

Article

# Multistage Segmentation of Prostate Cancer Tissues Using Sample Entropy Texture Analysis

Tariq Ali <sup>1</sup>, Khalid Masood <sup>2</sup>, Muhammad Irfan <sup>3,\*</sup>, Umar Draz <sup>4,\*</sup>, Arfan Ali Nagra <sup>2</sup>, Muhammad Asif <sup>2</sup>, Bandar M. Alshehri <sup>5</sup>, Adam Glowacz <sup>6,\*</sup>, Ryszard Tadeusiewicz <sup>7</sup>, Mater H. Mahnashi <sup>8</sup> and Sana Yasin <sup>9</sup>

<sup>1</sup> Department of Computer Science, Sahiwal Campus, COMSATS University Islamabad, Sahiwal 57000, Pakistan; tariqali@cuisahiwal.edu.pk

<sup>2</sup> Department of Computer Science, Lahore Garrison University, Lahore 54792, Pakistan; khalid.masood@lgu.edu.pk (K.M.); arfanalinagra@lgu.edu.pk (A.A.N.); drmuhammasif@lgu.edu.pk (M.A.)

<sup>3</sup> Electrical Engineering Department, College of Engineering, Najran University, Najran 61441, Saudi Arabia

<sup>4</sup> Department of Computer Science, University of Sahiwal, Sahiwal, Punjab 57000, Pakistan

<sup>5</sup> Department of Clinical Laboratory, Faculty of Applied Medical Sciences, Najran University, P.O. Box 1988, Najran 61441, Saudi Arabia; Bmalshehri@nu.edu.sa

<sup>6</sup> Department of Automatic Control and Robotics, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, AGH University of Science and Technology, al. A. Mickiewicza 30, 30-059 Kraków, Poland

<sup>7</sup> Department of Biocybernetics and Biomedical Engineering, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, AGH University of Science and Technology, al. A. Mickiewicza 30, 30-059 Kraków, Poland; rtad@agh.edu.pl

<sup>8</sup> Department of Medicinal Chemistry, Pharmacy School, Najran University, Najran 61441, Saudi Arabia; matermaha@gmail.com

<sup>9</sup> Department of Computer Science, University of Okara, Okara 56130, Pakistan; sanayaseen42@yahoo.com

\* Correspondence: irfan16.uetian@gmail.com (M.I.); sheikhumar520@gmail.com (U.D.); adglow@agh.edu.pl (A.G.)

Received: 10 October 2020; Accepted: 1 December 2020; Published: 4 December 2020



**Abstract:** In this study, a multistage segmentation technique is proposed that identifies cancerous cells in prostate tissue samples. The benign areas of the tissue are distinguished from the cancerous regions using the texture of glands. The texture is modeled based on wavelet packet features along with sample entropy values. In a multistage segmentation process, the mean-shift algorithm is applied on the pre-processed images to perform a coarse segmentation of the tissue. Wavelet packets are employed in the second stage to obtain fine details of the structured shape of glands. Finally, the texture of the gland is modeled by the sample entropy values, which identifies epithelial regions from stroma patches. Although there are three stages of the proposed algorithm, the computation is fast as wavelet packet features and sample entropy values perform robust modeling for the required regions of interest. A comparative analysis with other state-of-the-art texture segmentation techniques is presented and dice ratios are computed for the comparison. It has been observed that our algorithm not only outperforms other techniques, but, by introducing sample entropy features, identification of cancerous regions of tissues is achieved with 90% classification accuracy, which shows the robustness of the proposed algorithm.

**Keywords:** wavelet packets; sample entropy; mean-shift segmentation; dice coefficient

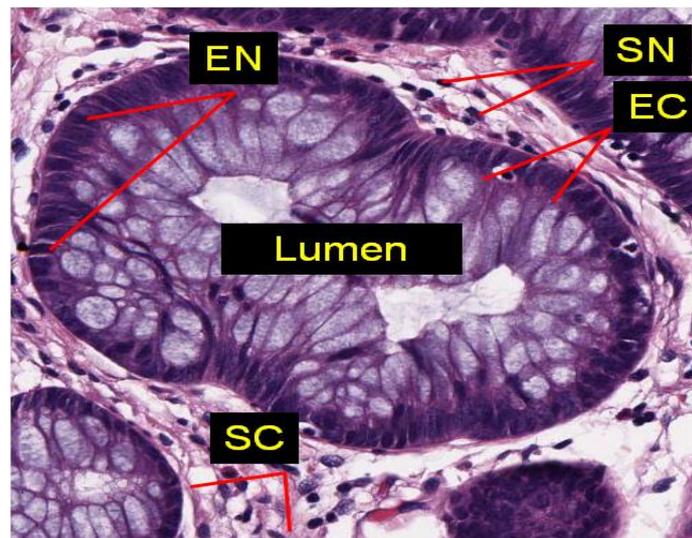
## 1. Introduction

There are around 700,000 deaths from prostate cancer per year worldwide, and it is one of the top five cancers that cause fatality [1,2]. Similar to other cancers, if prostate cancer is diagnosed at earlier stages, the chances of survival for patients are greater than 90% [3]. Experts agree that early diagnosis is key, as demonstrated by their emphasis on regular screening [4]. Screening is often targeted at high risk groups (those known to have premalignant lesions, previously treated cancer, predisposing conditions such as ulcerative colitis, or a family history) who can be followed up by episodic colonoscopy. Screening is aimed at finding polyps, a collection of tiny cells on the prostate surface that are either benign or malignant. Benign cells are not harmful, but malignant cells may develop into cancer in less than five years [5]. It is known that manual screening increases the burden on physicians, and they spend most of their time looking into benign tissues [6]. Furthermore, significant inter- and intra-observer variations in the manual diagnosis of prostate cancer have been reported in the literature [7,8]. Hence, automation in the diagnosis, especially in benign cases, will reduce the burden on pathologists, who will only spend time in examining cancer tissues and can focus more on difficult cases. Automated technologies may not provide complete tumor assessment, but they may assist the pathologists in diagnosis and providing a more accurate estimate for the prognosis [9]. Thus, it is said that histopathology is the profession that has not changed much over the last 100 years [10].

