ORIGINAL PAPER

Automated Gleason grading of prostate cancer using transfer learning from general-purpose deep-learning networks

Mircea-Sebastian Şerbănescu[1], Nicolae Cătălin Manea[1], Liliana Streba[2], Smaranda Belciug[3], Iancu Emil Pleşea[4–6], Ionica Pirici[7], Raluca Maria Bungărdean[8], Răzvan Mihail Pleşea[9]

[1]*Department of Medical Informatics and Biostatistics, University of Medicine and Pharmacy of Craiova, Romania*

[2]*Department of Medical Oncology, University of Medicine and Pharmacy of Craiova, Romania*

[3]*Department of Computer Science, University of Craiova, Romania*

[4]*Department of Pathology, Carol Davila University of Medicine and Pharmacy, Bucharest, Romania*

[5]*Department of Pathology, Fundeni Clinical Institute, Bucharest, Romania*

[6]*Doctoral School, Carol Davila University of Medicine and Pharmacy, Bucharest, Romania*

[7]*Department of Human Anatomy, University of Medicine and Pharmacy of Craiova, Romania*

[8]*Department of Pathology, Iuliu Haţieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania*

[9]*Department of Cellular and Molecular Biology, University of Medicine and Pharmacy of Craiova, Romania*

Abstract

Two deep-learning algorithms designed to classify images according to the Gleason grading system that used transfer learning from two well-known general-purpose image classification networks (AlexNet and GoogleNet) were trained on Hematoxylin–Eosin histopathology stained microscopy images with prostate cancer. The dataset consisted of 439 images asymmetrically distributed in four Gleason grading groups. Mean and standard deviation accuracy for AlexNet derivate network was of 61.17±7 and for GoogleNet derivate network was of 60.9±7.4. The similar results obtained by the two networks with very different architecture, together with the normal distribution of classification error for both algorithms show that we have reached a maximum classification rate on this dataset. Taking into consideration all the constraints, we conclude that the resulted networks could assist pathologists in this field, providing first or second opinions on Gleason grading, thus presenting an objective opinion in a grading system which has showed in time a great deal of interobserver variability.

Keywords: deep learning, prostate cancer, Gleason grading system, transfer learning, AlexNet, GoogleNet.

⊟ Introduction

Prostate cancer is the second most common cancer diagnosed in men, with over 10% of the male population being diagnosed during their lifetime [1].

Despite the advances in all imaging medical fields [ultrasound, computed tomography (CT), magnetic resonance imaging (MRI)], the "gold standard" for diagnosis remains the microscopic tissue examination performed by a pathologist.

Developed in 1966, the Gleason grading system (GGS), together with more recent revisions [2, 3], stratifies prostate cancers based on architectural patterns as a reflection of their biology. GGS remains the most powerful predictor of prognosis in almost all prostate cancer studies, being widely used in standardized patient management [4]. The GGS classifies prostate cancer growth patterns in five grades (some having sub-grades), and summing up the two most common grades results the final Gleason score, which ranges between two and 10, and which is supposed to stratify patient's outcome.

Without any intention to alter the role of GGS in prognostic and patient management, different studies have shown that the system suffers from two major drawbacks: the first being related to the grading itself, while the second is related to the quantity of the biological product that is being analyzed. The first drawback refers to the suboptimal interobserver and intraobserver variability, with reported discordance ranging from 30% to 53% [5–10], and with imprecise differences between classes on standard feature extraction algorithms, such as fractal analysis [11–13]. The second drawback refers to the fact that the score is computed using the dominant and subdominant patterns of the cancer. In healthy subjects, the size of the prostate is approximately $3\times4\times5$ cm, whereas in pathological conditions the size increases three times. The size of the prostate, combined with the possible transurethral resection of the prostate, produce a considerable sample size thus requiring a careful assessment of first and second most frequent patterns.

Taking into consideration the mentioned drawbacks, the GGS proves that these tasks are time-consuming if they are performed by the pathologist, and also that they imply material cost with high interobserver and intra-observer variability. Hence, this task is suitable for computer-aided medical diagnosis systems.

Multiple computer-aided diagnosis (CAD) systems have been proposed for GGS automatization with different approaches, from standard artificial intelligence (AI) algorithms [14–16] to the newer deep-learning (DL) approaches [17–19].

Aim

The aim of current research was to develop a DL algorithm that uses transfer learning from well-known pre-trained networks capable of classifying histological images according to the GGS with high accuracy (ACC).

⯐ Patients, Materials and Methods

Patient inclusion, ethical data and image retrieval

We prospectively included 439 images from 83 patients who underwent total prostate resection, following a diagnosis of prostate cancer, between January 2013 and December 2015 at the Municipal Clinical Hospital of Cluj-Napoca, Romania. All presumptive diagnoses were made by combining clinical and imaging data and confirmed through pathology. All patients signed informed consent forms and agreed to tissue harvesting for research purposes, as per usual Hospital Guidelines. We ascertained that our study did not interfere with therapeutic or diagnostic procedures. We harvested tissue from whole-organ resection specimens for usual diagnosis and staging by two expert pathologists; afterwards, images were obtained as per the below protocol.

