Computer Aid-System to Identify the First Stage of Prostate Cancer Through Deep-Learning Techniques

José Gabriel García*, Adrián Colomer*, Fernando López-Mir*, José M. Mossi†, Valery Naranjo*
*Instituto de Investigación e Innovación en Bioingeniería, I3B, Universitat Politècnica de València (UPV), Valencia, Spain
E-mail: {jogarpa7, adcogra, ferlomi, vnaranjo}@i3b.upv.es
†ITEAM Research Institute, Universitat Politècnica de València (UPV), Valencia, Spain
E-mail: jmmossi@dcom.upv.es

***Abstract*—Nowadays, there are high rates of discordance between pathologists when they analyse the biopsy samples to diagnose prostate cancer according to the Gleason scale. Thus, we designed a computer-aid system capable of accurately differentiating between normal tissues and pathological ones at the first stage. Specifically, we made use of an original segmentation algorithm to identify regions of interest and distinguish from them between artefacts (false glands), benign glands and Gleason grade 3 glands. Regarding the building of predictive models, we applied, for the first time, deep-learning algorithms on the previously segmented gland candidates. We compared the results reported by two different convolutional neural networks (CNNs) addressed with distinct classification strategies. The best model reached a multi-class classification accuracy of 0.812 ± 0.033, after performing an in-depth data partitioning per medical history.**

***Index Terms*—Convolutional neural networks, gland segmentation and classification, histological image, prostate cancer**

I. Introduction

Currently, the procedure to diagnose prostate cancer lies in a very tedious manual task that entails a high workload for the pathologists. They use the Gleason score [1] to classify the samples into different grades according to the severity of cancer. Gleason grade 1 and 2 are closely similar to normal tissues, whereas Gleason grades 3, 4 and 5 correspond to cancerous tissues, from less to more severe.

Many authors of the state of the art developed computer-aid and prognosis Gleason systems to help the specialists to reduce its subjectivity level. Some of them carried out an approach based on regions (patches) from which they extracted morphological and textural features [2], [3], or colour and fractal dimensions [4], [5]. Other authors performed a previous segmentation stage to identify regions of interest (ROIs) that allow to include more specific information in their predictive models [6], [7]. They applied features based on texture, morphology, intensity and orientation of the interesting areas to distinguish either between healthy and pathological tissues or between the different grades of the Gleason scale. Note that some studies, such as [8], [9], made use of contextual and structural features to address a strategy based on the classification and evaluation of individual glands, instead of reporting results per image. Nevertheless, the traditional hand-driven learning approach has evolved in recent years towards deep learning, which is based on examples rather than on hand-crafted features [10], [11]. So, in this paper, we make use of convolutional neural networks (CNNs) to create predictive models from the previous segmented gland candidates.

In particular, the main objective of this work is to discriminate between normal tissues (Gleason grade 2) and pathological tissues in their first stages of cancer (Gleason grade 3), since both describe a pattern corresponding to moderately differentiated carcinomas. The pathologists that collaborate in this project maintain that Gleason grades 2 and 3 are characterised as the only ones that present individual glands, so we focused on the accurate identification and classification of the gland units, like the aforementioned studies [8], [9]. However, unlike in [7], [9], where the researchers used the Nuclei-Lumen Algorithm (NLA) to perform a polygonal segmentation of the glands, we applied an original method called *Locally Constrained Watershed Transform* (LCWT) [12] which allows to be faithful to the gland segmentation stipulated by the medical literature. It should be noted that, to the best of the author's knowledge, we are the first who perform a multi-class classification approach applying deep-learning algorithms on segmented gland candidates, instead of using patches or sub-regions. It is also remarkable that, similarly to [9], we discerned between artefacts, benign and pathological glands, as shown in Fig. 1, unlike in [8], where the authors did not take into account the artefact elements.