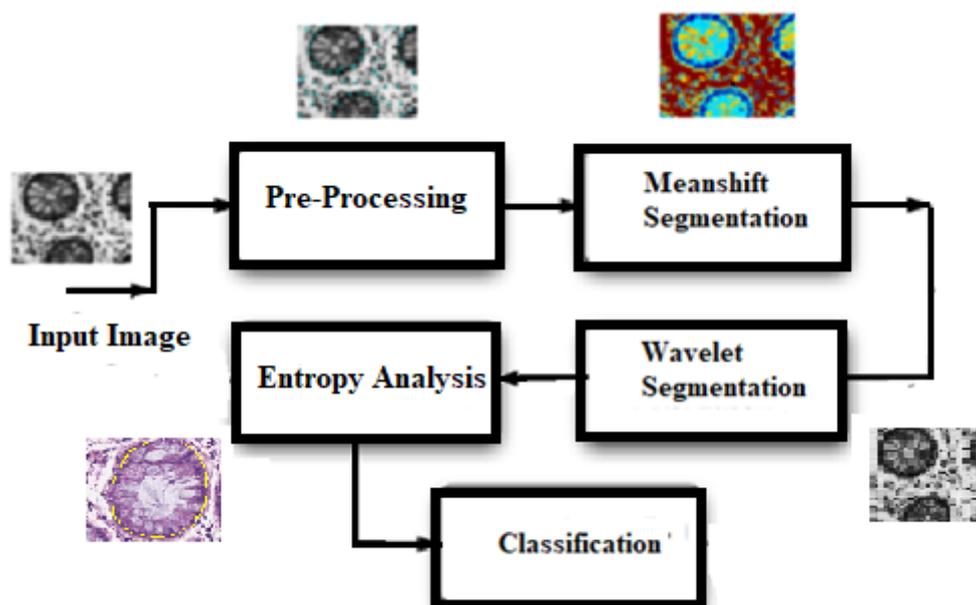
In digital pathology, digital cameras with high resolutions are used to capture the images of pathology slides. A huge amount of data is generated when digitization of glass slides is performed. There are a number of advantages for digitization that include expedite transmission, collective reporting, and forum discussions. A few challenges for digitization include lighting and staining artifacts, processing of a large amount of data and its storage, along with the absence of standards [11]. However, with the advent of advanced technologies and progress in big data and cloud computing, traditional and film-based technology is about to be eradicated and digital histopathology has replaced it completely [12].

In this study, we have developed an automated algorithm using texture based on sample entropy to identify cancerous tissues of prostate cancer. Texture is a feature that is determined by the spatial arrangement of the intensity levels of pixels in a closed neighborhood. It is a pattern that repeats locally for variations of the intensity of an image [13]. In a fine texture, its elements are small and variations are large, whereas in a coarse texture, its elements are large and contain several pixels. Texture can be computed by statistical, structural, or modelling methods. In our proposed technique, statistical methods using entropy and homogeneity are used for classification of benign and cancerous regions of the tissue. Sample entropy (denoted by SampEn) determines the randomness of an image and is proportional to the patterns in an image [14]. As shown in Figure 1, the tissue of a prostate gland is generally tubular in shape and its constituents are epithelial nuclei, stroma nuclei, lumen, and stroma cytoplasm. The presence of epithelial nuclei on the boundary of a prostate gland shows that the tissue structure is intact and it is benign, whereas the spread of epithelial nuclei all over the stroma and the absence of a regular shape of the gland determine that the tissue biopsy is malignant [15].

In our proposed algorithm, a multistage segmentation process for glandular structures of epithelial tissue, choosing prostate tissue as an example, is proposed. There are three stages that work in sequence to model the texture of tissue glands. After doing a few pre-processing steps, such as affine transformation and morphological operations including closing and top hat transforms, a coarse segmentation of the input images is performed using mean-shift segmentation. In the next stage, wavelet packets are used that extract glandular regions of the tissue. The texture of the extracted regions is modeled using sample entropy values and, applying a non-linear threshold, cancerous tissues are differentiated from benign tissues. The block diagram of the proposed system is presented in Figure 2.



**Figure 1.** A microscopic tissue component showing various gland regions such as EN (epithelial nuclei), SN (stroma nuclei), EC (epithelial cytoplasm), SC (stroma Cytoplasm) and lumen.



**Figure 2.** The proposed framework.

The organization of the paper is presented as follows. Related work and background for the proposed algorithm is introduced in the next section. Section 3 presents the multistage segmentation methods of mean-shift and wavelet packets. The algorithmic framework employing the sample entropy is described in Section 4. The experimental results for segmentation of glandular structures in an experimental dataset of prostate biopsy histology images are presented and discussed in Section 5. Conclusions for the proposed model and future directions are presented in the final section.

## 2. Literature Review

A variety of algorithms have been proposed in the literature for the analysis of colon and prostate histology images (see, for example, [16,17]), but relatively few methods, such as [18–20], are aimed at the modeling or segmentation of glandular structures. These methods can be categorized into one of the three classes: (a) structural or texture-based segmentation, (b) boundary segmentation, and (c) region-based segmentation. Farjam et al. [21] employed variance and Gaussian filters to extract textural

features from glandular structures. The authors then extract features related to size and shape of these glands and combine them to generate a malignancy index for classifying a given image into benign or malignant. However, as shown in the results, techniques that employ texture lack the prior knowledge and their performance may degrade.

In boundary segmentation methods, Naik et al. [22] have used a level set method for active contours [23] to segment prostate tissues. The proposed technique uses a level set curve for the segmentation and employs a Bayesian classifier for the detection. Morphological features are then computed from the glandular regions and the proposed machine learning technique assigns a Gleason score to the given image. A limitation of the level set framework, however, is that the initialization of level sets may require manual intervention. Furthermore, level sets may generate errors in the case of occluded glands. Monaco et al. [24] employed a region growing algorithm to segment glands in prostate histology images to address these limitations. More recently, Monaco & Madabhushi [25] proposed a weighted version of the maximum posterior marginal estimation, designed to penalize classification errors associated with certain classes more heavily than others, and employed an extended Markov chain Monte Carlo (E-MCMC) method to detect cancerous glands in prostate histology sections.

A few regions-based approaches for glandular segmentation, such as object graphs [26] and iterative region growing [27], can also be found in the literature. The region growing algorithm of Wu et al. [28] perform segmentation for the images of intestines using adaptive thresholds for intensity values of glands. A major drawback with thresholding is that it requires manual selection, which may vary from one sample to another and is proportional to the number of samples. Moreover, the stoppage criterion for the algorithm is determined on the maximum number of iterations and the criterion varies with the number of samples in the dataset. Demir et al. [29] extended the idea of cell graphs [30] to construct object graphs for segmentation of glandular structures in colon histology images. Nuclei and lumen parts of glands are segmented using object texture information. The boundary of glands is determined using a graph-based technique that connects centroids of the nuclei. One of the limitations of this algorithm is that it adopts a semi-automatic approach to label the gland as a Boolean value to be true or false and it can be subjective. More recently, Tosun et al. [31] have proposed the idea of color graphs using color edges that carry different colors depending on the types of nodes they connect.

Nguyen et al. [32] proposed a hybrid approach involving both boundary segmentation and region growing by employing the prior knowledge about the structure of tissues for the segmentation. In the proposed algorithm, the structure of the gland is constructed by joining nuclei and cytoplasm using a technique for region growing. For pixel-level segmentation, they used the nearest neighbor classifier on the  $L^*a^*b^*$  color space and classified each pixel into various classes such as epithelial nuclei, epithelial cytoplasm, lumen, and stroma. The authors also employed glandular morphological features for the purposes of Gleason grading and reported an accuracy of 88% for classifying input images into benign, grade 3, and grade 4 prostate cancer. The algorithm, however, seems to rely heavily on size constraints where there are holes in the rough glandular boundary extracted by joining nuclei and cytoplasm. Because of the separation of boundary extraction and region growing, the algorithm does not take into account the very intimate spatial relationship between the epithelial boundary of glandular structures and lumen texture of the inside of glandular regions.

In the work of [33], a stack of five convolution neural network (CNN) layers is used to determine prostate cancer. An automatic process was adopted to generate region of interests (ROIs), and this is both the advantage and limitation of the proposed algorithm. There is no ground truth, so the validation lacks the input from an expert. The algorithm performs superiorly with an 84% area under the curve. In [34], the transfer learning approach using deep learning was used to classify prostate cancer. The deep learning approach using CNN was compared with various machine learning algorithm such as decision tree, support vector machine (SVM), and Naive Bayes, and it was found that the CNN outperforms all the machine learning techniques. The limitation of the approach is that transfer learning is employed, which normally suffers from overfitting.

### 3. Materials and Methods

Modeling of the glandular structures in the proposed algorithm, as shown in Figure 2, is described briefly as follows:

- (1) As a first step, segmentation of the tissue is achieved by applying the color mean-shift (MS) algorithm [34] on the input image  $I(z)$ . The tissue is divided into four parts that contain epithelial nuclei, cytoplasm, lumen, and stroma nuclei.
- (2) The above segmentation result can be used to get a rough idea of pixels forming nuclei, both stromal and epithelial. In the second step, textural features for pixels of the image are used to model the glands. The wavelet packet features [35] and AdaBoost classifier are used in assigning lumen-ness labels  $T(z)$  (1 for lumen, and 0 for non-lumen) for all pixel locations  $z \in Z$  in the image  $I(z)$ .
- (3) Finally, the texture of the image is modeled using sample entropy analysis. Using empirical methods, a threshold is obtained that distinguishes cancerous regions from the benign tissue samples.

Our hypothesis is that glandular structures are formed by epithelial nuclei at the boundary of glands and the lumen is inside them. We employed the color mean-shift (MS) algorithm [36] for a coarse segmentation of prostate tissue into its four constituents that contain epithelial/stromal nuclei, cytoplasm, lumen, and stromal background. The color MS algorithm is an unsupervised clustering algorithm that assigns each pixel the weighted mean of its neighboring pixels, where the weight assigned to each of the neighboring pixels is given by a Gaussian kernel, which determines the photometric similarity between a pixel  $z$  under consideration and all pixels in its neighborhood  $N(z)$ . Iterative application of the color-based MS algorithm results in a coarse segmentation of the tissue types, as shown in Figures 3 and 4. The MS algorithm performs an initial segmentation, followed by the fine segmentation using wavelet packets.

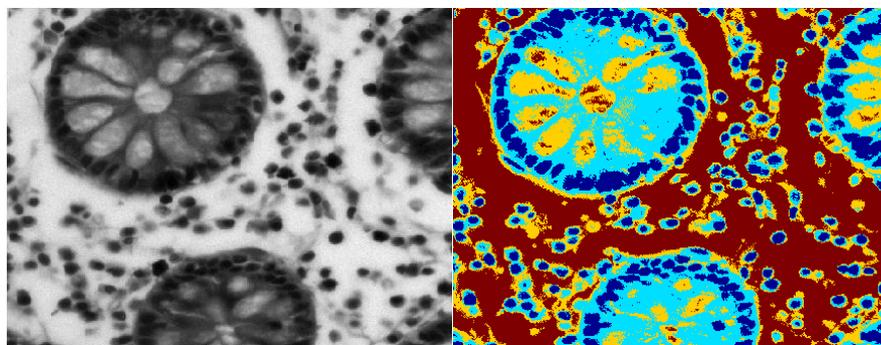


Figure 3. Mean shift segmentation for benign circular glands.

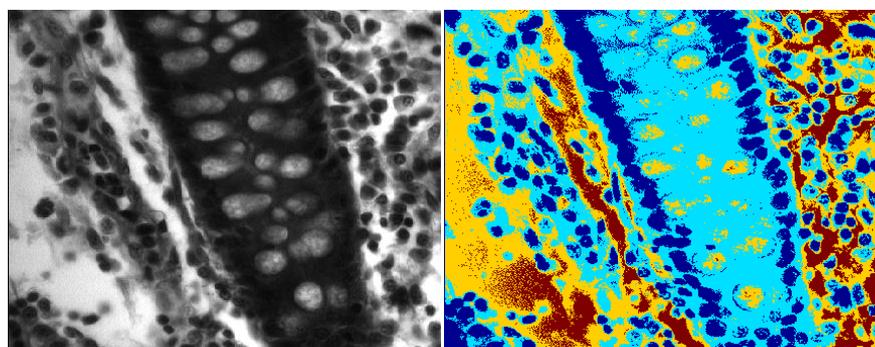


Figure 4. Mean shift segmentation for benign elliptical glands.

In wavelet packets, both low and high frequency components are decomposed in each sub-band. An analogy of binary trees is presented for wavelet packets where every low or high sub-band

represents a node with two children. As the binary tree may have either 0 or 2 children, the basis of wavelet packets is  $2^{2n}$ , where  $n$  is the level of decomposition. In the wavelet packet analysis, the image is examined at different levels. The over-complete wavelet packet decomposition provides a feature vector that is computed from the energy and entropy features of the texture that is obtained by wavelet packets and sample entropy analysis. In wavelet packet decomposition, the approximate and detail images at each level are decomposed into four levels and a quad tree image filter is achieved. In conventional wavelet packets transform, the down-sampled sub images are one-quarter of the size of the original image. However, in over-complete filter decomposition, there is no down-sampling and the sub-bands are of the same size as the original. Thus, a robust feature vector is obtained. A high number of features can be obtained with an increase in the decomposition level, as the number of sub-bands increases exponentially. In two-level decomposition, eight feature sub images are retained only and these images contain maximum variance in each level.

In sample entropy, the spatial distributions of pixels are correlated to compute the entropy distributions [37]. The entropy is an irregularity measure of the texture in images. It determines the similarity of two patterns  $k$  and  $k + 1$ , if their pixel distribution is alike. The similarity is based on the distance ( $d$ ) of each corresponding pixel in both of the patterns. The probability is computed based on the pattern matches in  $k$  and  $k + 1$  pattern. The probability of  $k + 1$  patterns of the same distribution is determined by the ratio of the similarity match in  $k$  patterns. In computing the sample entropy, each  $k$  length pattern is tested for all  $k$  length patterns in the entire image. A match is recorded if each pixel on  $k + 1$ th pattern differs by not more than distance ( $d$ ) from the corresponding pixels in  $k$  window (pattern). Thus, sample entropy reflects the repeated-ness of patterns in images. The low values of entropy show more regular texture and fine details, whereas high values of entropy reflect irregular texture with coarse representation.

Sample entropy is a measure of irregularity and periodic or regular patterns show small values of sample entropy. Let us consider two patterns  $P^m(d)$  and  $P^{m+1}(d)$ ; by determining their logarithmic ratio, the sample entropy can be computed. An image  $I(i, j)$  with  $n_i$  rows and  $n_j$  columns is considered that contains  $x_k(i, j)$  set of pixels varying from  $j$  to  $j + m - 1$  column lines and  $i$  to  $i - m + 1$  row. The window size is  $x_k(i, j)$  and it is centered at  $i$  and  $j$ . There are only  $(n_i - k) \times (n_j - k)$  patterns as the last  $m$  lines cannot be used for square patterns as the total number of pixels is equal to  $N$ , where  $N = n_i \times n_j$ . To determine the mathematical formulation of the sample entropy, the following derivations are used.

$$SEn = -\ln \frac{P^{m+1}(r)}{P^m(r)}$$

where

$$P^m(r) = \frac{1}{N_p} \sum_{i=1}^{N_p} P_i^m$$

$$P^{m+1}(r) = \frac{1}{N_p} \sum_{i=1}^{N_p} P_i^{m+1}$$

where  $r$  represents the number of grey levels and  $N_p$  is the total number of  $m$  length different squared patterns in an image, whereas  $P^m$  and  $P^{m+1}$  are  $m$  and  $m + 1$  length pattern in the image, respectively. The sample entropy is the logarithmic ratio for these patterns. The entropy and energy are computed for each segmented sub image with the square local area obtained by sample entropy window size. Energy is the content of information in an image. It is defined on the normalized histogram of an image and it shows how grey levels are spread in the image. The energy of an image is inversely proportional to Shannon entropy and is closely related to randomness. A high energy image has a low number of grey levels. For each squared area, the feature vector represents values for the central pixel of the sub image.

#### 4. Results

We use texture as a cue for assigning lumen-ness labels to all pixels in the image. The wavelet packet transform is used to compute pixel-based features. In the wavelet transform, the approximation filter is not divided and only the last level is split for the high levels of approximation and detailed images. The idea behind the wavelet packet transform is to further decompose both approximation as well as detail sub-bands at each but the last level. The advantage is that useful information in the high signal wavelengths is computed for the richer analysis of textural features. The wavelet packet features perform fine segmentation of the tissue. AdaBoost is added in the model that enhances the performance of the classifier. In AdaBoost modeling, weak classifiers are converted into a strong classifier.

Daubechies 4 filters of wavelet packets are used to extract a 64-dimensional feature vector. Pixel labeling algorithm using wavelet packet features is trained on 50 random samples in the training data. There were 25 prostate cancer (suspected) patients and 2 samples were collected from each patient. We trained and tested the proposed technique on a dataset of 50 Hematoxylin and Eosin (H&E) stained prostate histopathology images. The dataset is collected from National Cancer Institute (NCI), where regions of interest (ROIs) from whole slide images are extracted at 40X magnification [38]. The resolution of each image is  $2048 \times 2048$ . For the ground truth, manual segmentation was performed by the pathologist who assisted in extracting the ROIs as well. There was one expert who marked all the ground truths (GTs). However, the digitization of the ground truth samples was sent to our collaborator in the other institute, who endorsed most of the GT samples. The difference in opinion was for only two samples, which was later approved and accepted by the expert and the GTs for the two samples were updated accordingly. The Dice similarity coefficient was used to evaluate the performance of our proposed algorithm that computes the area of overlap for automated segmentation with the ground truth images.

An average patch size of  $50 \times 50$  pixels was selected such that half of them were from lumen areas and the other half were from the stromal areas for each of the training images. The wavelet packet transform for the third level generates features for the energy and entropy of each patch and a 64-dimensional feature vector is produced. The Adaboost classifier is used to enhance the accuracy of SVM. The Adaboost technique enhances the performance of a weak classifier, which has accuracy of slightly more than guessing, into a strong classifier. In each iteration of the Adaboost algorithm, the weight of the misclassified observation is increased, otherwise the weight is reduced. In a few iterations, the training error is reduced exponentially, and the algorithm achieves rapid convergence. The trained classifier is then used to assign pixel labels to each image pixel in our dataset based on their 64-dimensional wavelet packet features.

#### 5. Discussion

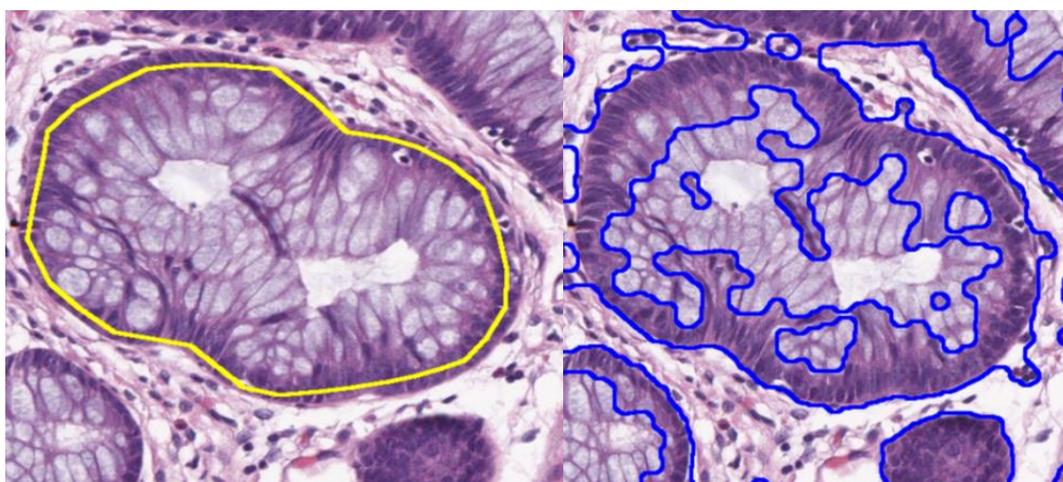
To test the quality of the segmentation results by our algorithm, ground truth glandular regions are computed for each gland marked with its boundary. In Figure 5, the segmented image of the proposed algorithm (TPA) is compared with the Gaussian filtering technique (GFT) in [39]. A major strength of our algorithm is that it does not result in holes, formed during segmentation of the lumen when only texture is used as a cue, because of its use of prior knowledge about gland formation and sequentially connecting epithelial nuclei on the glandular boundary. The classification criterion is based on the sample entropy and wavelet packet features for a patch. The low values of the sample entropy represent a uniform structure and, by computing an empirical threshold, a tissue is considered to be normal if the values for a patch are less than the threshold, and it can be in various stages of cancer if the patch values are above the determined threshold.

To evaluate the performance of our algorithm, the Sorensen–Dice index or Dice Similarity Coefficient (DSC) is used as a quality measure with its value ranging between 0 and 1 [40]. It is a similarity index between two sets of samples. The Dice similarity coefficient (DSC) of zero shows the least agreement between automated segmentation and the ground truth, while a value approaching one represents good agreement between the two. The DSC is computed as 2 times the area of overlap

divided by the total number of pixels in both images as in our case, one is the automated segmented image and the other is the ground truth image. The DSC is the same as the F1 score, which is computed as follows:

$$DICE = (2 \times TP)/(TP + TP + FP + FN)$$

where  $TP$  represents total number of positive samples that are also positive in the ground truth, whereas  $FP$  is the false positive cases, which are false in the ground truth, but predicted to be true by the algorithm.  $FN$  represent the false negatives that the algorithm predicts to be negative, while in the ground truth, they are positive. Figure 6 presents a comparison of the segmentation accuracy of the two algorithms, GFT of [39] and the proposed algorithm (TPA) using DSC. As shown in Figure 6, the proposed model always produces better results than the Gaussian filtering-based segmentation, with average DSC coefficient values being 0.79 and 0.91 for [39] and our method, respectively.



**Figure 5.** Comparison of the Gland segmentation by the proposed algorithm (yellow boundary line) and Gaussian filtering algorithm (GFT) by [39] (blue boundary lines).

To evaluate the experimental results, we have computed quality metrics such as accuracy, precision, and recall. They are defined as follows:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$Precision = \frac{TP}{(TP + FP)}$$

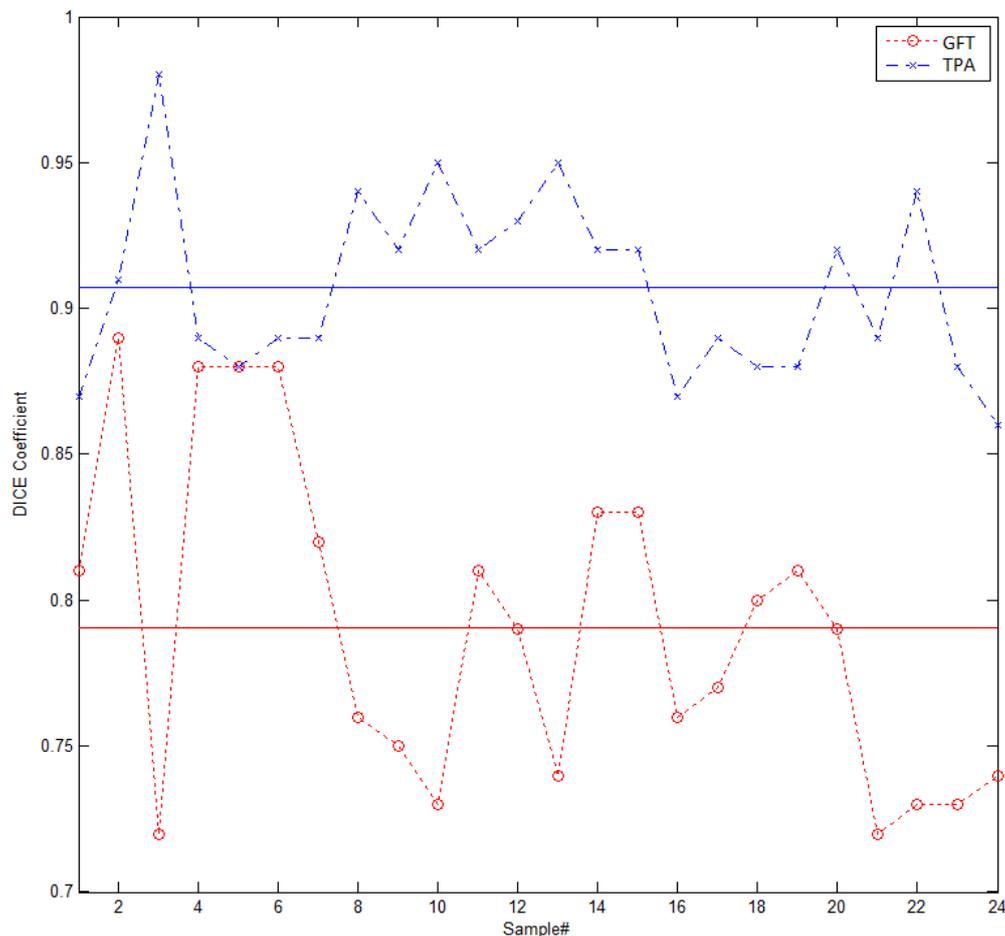
$$Recall = \frac{TP}{(TP + FN)}$$

The parameters in the confusion matrix are explained as follows.

- True-positive ( $TP$ ): The algorithm detects it as cancerous and the ground truth also labels it as cancerous.
- False-positive ( $FP$ ): The algorithm detects it as cancerous, while the ground truth labels it as normal.
- True-negative ( $TN$ ): The algorithm detects it as normal, whereas ground truth also labels it as normal.
- False-negative ( $FN$ ): The algorithm detects it normal, whereas ground truth labels it as cancerous.

Table 1 presents the classification attributes of samples in our experimental dataset. The texture of tissues is modeled using sample entropy values. The support vector machine (SVM) classifier using a Gaussian kernel is employed to separate the two classes of data samples. The model achieves 90%

classification accuracy, which supports the notion that sample entropy values contain discriminative features that separate benign samples from malignant samples. However, a large-scale validation of these results is required for endorsing the results achieved by the proposed algorithm.



**Figure 6.** Dice similarity coefficient (DSC) for the proposed algorithm (TPA) and Gaussian filtering technique (GFT) in [39] for the first 24 samples in our dataset. The average DSC for our technique is 0.91, whereas for GFT, it is 0.79.

**Table 1.** Quality metrics for classification of tissue regions into benign and malignant samples. SVM, support vector machine.

Quality Metrics	Benign	Malignant
	SVM (%)	SVM (%)
Accuracy	90	91
Precision	86	89
Recall	81	84
F1 Score	85	87

## 6. Conclusions

In this study, our proposed algorithm identifies cancerous regions of prostate histology images using the multi-segmentation technique. The model performs two-step segmentation for prostate tissue images. Initially, the MS (mean-shift) algorithm is used to perform the coarse segmentation to split the tissue constituents in four parts. In the second stage, wavelet filters are used to perform fine segmentation of the tissue glands. For classification, sample entropy features are extracted from the segmented regions along with the energy parameters. It has been observed that sample entropy achieves more than 90% classification accuracy. An added advantage of performing multistage

segmentation using sample entropy values is that one could easily separate epithelial nuclei from the stroma nuclei in standard H&E stained images without using any additional immunohistochemical (IHC) markers. Furthermore, morphological signatures of glandular structures modeled using our algorithm can also be fused with disease signatures from other modalities such as mass spectrometry (MS), IHC, flow cytometry (FC), and/or cytogenetic (CG) testing for integrated diagnostic/prognostic purposes [41,42]. Future directions of this work include extension of this model for application to whole-slide histology images of other epithelial cancers such as colon and lung. Moreover, additional data samples will be added for large-scale validation of the model. For robust and efficient utilization of the proposed model, autoencoders with zero shot learning that do not require large training samples will be incorporated into the classification algorithm.

**Author Contributions:** T.A. and K.M. proposed the research conceptualization and methodology. The technical and theoretical framework was prepared by M.I. and U.D. The technical review and improvement were performed by A.A.N. and M.A. The overall technical support, guidance, and project administration were done by B.M.A. and A.G. S.Y. provided her support in drafting this article. The editing and, finally, proofreading were done by R.T. and M.H.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the AGH University of Science and Technology, grant No. 16.16.120.773.

**Acknowledgments:** The authors acknowledge the support from Deanship of Scientific Research, Najran University, Kingdom of Saudi Arabia.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Parkin, D.M.; Bray, F.; Ferlay, J.; Pisani, P. Global cancer statistics, 2002. *CA Cancer J. Clin.* **2019**, *55*, 74–108. [[CrossRef](#)] [[PubMed](#)]
- WHO. *Cancer: Fact Sheet No. 29*; WHO: Geneva, Switzerland, 2019.
- National Cancer Intelligence Network. *Colorectal Cancer Survival by Stage*; NCIN Data Briefing; National Cancer Intelligence Network: London, UK, 2019.
- Smith, R.A.; Cokkinides, V.; Brawley, O.W. Cancer screening in the united states, 2018: A review of current American cancer society guidelines and issues in cancer screening. *CA Cancer J. Clin.* **2018**, *59*, 27–41. [[CrossRef](#)]
- Murgia, C.; Pritchard, J.; Kim, Y. Colon origin and evolution of a transmissible cancer. *Cell* **2016**, *126*, 477–487. [[CrossRef](#)] [[PubMed](#)]
- Gurcan, M.; Boucheron, L.; Can, A.; Madabhushi, A.; Rajpoot, N.; Yener, B. Histopathological image analysis: A review. *IEEE Rev. Biomed. Eng.* **2009**, *2*, 147–171. [[CrossRef](#)] [[PubMed](#)]
- Ali, G.; Ali, A.; Ali, F.; Draz, U.; Majeed, F.; Yasin, S.; Ali THaider, N. Artificial neural network based ensemble approach for multicultural facial expressions analysis. *IEEE Access* **2020**, *8*, 134950–134963. [[CrossRef](#)]
- Van Putten, P.G.; Hol, L.; van Dekken, H.; van Krieken, J.H.; van Ballegooijen, M.; Kuipers, E.J.; van Leerdam, M.E. Inter-observer variation in the histological diagnosis of polyps in colorectal cancer screening. *Histopathology* **2011**, *58*, 974–981. [[CrossRef](#)]
- Masood, K.; Alghamdi, M. Modelling mental stress using a deep learning framework. *IEEE Access* **2019**, *7*, 68446–68454. [[CrossRef](#)]
- May, M. A better lens on disease. *Sci. Am.* **2010**, *302*, 74–77. [[CrossRef](#)]
- Masood, K. Adaptive wavelet packets for minimization of error in computation of mental stress. *Int. J. Wavelets Multiresolution Inf. Proc. (IJWMIP)* **2015**, *13*, 82–90. [[CrossRef](#)]
- Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [[CrossRef](#)]
- Grassberger, P.; Procaccia, I. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev.* **1983**, *2*, 25–91. [[CrossRef](#)]
- Doyle, S.; Feldman, M.; Tomaszewski, J.; Madabhushi, A. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Trans. Biomed. Eng.* **2010**, *59*, 1205–1218. [[CrossRef](#)]

15. Rajpoot, N. Local discriminant wavelet packet basis for texture classification. In Proceedings of the SPIE Wavelets X, San Diego, CA, USA, 13 November 2003; Volume 5207, pp. 774–783.
16. Teverovskiy, M.; Kumar, V.; Ma, J.; Kotsianti, A.; Verbel, D.; Tabesh, A.; Pang, H.Y.; Vengrenyuk, Y.; Fogarasi, S.; Saidi, O. Improved prediction of prostate cancer recurrence based on an automated tissue image analysis system. In Proceedings of the IEEE International Symposium on Nano to Micro, Arlington, VA, USA, 18 April 2004; pp. 257–260.
17. Tabesh, A.; Teverovskiy, M.; Pang, H.Y.; Kumar, V.P.; Verbel, D.; Kotsianti, A.; Saidi, O. Multi-feature prostate cancer diagnosis and Gleason grading of histological images. *IEEE Trans. Med. Imaging* **2007**, *26*, 1366–1378. [[CrossRef](#)]
18. Huang, P.; Lee, C. Automatic classification for pathological prostate images based on fractal analysis. *IEEE Trans. Med. Imaging* **2010**, *28*, 1037–1050. [[CrossRef](#)] [[PubMed](#)]
19. Moradi, M.; Mousavi, P.; Tomaszewski, J.; Feldman, M.; Hagemann, I.; Adolmaesumi, P.; Madabhushi, A. High-throughput detection of prostate cancer in histological sections using probabilistic pairwise Markov models. *Med. Image Anal.* **2010**, *14*, 617–629.
20. Glotsos, D.; Kalatzis, I.; Spyridonos, P.; Kostopoulos, S.; Daskalakis, A.; Athanasiadis, E.; Ravazoula, P.; Nikiforidis, G.; Cavouras, D. Improving accuracy in astrocytomas grading by integrating a robust least squares mapping driven support vector machine classifier into a two level grade classification scheme. *Comput. Methods Prog. Biomed.* **2008**, *90*, 251–261. [[CrossRef](#)] [[PubMed](#)]
21. Farjam, R.; Soltanian-Zadeh, H.; Khouzani, K. An image analysis approach for automatic malignancy determination of prostate pathological images. *Clin. Cytom.* **2007**, *72*, 227–240. [[CrossRef](#)]
22. Naik, S.; Doyle, S.; Fieldman, M. Gland segmentation and Gleason grading of prostate histology by integrating low-high-level and domain specific information. In Proceedings of the 2nd Workshop on Microscopic Image Analysis with Applications in Biology, Piscataway, NJ, USA, 2007. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.6547&rep=rep1&type=pdf> (accessed on 2 December 2020).
23. Tariq, A.; Yasin, S.; Draz, U.; Tariq, T.; Javaid, S.; Arshad, M.A. Motif Detection in Cellular Tumor p53 Antigen Protein Sequences by using Bioinformatics Big Data Analytical Techniques. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 330–338.
24. Monaco, J.; Madabhushi, A. Weighted maximum posterior marginal for random fields using an ensemble of conditional densities from multiple Markov chain Monte Carlo simulations. *Med. Imaging IEEE Trans.* **2011**, *30*, 1335–1364. [[CrossRef](#)]
25. Madabhushi, A. Digital pathology image analysis: Opportunities and challenges. *Imaging Med.* **2009**, *1*, 7–10. [[CrossRef](#)]
26. Waqas, M.; Ghamdi, M.A.; Hussain, M.; Adnan, M.; Masood, K. Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges. *Brain Sci.* **2020**, *10*, 118–130.
27. Yasin, S.; Ali, T.; Draz, U.; Jung, L.T.; Arshad, M.A. Formal Analysis of Coherent Non-Redundant Partition-based Motif Detection Algorithm for Data Visual Analytics. *J. Appl. Environ. Biol. Sci.* **2018**, *8*, 23–30.
28. Wu, H.; Xu, R.; Harpaz, N. Segmentation of gland images with iterative region growing. *J. Microsc.* **2005**, *220*, 190–204. [[CrossRef](#)]
29. Demir, C.; Kandemir, M.; Tosun, A. Automatic segmentation of colon glands using object graphs. *Med. Image Anal.* **2009**, *14*, 1–12. [[CrossRef](#)] [[PubMed](#)]
30. Gunduz, C.; Yener, B.; Gultekin, S.H. The cell graphs of cancer. *Bioinformatics* **2004**, *2*, 145–155. [[CrossRef](#)] [[PubMed](#)]
31. Tosun, A.B.; Gunduz-Demir, C. Graph run-length matrices for histopathological image segmentation. *Med. Imaging IEEE Trans.* **2011**, *30*, 721–732. [[CrossRef](#)]
32. Nguyen, K.; Jain, A.K.; Allen, R.L. Automated gland segmentation and classification for Gleason grading of prostate tissue images. In Proceedings of the 20th International Conference on Pattern Recognition (ICPR), Istanbul, Turkey, 23–26 August 2010; Volume 1, pp. 1497–1500.
33. Yoo, S.; Gujrathi, I.; Haider, M.; Khalvati, F. Prostate Cancer Detection using Deep Convolutional Neural Networks. *Sci. Rep.* **2019**, *9*, 19518. [[CrossRef](#)]
34. Abbasi, A.; Hussain, L.; Awan, I.; Abbasi, I.; Majid, A.; Nadeem, M.A.; Chaudhary, Q. Detecting prostate cancer using deep learning convolution neural network with transfer learning approach. *Cogn. Neurodynamics* **2020**, *14*, 523–533. [[CrossRef](#)]

35. Caselles, V.; Kimmel, R.; Sapiro, G. Geodesic active contours. *Int. J. Comput. Vis.* **1997**, *22*, 61–79. [[CrossRef](#)]
36. Comaniciu, D.; Meer, P. Mean shift: A robust approach towards feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 603–619. [[CrossRef](#)]
37. Coifman, R.R.; Wickerhauser, M.V. Entropy-based algorithms for best basis selection. *IEEE Trans. Inf. Theory* **2002**, *38*, 713–718. [[CrossRef](#)]
38. Rawla, P. Epidemiology of prostate cancer. *World J. Oncol.* **2019**, *10*, 63–89. [[CrossRef](#)]
39. Gopinath, N. Tumor detection in prostate organ using canny edge detection technique. *Int. J. Pure Appl. Math.* **2018**, *118*, 211–217.
40. Cardenas, C.E.; McCarroll, R.E.; Court, L.E.; Elgohari, B.A.; Elhalawani, H. Deep learning algorithm for auto-delineation of high-risk oropharyngeal clinical target volumes with built-in dice similarity coefficient parameter optimization function. *Int. J. Radiat. Oncol. Biol. Phys.* **2018**, *101*, 468–478. [[CrossRef](#)] [[PubMed](#)]
41. Draz, U.; Ali, T.; Yasin, S. Towards Pattern Detection of Proprotein Convertase Subtilisin/kexin type 9 (PCSK9) Gene in Bioinformatics Big Data. *NFC IEFER J. Eng. Sci. Res.* **2018**, *6*, 160–165.
42. Draz, U.; Ali, T.; Yasin, S.; Waqas, U.; Zahra, S.B.; Shoukat, M.A.; Gul, S. A Pattern Detection Technique of L-MYC for Lungs Cancer Oncogene in Bioinformatics Big Data. In Proceedings of the 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 14–18 January 2020; pp. 218–223.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).