In the original GGS, revised several times by the author himself [20–23], the architectural patterns of tumor proliferation were labeled based on the five main classes and subtypes. Thus, the Gleason pattern 1 (very rare) is characterized by a very well-differentiated proliferation, consisting of medium-sized, round or oval, uniform glands, arranged very compactly but separated from each other. Compared to pattern 1, the glands from pattern 2 have a greater variability in size and shape and are separated by stromal bays, with an average distance of interglandular separation smaller than the diameter of a gland. Pattern 3 is considered the form with moderate differentiation. Gleason described this pattern as having three distinct architectural morphological aspects, designated as patterns 3A, 3B, and 3C. The 3A subtype is characterized by the presence of isolated glands of medium size, with a variable

shape, consisting of elongations, twists and angles that can also have sharp angles. Subtype 3B has, in principle, the same architectural appearance as pattern 3A, with the only difference that the tumor glands are smaller. Subtype 3C is composed of ducts or ducts expanded with sieve or intraluminal papillary tumor masses which, in accordance with the hypothesis that this pattern would really represent an intraductal proliferation, have smooth, rounded contours, as of relaxed ductal profiles. Pattern 4 is considered a poorly differentiated high-grade proliferation. Gleason described this pattern as having two distinct morphological aspects – patterns 4A and 4B. Subtype 4A – tumor proliferation is composed of cells that may have either a fused micro-acinar arrangement or a cribriform or a papillary one. Tumor cells form either infiltrative masses with a totally irregular appearance or strings or cords of epithelial malignant cells. Subtype 4B – carcinomatous proliferation of this type is identical in terms of architectural appearance to the other subtype of pattern 4, with the only difference that the cells that form it have a clear cytoplasm. Pattern 5 is the weakest differentiated form of prostate cancer and was also divided by Gleason into two subtypes: 5A and 5B. Subtype 5A resembles the "comedo" type of intra-ductal breast carcinoma, presenting as tumor masses in which the cells have a chordal or cylindrical arrangement, with a cribriform, papillary appearance (as in subtype 3C) or solid, with smooth, rounded edges, whose central area is typically occupied by necrotic detritus. Subtype 5B consists of tumor areas with irregular edges formed by anaplastic tumor cells.

The recent revisions simplified somehow the original system. Thus, pattern 3 remained with only two subtypes mainly (original 3A and 3B). Pattern 4 included cribriform glands larger than benign glands and with an irregular border, finally consisting of poorly formed glands of either cribriform or fused architecture [2, 3, 24].

We included in our study 439 Hematoxylin and Eosin (HE) images with monotonous patterns that were classified according to GGS by two pathologists in four groups: Gleason pattern 2 ($n=57$), Gleason pattern 3 ($n=166$), Gleason pattern 4 ($n=182$), and Gleason pattern 5 ($n=34$). The dataset had no image with pattern 1. The images, 32-bit red, green, blue (RGB) color space, were cropped at 512-by-512 pixels from whole slide images scanned with Leica Aperio AT2, using a 20× apochromatic objective. A sample of each pattern is presented in Figure 1.
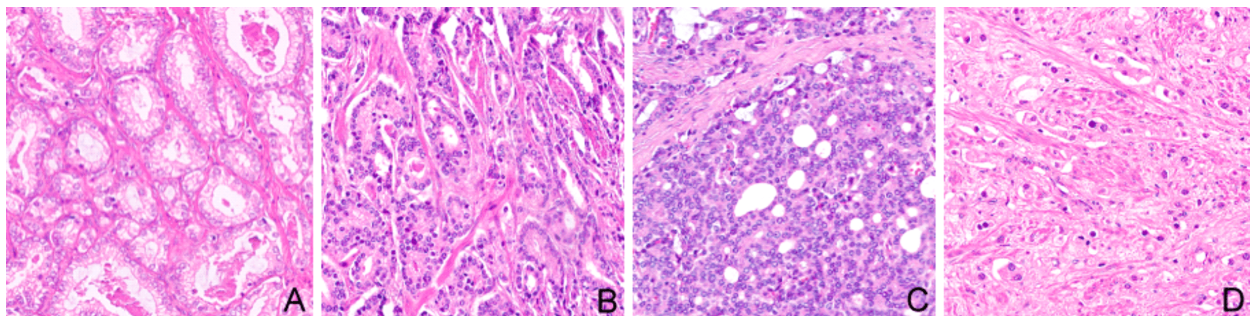

Figure 1 – *Samples from the dataset (HE staining, ×200): (A) Gleason pattern 2; (B) Gleason pattern 3; (C) Gleason pattern 4; (D) Gleason pattern 5.*

Deep neural network algorithms and methods

Two DL algorithms were developed using transfer learning from AlexNet [25] and GoogleNet [26] networks.

AlexNet is a convolutional neural network that has been trained on more than a million images from the ImageNet database available free of charge on http://www.image-net.org. The network is eight layers deep and classifies images into 1000 categories of objects from the real world. The network has an input image size of 227-by-227 pixels, with a 32-bit RGB color space. In order to fit the input layer of the network, we resized the images at 227-by-227 pixels keeping the 32-bit RGB color space. The last layers of the network were replaced in order to classify the input images in four classes according to the GGS patterns.

GoogleNet is a convolutional neural network that is 22 layers deep, also trained on ImageNet. The network classifies images in the same 1000 object categories as AlexNet. The network has an input image size of 224-by-224 pixels, with a 32-bit RGB color space. In order to fit the input layer of the network, the images processed for AlexNet were resized at 224-by-224 pixels keeping the 32-bit RGB color space. The last layers of the network were replaced in order to classify the input images in four classes according to the GGS patterns.

We used 85% of the images for training and the remaining 15% for testing.

We performed the algorithm implementation and the statistical assessment in MATLAB (MathWorks, USA).

⊟ Results

DL being a stochastic algorithm, a certain number of runs is needed to be performed in order to obtain robust and trustworthy results. A suitable statistical power (two-tailed type of null hypothesis with default statistical power goal $p \geq 95\%$ and type I error $\alpha = 0.05$ – level of significance) can be achieved through 100 independent computer runs. When designing an experiment, one needs to perform statistical power analysis together with sample size. Precision and ACC may lack due to a low sample size, whereas a high sample size may lead to an increase of computational and time costs, without a gain in performance. The standard 10-fold cross-validation has been used in our study. The DL algorithms have been independently run 100 times in a complete 10-fold cross-validation cycle.

When running multiple times a stochastic algorithm, we encounter differences between ACCs, hence we need to compute also the standard deviation (SD) of the ACCs obtained. If we obtain low value SD, then our model is more stable.

In order to perform different statistical tests, we first need to verify whether the sample data has a normal distribution or not. If the data does not have a normal distribution, the results might be affected, due to the existence of outliers. In our study, we applied the Kolmogorov–Smirnov and the Shapiro–Wilk W tests.

Mean and SD on 100 algorithm runs for the ACC was of 61.17 ± 7 and of 60.9 ± 7.4 for AlexNet and GoogleNet, respectively.

Samples of the training sequences are presented in

Figure 2 and the confusion matrix from the same runs, applied to the whole dataset in Figure 3.

Due to its simpler architecture, we packed the resulted algorithm from AlexNet in a standalone application (Microsoft Windows) capable of learning on new image datasets, transferring knowledge form pre-trained networks, and classifying new images. A preview of the application is presented in Figure 4.

The results of the Kolmogorov–Smirnov and Shapiro–Wilk W tests are displayed in Table 1.

Table 1 – *Normal distribution assessment*

Algorithm	Kolmogorov–Smirnov (K–S)		Shapiro–Wilk (S–W) W	
	K–S max D	Lilliefors p	S–W W	p-level
AlexNet	0.107	0.2	0.976	0.41
GoogleNet	0.122	0.1	0.977	0.46

From Table 1, we can see that regardless the algorithm, the sample data is normally distributed. Hence, we can proceed to compare how the two algorithms perform using the t-test for independent samples. The results could be objectively compared because both algorithms have been run in the same conditions (100 computer-runs/10-fold cross validation).

The results from the t-test are displayed in Table 2.

Table 2 – *Statistical assessment the means of the two algorithms*

Variable	t-test / p-level
AlexNet *vs.* GoogleNet	0.62 / 0.53

As shown in Table 2, there is no significant difference in means (p-level >0.05) between the two networks concerning the testing performances. Thus, both algorithms perform the same on this dataset.

⊟ Discussions

When using the GGS to stratify the aggressiveness of prostate cancer, pathologists are often confronted with a classification problem pertaining to the varying nature of tissue loss and resulting morphology. Thus, it is imperative that common morphological descriptors should be identified and applied in medical practice, in order to unify the opinions of different clinician pathologists and provide a closer-to-accurate prognosis. Survival as well as different approaches for treatment depend on this step. Diagnosis between malignant and benign histological tissue is possible by using semi-automated computer-assisted methods [27, 28]. Previous approaches relied on identifying distinctive features [29–31] and training neural networks to identify and quantify such pre-determined and operator-dependent markers. Recently, DL greatly reduced this effort, at the expense of transparency – basically, the technique can be labeled as a "black box" approach, since the operator is "blinded" to the way, the computer identifies the significant features. As previous authors stated, this may be a serious impediment towards widespread acceptance and regulatory approval [32].

In our study, we used digitized images of prostate cancer microscopy, classified as Gleason patterns 2 to 5. We have proven here that it is feasible to use a DL approach to tackle this medical classification problem,

irrespective of the network architecture – either using AlexNet or GoogleNet, which produced comparable results. This computerized approach may lead to successful implementation of medical-grade tools aimed to both second the opinion of a medical expert, or to provide intermediate diagnosis in tertiary medical centers that lack immediate access to a pathologist expert or which can rely on telemedicine for faster decision making. In our opinion, it is not, however, advisable to not involve a human specialist in the process, since the algorithms also have specific limitations that are inherent due to the imperfect nature of the Gleason scoring system.
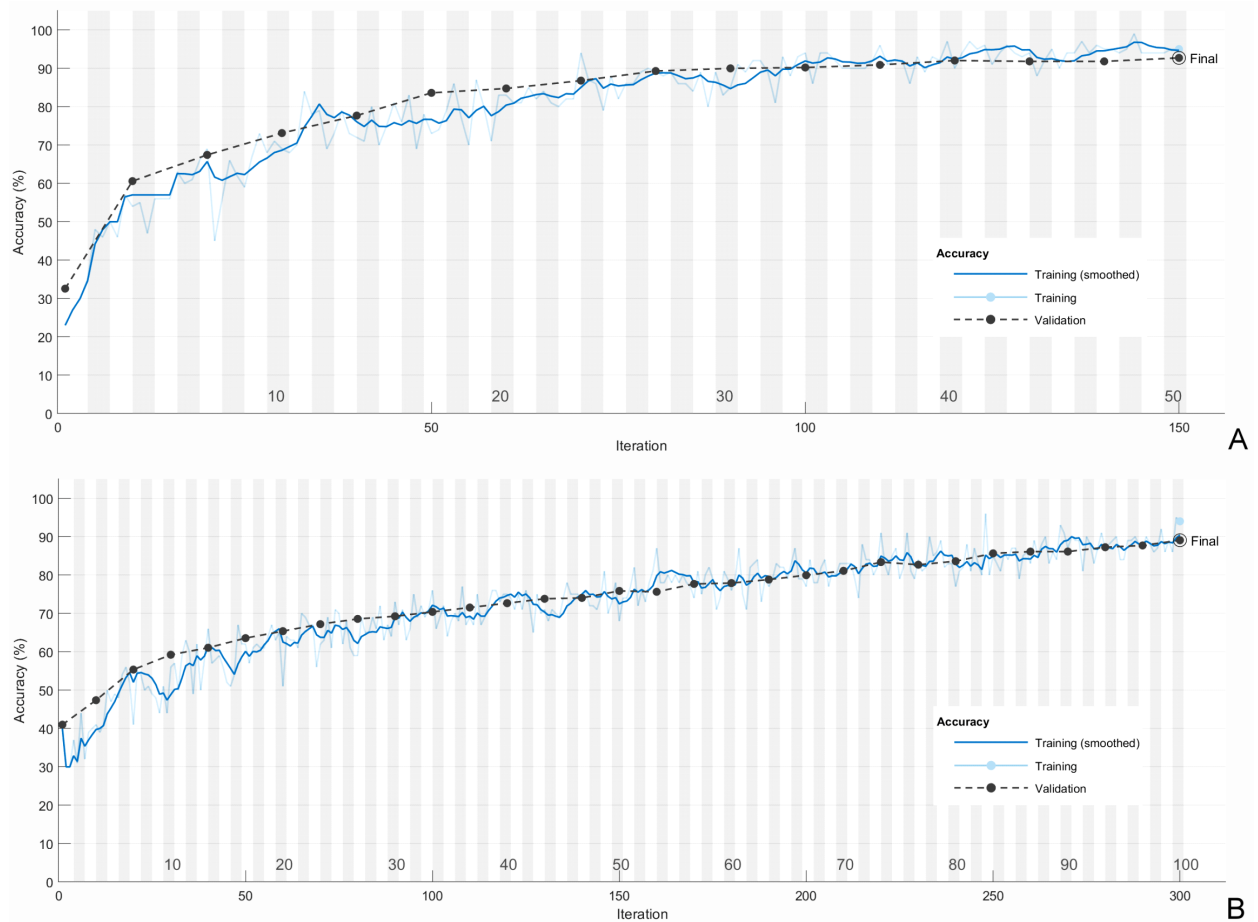

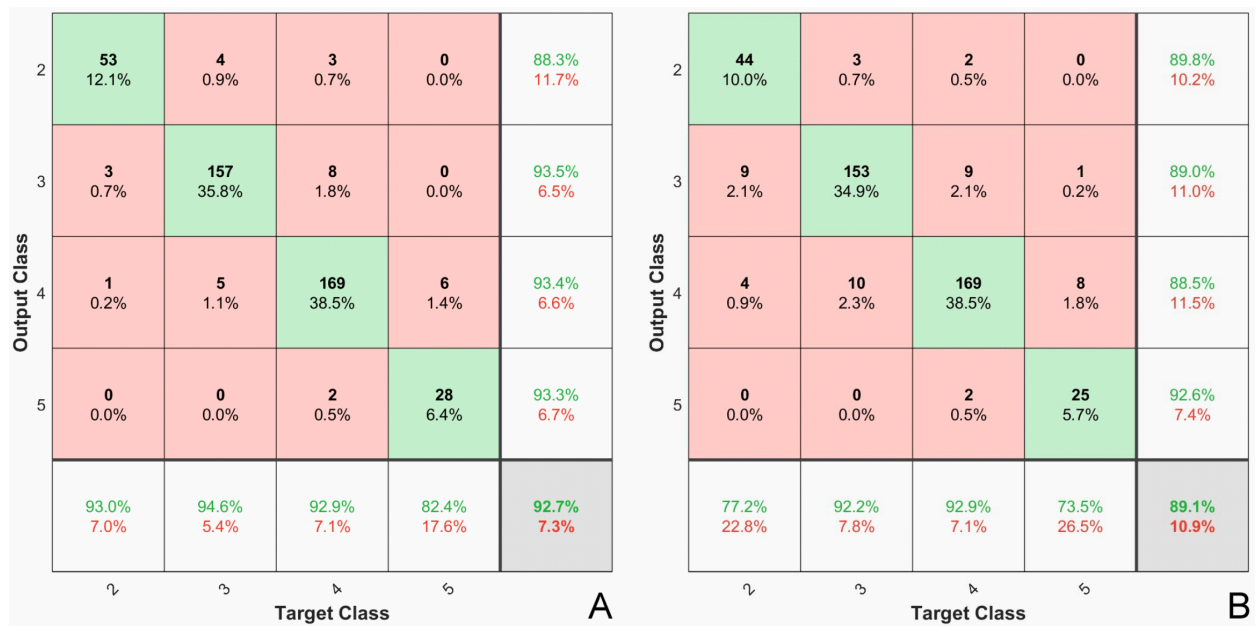**Figure 2 – *Training process: (A) AlexNet; (B) GoogleNet.***

Figure 3 – *Confusion matrix heatmap: (A) AlexNet; (B) GoogleNet.*
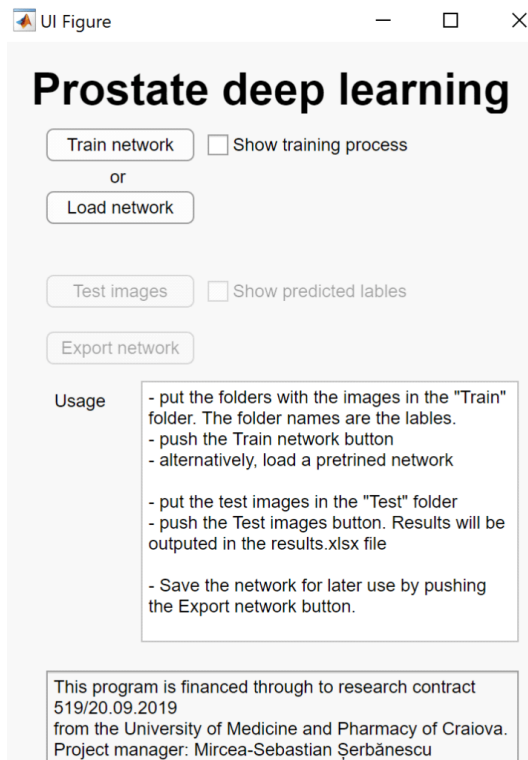
Figure 4 – *Standalone prostate cancer image classifier application interface.*

A common observed problem of the GGS scoring is that score 7 (grades 4+3) cancers were associated with a three-fold increase in prostate cancer outcome compared with grades 3+4 cancers [95% confidence interval (CI), 1.1 to 8.6) [33]. The same conclusion was reported by Chan *et al.* (2000) [34]; the authors have concluded that Gleason score 4+3 tumors had an increased risk of progression independent of stage and margin status ($p<0.0001$). They also reported that the 5-year actuarial risk of progression was 15% for Gleason score 3+4 and 40% for 4+3 tumors. A close look to the data in Figure 4 shows a higher error rate of the classification exactly between pattern 3 and pattern 4. Thus, AlexNet incorrectly labeled five images with pattern 4 as pattern 3 and eight images with pattern 3 as pattern 4. GoogleNet incorrectly labeled nine images with pattern 4 as pattern 3, and 10 images with pattern 3 as pattern 4. This represents 25% of all the errors of the AlexNet classifier and 40% of GoogleNet. This can be explained on one hand by the fact the classifier was trained on images labeled by pathologists and is thus subjected to their subjectivity and, on the other hand, it could show a problem related to the GGS itself as reports show higher interobserver variability between these classes. In a similar report [35], the 24 cases that had score changes, five cases were upgraded from grade 3 to 4 and 15 were downgraded from grade 4 to 3, this representing 80% of the reported changes.

This study has limitations that would need to be addressed to prior to clinical usage, and these will need future work and improvements. However, based on its performance, the resulted application could be used for research purposes.

First, the specimens that were used to develop the DL algorithms originated from a single medical center, the slides were stained in one laboratory, using only the HE staining, the digital data were acquired using a single slide scanner, and, the classes are unbalanced. Only after the results would be confirmed on multi-center studies, with different staining protocols, the images would be digitalized using slide scanners from different vendors and the dataset would be large enough and balanced, then the proposed method could be considered for clinical deployment.

Second, the study focuses on classifying images of acinar prostatic adenocarcinoma excluding other types of prostate cancer or invading nearby tumors, and there were no normal glands available as reference. Used incorrectly, it could generate miss classification and miss understanding.

Last, a more serious limitation is the subjective nature of the GGS. Inter-pathologist variability is a non-negligible aspect, as also shown in other studies [18], and it can be overcome considering two different approaches. One possible approach would be to replace the "gold standard" classification – GGS – with a simpler and more objective one. This is unlikely due to the large usage of the GGS in practice. From our experience [36, 37], a possible alternative could be the Srigley grading system, which, at least, has more clearly defined classes. Another possible approach could be the use of a large trading set, thus reducing the error probability.

We conducted our study on resection specimens; however, CAD on needle core biopsies would have a higher clinical impact. Since there is no difference on the GGS training and functionality for core biopsies [19], we expect a constant behavior for the classifiers, but the theory remains to be proved in further work.

Different from other approaches, this study describes the transfer learning from general-purpose DL networks to a diagnostic system of prostate cancer grading through GGS using routine histopathology images. The technique has been successfully used on ultrasound and MRI images [38–41].

⊟ Conclusions

In this paper, we present two DL algorithms design to classify images according to GGS. The algorithms use transfer learning from two well-known general-purpose image classification DL networks – AlexNet and GoogleNet –, and are further trained on histopathology images of prostate cancer. With a reported ACC of 61.17±7 for AlexNet and of 60.9±7.4 for GoogleNet, with a small dataset of only 439 asymmetrically distributed cases in four GGS classes, we find the result to be promising. The similar results obtained by the two networks with very different architecture, together with the normal distribution of classification error for both algorithms show that we have reached a maximum classification rate on this dataset. With further evaluation, the resulted networks could assist pathologists by presenting an objective first or second opinion in a grading system with high interobserver and intraobserver variability.

Conflict of interests

The authors declare that they have no conflict of interests.

Authors' contribution

Mircea-Sebastian Şerbănescu and Nicolae Cătălin Manea contributed equally to this article and share main authorship.

References

[1] ***. Cancer stat facts: prostate cancer – 2020. Surveillance, Epidemiology, and End Results Program [online], National Cancer Institute (NIH), available at: https://seer.cancer.gov/statfacts/html/prost.html, accessed: December 27, 2019.

[2] Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL, ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol, 2005, 29(9):1228–1242. https://doi.org/10.1097/01.pas.0000173646.99337.b1 PMID: 16096414

[3] Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, Grading Committee. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol, 2016, 40(2):244–252. https://doi.org/10.1097/PAS.0000000000000530 PMID: 26492179

[4] National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines in Oncology. Prostate cancer – 2020. NCCN Evidence-Based Cancer Guidelines, Oncology Drug Compendium, Oncology Continuing Medical Education [online], available at: https://www.nccn.org/professionals/physician_gls/default.aspx, accessed: December 27, 2019.

[5] Persson J, Wilderäng U, Jiborn T, Wiklund PN, Damber JE, Hugosson J, Steineck G, Haglind E, Bjartell A. Interobserver variability in the pathological assessment of radical prostatectomy specimens: findings of the Laparoscopic Prostatectomy Robot Open (LAPPRO) Study. Scand J Urol, 2014, 48(2):160–167. https://doi.org/10.3109/21681805.2013.820788 PMID: 23906418

[6] Veloso SG, Lima MF, Salles PG, Berenstein CK, Scalon JD, Bambirra EA. Interobserver agreement of Gleason score and modified Gleason score in needle biopsy and in surgical specimen of prostate cancer. Int Braz J Urol, 2007, 33(5):639–646; discussion 647–651. https://doi.org/10.1590/s1677-55382007000500005 PMID: 17980061

[7] Netto GJ, Eisenberger M, Epstein JI, TAX 3501 Trial Investigators. Interobserver variability in histologic evaluation of radical prostatectomy between central and local pathologists: findings of TAX 3501 Multinational Clinical Trial. Urology, 2011, 77(5):1155–1160. https://doi.org/10.1016/j.urology.2010.08.031 PMID: 21146858 PMCID: PMC3449146

[8] Allsbrook WC Jr, Mangold KA, Johnson MH, Lane RB, Lane CG, Amin MB, Bostwick DG, Humphrey PA, Jones EC, Reuter VE, Sakr W, Sesterhenn IA, Troncoso P, Wheeler TM, Epstein JI. Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. Hum Pathol, 2001, 32(1):74–80. https://doi.org/10.1053/hupa.2001.21134 PMID: 11172298

[9] Allsbrook WC Jr, Mangold KA, Johnson MH, Lane RB, Lane CG, Epstein JI. Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologist. Hum Pathol, 2001, 32(1):81–88. https://doi.org/10.1053/hupa.2001.21135 PMID: 11172299

[10] Mikami Y, Manabe T, Epstein JI, Shiraishi T, Furusato M, Tsuzuki T, Matsuno Y, Sasano H. Accuracy of Gleason grading by practicing pathologists and the impact of education on improving agreement. Hum Pathol, 2003, 34(7):658–665. https://doi.org/10.1016/s0046-8177(03)00191-6 PMID: 12874761

[11] Serbanescu MS, Plesea IE. R-VA a new fractal parameter for grayscale image characterization. Ann Comput Sci Ser, 2015, 13(1):9–14.

[12] Serbanescu MS, Plesea RM, Pop OT, Bungardean C, Plesea IE. SY14.04/Image Analysis III: Fractal behavior of Gleason and Srigley grading systems. Proceeding of the 13th European Congress on Digital Pathology, May 25–28, 2016, Berlin, Germany, Diagn Pathol, 2016, 1(8):145. https://doi.org/10.17629/www.diagnosticpathology.eu-2016-8:145

[13] Serbanescu MS, Plesea RM, Covica V, Plesea IE. PS-18-003: Fractal dimension of stromal fibrillar network: a new approach to prostate carcinoma architectural assessment. Proceedings of the 27th European Congress of Pathology, September 5–9, 2015, Belgrade, Serbia, Virchows Arch, 2015, 467(1 Suppl):S235. https://doi.org/10.1007/s00428-015-1805-9

[14] Leo P, Elliott R, Shih NNC, Gupta S, Feldman M, Madabhushi A. Stable and discriminating features are predictive of cancer presence and Gleason grade in radical prostatectomy specimens: a multi-site study. Sci Rep, 2018, 8(1):14918. https://doi.org/10.1038/s41598-018-33026-5 PMID: 30297720

[15] Sparks R, Madabhushi A. Statistical shape model for manifold regularization: Gleason grading of prostate histology. Comput Vis Image Underst, 2013, 117(9):1138–1146. https://doi.org/10.1016/j.cviu.2012.11.011 PMID: 23888106 PMCID: PMC3718190

[16] Nguyen K, Jain AK, Allen RL. Automated gland segmentation and classification for Gleason grading of prostate tissue images. In: ***. Proceedings of the 20th International Conference on Pattern Recognition (ICPR), August 23–26, 2010, Istanbul, Turkey. Institute of Electrical and Electronics Engineers (IEEE) Computer Society, Los Alamitos, California, USA, 2010, 1497.

[17] Gummeson A, Arvidsson I, Ohlsson M, Overgaard NC, Krzyzanowska A, Heyden A, Bjartell A, Aström K. Automatic Gleason grading of H and E stained microscopic prostate images using deep convolutional neural networks. In: Gurcan MN, Tomaszewski JE (eds). Medical imaging 2017: digital pathology. The International Society for Optical Engineering (SPIE) Proceedings, vol. 10140, Medical Imaging, Orlando, USA, 2017, 1–286. https://doi.org/10.1117/12.2253620

[18] Arvaniti E, Fricker KS, Moret M, Rupp N, Hermanns T, Fankhauser C, Wey N, Wild PJ, Rüschoff JH, Claassen M. Automated Gleason grading of prostate cancer tissue microarrays *via* deep learning. Sci Rep, 2018, 8(1):12054. https://doi.org/10.1038/s41598-018-30535-1 PMID: 30104757 PMCID: PMC6089889

[19] Lucas M, Jansen I, Savci-Heijink CD, Meijer SL, de Boer OJ, van Leeuwen TG, de Bruin DM, Marquering HA. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. Virchows Arch, 2019, 475(1):77–83. https://doi.org/10.1007/s00428-019-02577-x PMID: 31098801 PMCID: PMC6611751

[20] Gleason DF. Classification of prostatic carcinomas. Cancer Chemother Rep, 1966, 50(3):125–128. PMID: 5948714

[21] Gleason DF; The Veterans Administration Cooperative Urological Research Group. Chapter 9: Histologic grading and clinical staging of prostatic carcinoma. In: Tannenbaum M (ed). Urologic pathology: the prostate. Lea & Febiger, Philadelphia, USA, 1977, 171–197.

[22] Gleason DF. Histologic grading of prostate cancer: a perspective. Hum Pathol, 1992, 23(3):273–279. https://doi.org/10.1016/0046-8177(92)90108-f PMID: 1555838

[23] Gleason DF, Mellinger GT; The Veterans Administration Cooperative Urological Research Group. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol, 1974, 111(1):58–64. https://doi.org/10.1016/s0022-5347(17)59889-4 PMID: 4813554

[24] Pleşea RM, Şerbănescu MS, Ciovică DV, Roşu GC, Moldovan VT, Bungărdean RM, Popescu NA, Pleşea IE. The study of tumor architecture components in prostate adenocarcinoma using fractal dimension analysis. Rom J Morphol Embryol, 2019, 60(2):501–519. PMID: 31658324

[25] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the Association for Computing Machinery (ACM), 2017, 60(6):84–90. https://doi.org/10.1145/3065386

[26] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. arXiv, 2014, arXiv:1409.4842.

[27] Gheonea DI, Streba CT, Vere CC, Şerbănescu M, Pirici D, Comănescu M, Streba LAM, Ciurea ME, Mogoantă S, Rogoveanu I. Diagnosis system for hepatocellular carcinoma based on fractal dimension of morphometric elements integrated

in an artificial neural network. BioMed Res Int, 2014, 2014: 239706. https://doi.org/10.1155/2014/239706 PMID: 25025042 PMCID: PMC4084678

[28] Ştefănescu D, Streba C, Cârţână ET, Săftoiu A, Gruionu G, Gruionu LG. Computer aided diagnosis for confocal laser endomicroscopy in advanced colorectal adenocarcinoma. PLoS One, 2016, 11(5):e0154863. https://doi.org/10.1371/journal.pone.0154863 PMID: 27144985 PMCID: PMC4856267

[29] Stoiculescu A, Pleşea IE, Pop OT, Alexandru DO, Man M, Şerbănescu M, Pleşea RM. Correlations between intratumoral interstitial fibrillary network and tumoral architecture in prostatic adenocarcinoma. Rom J Morphol Embryol, 2012, 53(4):941–950. PMID: 23303017

[30] Pleşea IE, Stoiculescu A, Şerbănescu M, Alexandru DO, Man M, Pop OT, Pleşea RM. Correlations between intratumoral vascular network and tumoral architecture in prostatic adeno-carcinoma. Rom J Morphol Embryol, 2013, 54(2):299–308. PMID: 23771073

[31] Streba L, Forţofoiu MC, Popa C, Ciobanu D, Gruia CL, Mogoantă SŞ, Streba CT. A pilot study on the role of fractal analysis in the microscopic evaluation of colorectal cancers. Rom J Morphol Embryol, 2015, 56(1):191–196. PMID: 25826505

[32] Madabhushi A, Feldman MD, Leo P. Deep-learning approaches for Gleason grading of prostate biopsies. Lancet Oncol, 2020, 21(2):187–189. https://doi.org/10.1016/S1470-2045(19)30793-4 PMID: 31926804

[33] Stark JR, Perner S, Stampfer MS, Sinnott JA, Finn S, Eisenstein AS, Ma J, Fiorentino M, Kurth T, Loda M, Giovannucci EL, Rubin MA, Mucci LA. Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3? J Clin Oncol, 2009, 27(21):3459–3464. https://doi.org/10.1200/JCO.2008.20.4669 PMID: 19433685 PMCID: PMC2717753

[34] Chan TY, Partin AW, Walsh PC, Epstein JI. Prognostic significance of Gleason score 3+4 *versus* Gleason score 4+3 tumor at radical prostatectomy. Urology, 2000, 56(5):823–827. https://doi.org/10.1016/s0090-4295(00)00753-6 PMID: 11068310

[35] Salmo EN. An audit of inter-observer variability in Gleason grading of prostate cancer biopsies: the experience of central pathology review in the North West of England. Integr Cancer Sci Ther, 2015, 2(2):104–106. https://doi.org/10.15761/ICST.1000123

[36] Şerbănescu MS, Pleşea RM, Ciovică V, Moldovan VT, Ştefan A, Grădinaru A, Leavitt MO, Pleşea IE. Clinical Poster 18: Assessment of MMP9 and MMP2 expression in prostate carcinoma in correlation with Gleason and Srigley grading systems. In: ***. Abstracts of the 15th European Congress on Digital Pathology (ECDP), April 10–13, 2019, University of Warwick, Coventry, UK, 2019, 55.

[37] Mitroi G, Pleşea RM, Pop OT, Ciovică DV, Şerbănescu MS, Alexandru DO, Stoiculescu A, Pleşea IE. Correlations between intratumoral interstitial fibrillary network and vascular network in Srigley patterns of prostate adenocarcinoma. Rom J Morphol Embryol, 2015, 56(4):1319–1328. PMID: 26743277

[38] Yuan Y, Qin W, Buyyounouski M, Ibragimov B, Hancock S, Han B, Xing L. Prostate cancer classification with multipara-metric MRI transfer learning model. Med Phys, 2019, 46(2): 756–765. https://doi.org/10.1002/mp.13367 PMID: 30597561

[39] Aljundi R, Lehaire J, Prost-Boucle F, Rouvière O, Lartizien C. Transfer learning for prostate cancer mapping based on multi-centric MR imaging databases. In: Bhatia KK, Lombaert H (eds). Machine learning meets medical imaging (MLMMI): First International Workshop, MLMMI 2015, held in conjunction with 32nd International Conference on Machine Learning (ICML) 2015, July 11, 2015, Lille, France, Revised Selected Papers. Lecture Notes in Computer Science, vol. 9487, Springer, Cham, 2015, 74–82. https://doi.org/10.1007/978-3-319-27929-9_8

[40] Azizi S, Mousavi P, Yan P, Tahmasebi A, Kwak JT, Xu S, Turkbey B, Choyke P, Pinto P, Wood B, Abolmaesumi P. Transfer learning from RF to B-mode temporal enhanced ultrasound features for prostate cancer detection. Int J Comput Assist Radiol Surg, 2017, 12(7):1111–1121. https://doi.org/10.1007/s11548-017-1573-x PMID: 28349507

[41] Chen Q, Hu S, Long P, