13, pp. 92–96, 2002.
- [21] M. Fiori, P. Musé, S. Aguirre, and G. Sapiro, "Automatic colon polyp flagging via geometric and texture features," in *International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, Jan. 2010, pp. 3170–3173.
- [22] D. Vining, Y. Ge, D. Ahn, and D. Stelts, "Virtual colonoscopy with computer-assisted polyps detection," *Computer-Aided Diagnosis in Medical Imaging*, pp. 445–452, 1999.
- [23] H. Yoshida and J. Näppi, "Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1261–74, 2001.
- [24] R. M. Summers, M. Franaszek, M. T. Miller, P. J. Pickhardt, J. R. Choi, and W. R. Schindler, "Computer-aided detection of polyps on oral contrast-enhanced CT colonography," *American Journal of Roentgenology*, vol. 184, no. 1, pp. 105–8, Jan. 2005.
- [25] D. S. Paik, C. F. Beaulieu, G. D. Rubin, B. Acar, R. B. Jeffrey, J. Yee, J. Dey, and S. Napel, "Surface normal overlap: a computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT," *IEEE Transactions on Medical Imaging*, vol. 23, no. 6, pp. 661–75, 2004.
- [26] D. Chen, Z. Liang, M. R. Wax, L. Li, B. Li, and A. Kaufman, "A novel approach to extract colon lumen from CT images for virtual colonoscopy," *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1220–6, Dec. 2000.
- [27] M. Franaszek, R. M. Summers, P. J. Pickhardt, and J. R. Choi, "Hybrid segmentation of colon filled with air and opacified fluid for CT colonography," *IEEE Transactions on Medical Imaging*, vol. 25, no. 3, pp. 358–68, Mar. 2006.
- [28] D. Chen, R. Fahmi, A. A. Farag, R. L. Falk, and G. W. Dryden, "Accurate and fast 3D colon segmentation in CT colonography," in *ISBI 2009*, 2009, pp. 490–493.
- [29] Z. Wang, Z. Liang, X. Li, L. Li, B. Li, D. Eremina, and H. Lu, "An improved electronic colon cleansing method for detection of colonic polyps by virtual colonoscopy," *IEEE Transactions on Bio-medical Engineering*, vol. 53, no. 8, pp. 1635–46, Aug. 2006.
- [30] S. Yamamoto, G. Iinuma, M. Suzuki, T. Tanaka, Y. Muramatsu, and N. Moriyama, "A simple image processing approach for electronic cleansing in computed tomographic colonography," *Biomed Imaging Interv J*, vol. 5, no. 3, p. e28, 2009.
- [31] I. W. O. Serlie, F. M. Vos, R. Truyen, F. H. Post, J. Stoker, and L. J. van Vliet, "Electronic Cleansing for Computed Tomography (CT) Colonography Using a Scale-Invariant Three-Material Model," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1306–1317, Jun. 2010.
- [32] R. M. Summers, J. Yao, P. J. Pickhardt, M. Franaszek, I. Bitter, D. Brickman, V. Krishna, and J. R. Choi, "Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population," *Gastroenterology*, vol. 129, no. 6, pp. 1832–44, 2005.
- [33] Z. Wang, Z. Liang, L. Li, X. Li, B. Li, J. Anderson, and D. Harrington, "Reduction of false positives by internal features for polyp detection in CT-based virtual colonoscopy," *Medical Physics*, vol. 32, no. 12, pp. 3602–16, 2005.

- [34] C. van Wijk, V. F. van Ravesteijn, F. M. Vos, and L. J. van Vliet, "Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 688–698, 2010.
- [35] J. L. Ong and A.-K. Seghouane, "From point to local neighborhood: polyp detection in CT colonography using geodesic ring neighborhoods," *IEEE Transactions on Image Processing*, vol. 20, no. 4, pp. 1000–10, Apr. 2011.
- [36] E. Konukoglu and B. Acar, "HDF: Heat diffusion fields for polyp detection in CT colonography," *Signal Process.*, vol. 87, no. 10, pp. 2407–2416, Oct. 2007.
- [37] L. Bogoni, P. Cathier, M. Dundar, A. Jerebko, S. Lakare, J. Liang, S. Periaswamy, M. E. Baker, and M. Macari, "Computer-aided detection (CAD) for CT colonography: A tool to address a growing need," *British Journal of Radiology*, vol. 78, no. suppl_1, pp. S57–62, 2005.
- [38] S. A. Taylor, S. Halligan, D. Burling, M. E. Roddie, L. Honeyfield, J. McQuillan, H. Amin, and J. Dehmeshki, "Computer-assisted reader software versus expert reviewers for polyp detection on CT colonography," *American Journal of Roentgenology*, vol. 186, no. 3, pp. 696–702, 2006.
- [39] S. B. Gökürk, C. Tomasi, B. Acar, C. F. Beaulieu, D. S. Paik, R. B. Jeffrey, J. Yee, and S. Napel, "A statistical 3-D pattern processing method for computer-aided detection of polyps in CT colonography," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1251–60, 2001.
- [40] E. Konukoglu, B. Acar, D. S. Paik, C. F. Beaulieu, J. Rosenberg, and S. Napel, "Polyp enhancing level set evolution of colon wall: method and pilot study," *IEEE Transactions on Medical Imaging*, vol. 26, no. 12, pp. 1649–56, Dec. 2007.
- [41] K. Suzuki, H. Yoshida, J. Näppi, S. G. Armato, and A. H. Dachman, "Mixture of expert 3D massive-training ANNs for reduction of multiple types of false positives in CAD for detection of polyps in CT colonography," *Medical Physics*, vol. 35, no. 2, p. 694, 2008.
- [42] M. P. Wand and M. C. Jones, *Kernel Smoothing (Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC, Dec. 1994.
- [43] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.
- [44] D. Chopp and J. Sethian, "Flow under curvature: singularity formation, minimal surfaces, and geodesics," *Experimental Mathematics*, vol. 2, no. 4, pp. 235–255, 1993.
- [45] V. Caselles and C. Sbert, "What is the best causal scale space for three-dimensional images?" *SIAM Journal on Applied Mathematics*, vol. 56, no. 4, pp. 1199–1246, 1996.
- [46] J. J. Koenderink, *Solid Shape*. Cambridge, USA: MIT Press, 1990.
- [47] O. Monga, S. Benayoun, and O. Faugeras, "From partial derivatives of 3-D density images to ridge lines," in *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1992, pp. 354–359.
- [48] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *SIGGRAPH '87: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, 1987, pp. 163–169.
- [49] E. Meinhardt, E. Zacur, A. F. Frangi, and V. Caselles, "3D edge detection by selection of level surface patches," *Journal of Mathematical Imaging and Vision*, vol. 34, no. 1, pp. 1–16, Oct. 2008.
- [50] R. Kimmel and J. Sethian, "Computing geodesic paths on manifolds," in *Proceedings of the National Academy of Sciences*, vol. 95, no. 15, 1998, pp. 8431–8435.
- [51] T. M. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.
- [52] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, "The WEKA data mining software: an update," *Special Interest Group on Knowledge Discovery and Data Mining Explorer Newsletter*, vol. 11, no. 1, pp. 10–18, Nov. 2009.
- [53] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [54] P. Domingos, "MetaCost: A general method for making classifiers cost-sensitive," in *International Conference on Knowledge Discovery and Data Mining*, 1999, pp. 155–164.
- [55] P. Pickhardt, J. Choi, I. Hwang, J. Butler, M. Puckett, H. Hildebrandt, R. Wong, P. Nugent, P. Mysliwiec, and W. Schindler, "Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults," *New England Journal of Medicine*, vol. 349, no. 23, pp. 2191–2200, 2003.
- [56] S. A. Macskassy, F. Provost, and S. Rosset, "Roc confidence bands: an empirical evaluation," in *International Conference on Machine Learning*. New York, NY, USA: ACM, 2005, pp. 537–544.

APPENDIX

We want to compute the number of iterations needed to make a certain sphere (of the size of the CT resolution) vanish, according to our proposed *PDE*. For a sphere, the shape index is constant, so the *PDE* becomes:

$$\frac{\partial \mathcal{S}}{\partial t} = \kappa_{min}^{1/4} \vec{\mathcal{N}}.$$

Then, as $k_{min} = 1/r$, the sphere radius satisfies the following differential equation:

$$r'(t) = \frac{-1}{r(t)^{1/4}}.$$

Therefore,

$$\int_0^T r'(t)r(t)^{1/4} dt = - \int_0^T 1 dt \Rightarrow \frac{4}{5} \left(r(T)^{4/5} - r(0)^{4/5} \right) = -T.$$

As we want to find the value of T so that $r(T) = 0$,

$$T = \frac{4}{5} r(0)^{4/5}.$$

The resolution in the z direction is $1mm$ in our examples, and the time step considered for the numerical method was $t_s = 0.055$. That gives a value N for the number of iterations $N = \lceil 14.54 \rceil = 15$.