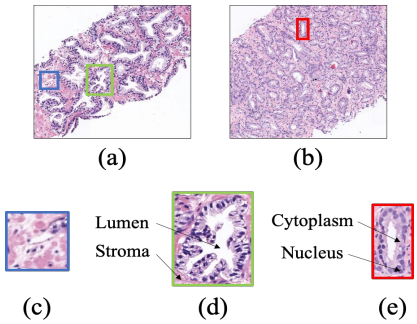


Fig. 1. (a) Tissue with a pattern of Gleason grade 2. (b) Sample of Gleason grade 3. (c) Artefact. (d) Benign gland unit. (e) Pathological gland unit.

This work has been funded by the Ministry of Economy, Industry and Competitiveness under the SICAP project (DPI2016-77869-C2-1-R). The work of José Gabriel García and Fernando López-Mir have been supported by the State Research Spanish Agency PTA2017-14610-I, and by the UPV (Grant PAID-10-18), respectively. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan V GPU used here.

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

II. Material

The experiments detailed in this paper are carried out using a private database composed of 35 whole-slide images (WSI) belonging to 25 different patients of the *Hospital Clínico Universitario de València*. In particular, an expert in pathological anatomy determined that 17 whole-slide images from 8 patients corresponded to benign samples, whereas the other 18 whole-slide images relative to the rest of patients corresponded to cancerous tissues in an early stage. In order to construct the database from the WSI, we performed a simple filtering to remove those pixels of the samples without tissue information. Then, we implemented a sliding window protocol to obtain sub-images of 1024×1024 dimensions corresponding to an optical magnification of $10\times$, which are characterised by presenting an optimal resolution to address the approach based on gland segmentation. Specifically, we obtained 814 sub-images relative to healthy tissues and 614 corresponding to pathological tissues, which form the final database.

III. Methods

A. Segmentation procedure

Clustering. In order to extract the necessary inputs for the $LCWT$ segmentation algorithm, we addressed a clustering step based on the *k-means* technique to group the pixels of the RGB image I into *k=4* different classes corresponding to the lumen, nucleus, cytoplasm and stroma components. Unlike in [7], [9], where the authors only use RGB images, we made use of different colour spaces with the aim of taking into account the optimal intensity level of the pixels, depending on the component mask that we wanted to obtain. Specifically, in order to compute the map of lumen candidates, whose pixels correspond to the lightest intensity levels of the RGB image, we extracted the saturation contribution S_{HSV} from the HSV colour space since it allows to highlight the brightest pixels according to (1), where MAX and MIN are the maximum and minimum of the R, G and B components.

$$S_{HSV} = \begin{cases} 0, & \text{if } MAX = 0 \\ 1 - \frac{MIN}{MAX}, & \text{otherwise} \end{cases} \quad (1)$$

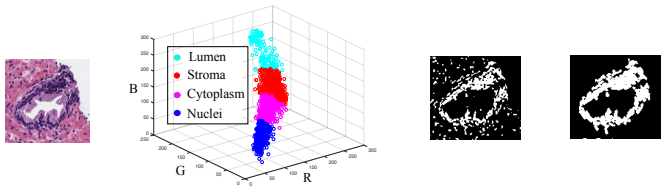
After extracting the S_{HSV} grey-scale image, we performed the *k-means* clustering, with *k=4*, to obtain the outputs corresponding to the lumen candidates (L). From the CMYK colour space, we obtained the cyan channel C_{CMYK} to achieve the maps relative to the cytoplasm and stroma components. Since the cyan colour allows to emphasise the blue channel and reduce the red contribution $C_{CMYK} \rightarrow (0, 1, 1)$, it is able to accurately discern between the purple and magenta contributions of the RGB images, which are related to the cytoplasm and stroma pixels, respectively. In the same way as before, we carried out the *k-means* algorithm and used the output to obtain two binary maps, (C) and (S), by means of exclusively selecting the pixels associated to the cytoplasm and stroma structures. Regarding the attainment of the nuclei mask (N) exposed in Fig. 2 (a-c), we used a reshaped RGB image similarly to [9], from which we randomly computed the 5% of the pixels to calculate the centroid μ of each cluster, improving the computational cost. In this case, after performing the *k-means* algorithm, we calculated the Euclidean distance between the intensity value of each pixel p and the coordinates of the centroid μ to group the pixels into its cluster.

Post-processing. Once all of the candidates' maps are obtained, we addressed a post-processing step with the aim of removing the noise of the binary masks that constitute the inputs to the $LCWT$ segmentation method. Specifically, in order to achieve the final mask of lumen candidates LM, we made use of a morphological operation called *area opening* (γ^a_λ) (2), to discard those 8-connected objects L_i with an area smaller than $\lambda = 20$ pixels from the mask L. Then, we also applied an operation of *dilation* $\delta_B(L) = L \oplus B$, where B is a disc structuring element (SE) with a radius $r = 1$ that allows to increase the area of the possible lumens.

$$\gamma^a_\lambda(L) = \bigcup \{L_i \mid i \in I,\ Area(L_i) \geq \lambda\} \quad (2)$$

To extract the cytoplasm CM and stroma SM final masks, we applied an *opening* filter described by $\gamma_B(X) = (X \ominus B) \oplus B$ (with $X = CM$ or $X = SM$), followed by an *area opening* with the same parameters as before to discard the structures of pixels with a low connectivity. Regarding the extraction of the nuclei mask NM (Fig. 2 (d)), we computed a *dilation* to link the pixels relative to each nucleus element N_i around the glands, and another *area opening* with $\lambda = 15$ pixels to remove those non-epithelial nuclei scattered across the stroma.

Gland unit segmentation. There are some state-of-the-art studies, such as [13], that proposed convolutional neural networks to address the segmentation of individual glands from the biopsy samples. However, in this paper, we use an algorithm that does not require any manual effort to build the segmentation models, unlike deep-learning methods. In particular, we made use of a variant of the traditional *watershed transform with markers* [14], which is based on mathematical morphology to perform the segmentation. This technique consists in converting a grey image I_G in a topographic surface where each defined regional minimum m_i represents a region called catchment basin CB_i. From them, an inundation process is simulated to increase the water level inside the basins, like a process of region growing in a 3D space. The points in which two or more basins come into contact are established as the watershed lines to segment the catchment



(a) (b) (c) (d)

Fig. 2. (a) RGB image of a specific gland. (b) 3D pixels distribution according to their intensity levels. (c) Output corresponding to the map of nucleus candidates. (d) Final nuclei mask after the post-processing step.

basin CB_i corresponding to each regional minimum m_i. Nevertheless, according to the medical literature, a gland unit is defined by the epithelial layer of non-connected nuclei that surrounds its lumen [9]. So, it is necessary to perform a variant of the aforementioned technique able to flood the basins without water leaking through the space between adjacent nuclei. For this reason, unlike Naik et al. [6] who described the segmentation by the cytoplasm component, instead of by the nucleus elements, we made use of the $LCWT$ algorithm [12] to determine the watershed lines from regions whose boundaries remain open, i.e. from the non-connected nuclei of each gland candidate. This technique was implemented for the first time (on histological images) in our previous work [15], with the final aim of removing the artefacts (false glands).

In this paper, making use of the masks previously achieved from the clustering and post-processing stages, we define the pixels $p_i \subset L_i$ as internal markers; whereas we use the pixels $p_i \subset S$ as external markers. The objective is to initialise the region growing step from the aforementioned markers, but considering the main novelty of the $LCWT$ technique: the constraints. $LCWT$ algorithm requires an input image, besides the I_G, that acts as a restriction to the markers' progress. Particularly, taking into account that the epithelial nuclei elements $N_i \subseteq N$ must be the gland boundaries, we defined the nuclei mask N as the input image, and imposed the constraints introducing different structuring elements SEs that allow to define the *constrained catchment basin* CCB_i for each regional minimum m_i, according to (3):

$$CCB_i = \bigcup_{\delta \geq 0} [V_i(\delta) \oplus SE] \tag{3}$$

V_i is the constrained partial catchment basin defined for each marker m_i as follows:

$$V_i(\delta) = R_i(\delta) \setminus \bigcup_{i \neq j} [R_j(\delta) \oplus SE \oplus SE], \tag{4}$$

where R_i is the threshold set obtained making use of the *Minkowski distance function* $\mathfrak{T}_f$ for each pixel p and marker m_i, according to:

$$R_i(\delta) = \{p | \mathfrak{T}_f(m_i, p) \leq \delta\}, \quad \forall \delta \geq 0 \tag{5}$$

Note that two disk SE with different radius values should be used depending on if m_i corresponds to an internal or an external marker. In particular, we used $r = 1$ and $r = 5$, respectively, to address the water flood from each CCB_i, similarly to the traditional *watershed* technique. However, in this case, the process was stopped when the size of some SE was greater than the distance between adjacent nuclei N_i. This fact allows to establish the watershed lines over the contact points between the different constrained catchment basins, as shown in Fig. 3. So, being $f : D \to \mathbb{N}$, the locally constrained watershed transform of f is:

$$LCWT(f) = D \setminus \bigcup_{i \in I_G} CCB_i \tag{6}$$

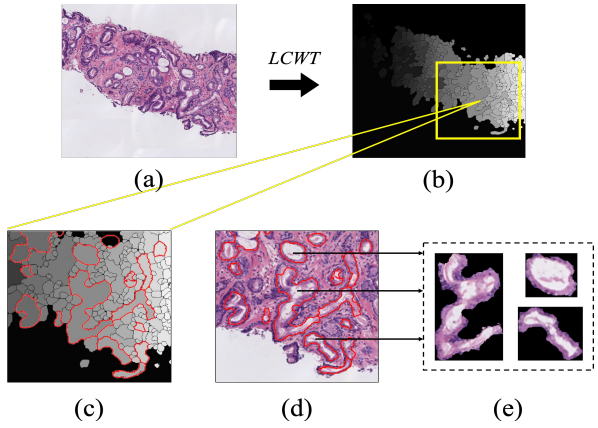


Fig. 3. (a) Original 1024×1024 RGB image. (b) Output of the $LCWT$ algorithm. (c) Segmented regions that contain some L_i highlighted in red. (d) Boundaries of the gland candidates overlapped over the original image. (e) Gland candidates extracted separately.

After applying the $LCWT$ algorithm, we discarded the output regions that did not contain any lumen candidate L_i inside them, as shown in Fig. 3 (c-d). We finally stored each gland candidate G_i separately (see Fig. 3 (e)) to construct, in an innovative way, the predictive models from them.

B. Classification strategy

Data partitioning. It should be noted that we did not perform a merely random data partitioning, but we designed a rigorous strategy based on separating the samples into five data sets (folds), in which each of them contained a similar number of glands candidates and a balanced number of elements of each class. In addition, we performed the data partitioning taking into account the patient medical history, so the gland candidates corresponding to a specific patient ID were always stored in the same data subset (fold). Once the images G_i were established in the different subsets according to its patient ID, we implemented an external *5-fold cross-validation* technique with the aim of providing reliable results trough avoiding the randomness effect of the partitioning. In this way, we carried out five iterations using in each of them, 4 folds to train a classification model and 1 fold to test it. Moreover, in each iteration, we performed a validation set composed of the 10% of the gland candidates images corresponding to the 4-training sets, with the aim of assessing the overfitting of the CNNs.

Network architectures. It is important to note that the depth of the neural networks is really determinant to reach high accuracy results in the classification stage. For this reason, we propose two different approaches based on deep-learning algorithms in order to build predictive models from two network architectures with a distinct structure and depth.

As we show in Fig. 4, we made use of a very popular CNN called *Very Deep Convolutional Networks for Large-Scale Image Recognition VGG19* [16]. In particular, in the same way as [17], we used 3×3 receptive fields to convolve them with the input at every pixel, and we defined a stride $= 1$ to move each filter of the convolutional layers along the

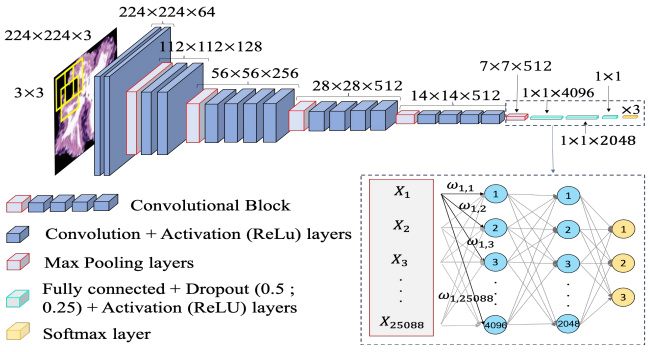


Fig. 4. *VGG19* network architecture used to build predictive models from gland candidates of histopathological prostate images.

image. However, in order to avoid the possible overfitting, we included two dropout layers (of 0.5 and 0.25) after each fully-connected layer with the aim of randomly removing the 50% and 25% of neurons, respectively. In addition, we defined the second hidden layer with 2048 neurons, instead of the 4096 proposed by the original *VGG19*. Moreover, we performed a final softmax layer composed of three neurons to discriminate between artefacts, benign glands and Gleason grade 3 glands.

Note that our classification database consists of 3.200 artefacts, 3.195 benign glands and 3.000 pathological glands, which is an insufficient number of specific images to train from scratch a neural network with the depth of the *VGG19* architecture, whose base model is composed of four convolutional blocks (see Fig. 4). For this reason, we computed a transfer learning technique called *fine-tuning* [18], which allows to make use of the wide knowledge acquired by the *VGG19* architecture when it was trained on the *ImageNet* data set. In this way, we used the weights ω_i pre-trained with around 14 million natural images belonging to 1000 classes for applying a *deep fine-tuning* [19] by means of freezing the coefficients of the three first convolutional blocks of the CNN. Then, we used our specific data set of prostate gland candidates to re-train the weights of the last convolutional block. We also applied *data augmentation* [20] to increase the database of specific images through the creation of synthetic samples similar to the original gland candidates and with the same ground-truth label. Particularly, we defined the aggressive factor ratio as $\alpha = 0.02$ to address the intensity and geometric transformations to obtain the artificial samples. We also initialised the batch size hyperparameter as $\beta = 16$ to update the weights when 16 images are forward propagated and its error is calculated, according to the categorical cross-entropy loss function (7):

$$\mathcal{L}(y, \hat{y}) = -\sum_{i} y_i \; log(\hat{y}_i) \tag{7}$$

where y_i is the ground-truth label of a specific image and $\hat{y}_i$ is the predicted label for such image. We updated the weights according to the Stochastic Gradient Descent (SGD) (8), defining a learning rate $\eta = 10^{-5}$ and a momentum $\gamma = 0.98$ to improve the convergence rate of the network.

$$\omega(t+1) = \omega(t) + V(t+1), \tag{8}$$

where $\omega(t+1)$ is the updated weight and $V(t+1)$ is described as follows, considering $V(t) = 0$:

$$V(t+1) = \gamma V(t) - \eta \frac{d}{d\omega} \mathcal{L}(y, \hat{y}) \tag{9}$$

We initialised $\omega(t)$ with random low values and established a maximum number of epochs $N = 100$ with a stop criterion $sc = 15$ epochs to avoid the overfitting in the training stage.

Besides that, we designed another CNN with lower depth to train a network architecture with the specific images from scratch. In this case, we did not carry out transfer learning techniques since we designed a shallow CNN (*gpNet*) whose base model is composed of two convolutional blocks with two and three convolutional layers, respectively. We also included batch normalisation and dropout layers (with coefficients of 0.25) after each convolutional block with the aim of reducing overfitting. Additionally, we designed a top model with a fully-connected layer composed of 512 neurons and a dropout layer of 0.25 followed by a softmax layer with three neurons. Note that we kept the initialisation of the rest of hyperparameters to analyse the importance associated with the depth of the CNN, as well as the relevance of the transfer learning techniques. It is also important to remark that we addressed this stage using the framework of *Keras* with *Tensorflow* as back-end [21].

IV. Results and Discussion

Regarding the segmentation stage, the modified $LCWT$ algorithm was evaluated in our previous work [15], in which we reported a considerable outperforming with respect to the publicly available NLA algorithm proposed in [9] and recently implemented in [7]. In particular, we achieved a Jaccard Index $JI = 0.739 \pm 0.108$ and a Dice coefficient $D = 0.849 \pm 0.121$, from 500 benign and pathological glands randomly selected. However, the main novelty of this paper lies in the application of deep-learning techniques, after the LCWT segmentation, to discriminate between false, benign and pathological glands, with the global aim of distinguishing between normal tissues and cancerous ones at the first stage. We expose in Table I the results obtained from different figures of merit after applying both the modified *VGG19* and the developed *gpNet* architectures. Additionally, in order to show the performance of each CNN, we represent in Fig 5, the ROC curves corresponding to both classification experiments, i.e. artefacts Vs. glands and benign Vs. pathological glands.

Note that Table I shows a very slight outperforming of the *VGG19* architecture over the *gpNet*, which may be due to the knowledge learned through training the model with a database composed of a much higher number of images. However, *gpNet* is much more efficient than *VGG19* because it has half of trainable parameters (67.148.803 Vs. 131.185.731, respectively), which allows to build the models more quickly during the learning process. For this reason, since *gpNet* neural network requires much fewer convolutions to address

TABLE I
CLASSIFICATION RESULTS ACHIEVED FROM TWO DIFFERENT APPROACHES BASED ON DEEP-LEARNING ALGORITHMS.

	Artefacts Vs. Glands		*Healthy Vs. Cancer*	
	VGG19	**gpNet**	**VGG19**	**gpNet**
Sensitivity	0.900 (0.020)	**0.913** (0.015)	**0.747** (0.109)	0.705 (0.138)
Specificity	**0.938** (0.026)	0.920 (0.035)	0.780 (0.123)	**0.800** (0.141)
F-Score	**0.891** (0.032)	0.887 (0.022)	**0.753** (0.058)	0.730 (0.095)
AUC	**0.974** (0.010)	0.972 (0.008)	**0.889** (0.036)	0.886 (0.079)
Accuracy	**0.925** (0.018)	0.919 (0.053)	0.816 (0.030)	**0.821** (0.084)

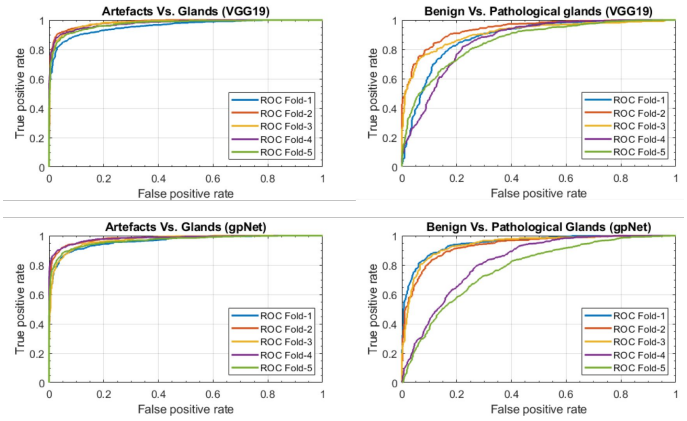


Fig. 5. ROC curves obtained after testing data from the different CNNs.

the forward-propagation step during the prediction, we can conclude that the *gpNet* model is better, also attending to the similarity of the results reached for all metrics with respect to the *VGG19* architecture. Additionally, we can also observe in Fig. 5 a great difference depending on the experiment under study. The discrimination between artefacts and glands is successfully achieved, reporting high values for the different figures of merit in both architectures. Nevertheless, the experiment focused on distinguishing between healthy and cancerous glands shows the most prominent differences, since the results provided by *gpNet* neural network differ a lot depending on the test data set, whereas *VGG19* reports similar results regardless of the fold, as shown in Fig. 5. Based on this fact, we can determine that *VGG19* is a more robust model because it is not affected by the partitioning criterion at patient level.

V. CONCLUSION

In this paper we presented an automatic system based on the identification of the first stage of prostate cancer from histopathological images. We made use of the LCWT segmentation algorithm to obtain the gland units separately to build predictive models from them. In particular, we performed two convolutional neural networks with a different depth to address the classification and prediction stages from diverse strategies. Finally, the proposed *gpNet* is considered the best model in terms of computational cost and quality of results.

REFERENCES

[1] D. Gleason, "Histological grading and clinical staging of prostatic carcinoma," *Urologic pathology. The prostate*, vol. 171, 1977.

[2] S. Doyle, M. Hwang, K. Shah, A. Madabhushi, M. Feldman, and J. Tomaszeweski, "Automated grading of prostate cancer using architectural and textural image features," in *Biomedical imaging: from nano to macro, 2007. ISBI 2007. 4th IEEE international symposium on*. IEEE, 2007, pp. 1284–1287.

[3] Á. Esteban, A. Colomer, V. Naranjo, and M. Sales, "Granulometry-based descriptor for pathological tissue discrimination in histopathological images," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 1413–1417.

[4] A. Tabesh, M. Teverovskiy, H.-Y. Pang, V. P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, "Multifeature prostate cancer diagnosis and gleason grading of histological images," *IEEE transactions on medical imaging*, vol. 26, no. 10, pp. 1366–1378, 2007.

[5] S.-K. Tai, C.-Y. Li, Y.-C. Wu, Y.-J. Jan, and S.-C. Lin, "Classification of prostatic biopsy," in *Digital Content, Multimedia Technology and its Applications (IDC), 2010 6th International Conference on*. IEEE, 2010, pp. 354–358.

[6] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi, "Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information," in *MIAAB workshop*. Citeseer, 2007, pp. 1–8.

[7] P. Leo, R. Elliott, N. N. Shih, S. Gupta, M. Feldman, and A. Madabhushi, "Stable and discriminating features are predictive of cancer presence and gleason grade in radical prostatectomy specimens: a multi-site study," *Scientific Reports*, vol. 8, no. 1, p. 14918, 2018.

[8] T. Xia, Y. Yu, and J. Hua, "Automatic detection of malignant prostatic gland units in cross-sectional microscopic images," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 1057–1060.

[9] K. Nguyen, A. Sarkar, and A. K. Jain, "Structure and context in prostatic gland segmentation and classification," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2012, pp. 115–123.

[10] O. J. del Toro, M. Atzori, S. Otálora, M. Andersson, K. Eurén, M. Hedlund, P. Rönnquist, and H. Müller, "Convolutional neural networks for an automatic classification of prostate tissue slides with high-grade gleason score," in *Medical Imaging 2017: Digital Pathology*, vol. 10140. International Society for Optics and Photonics, 2017, p. 101400O.

[11] E. Arvaniti, K. S. Fricker, M. Moret, N. J. Rupp, T. Hermanns *et al.*, "Automated gleason grading of prostate cancer tissue microarrays via deep learning," *bioRxiv*, p. 280024, 2018.

[12] R. Beare, "A locally constrained watershed transform," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 7, pp. 1063–1074, 2006.

[13] J. Ren, E. Sadimin, D. J. Foran, and X. Qi, "Computer aided analysis of prostate histopathology images to support a refined gleason grading system," in *Medical Imaging 2017: Image Processing*, vol. 10133. International Society for Optics and Photonics, 2017, p. 101331V.

[14] S. Beucher and F. Meyer, "The morphological approach to segmentation: the watershed transformation," *Optical Engineering-New York-Marcel Dekker Incorporated-*, vol. 34, pp. 433–433, 1992.

[15] J. G. García, A. Colomer, V. Naranjo, F. Peñaranda, and M. Sales, "Identification of individual glandular regions using lcwt and machine learning techniques," in *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 2018, pp. 642–650.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[17] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in *Advances in neural information processing systems*, 2014, pp. 568–576.

[18] S. Hoo-Chang, H. R. Roth, M. Gao, and et al., "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, no. 5, p. 1285, 2016.

[19] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.

[20] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: when to warp?" in *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE, 2016, pp. 1–6.

[21] F. Chollet, "Keras 2015," https://keras.io/, accessed: 25-10-2018.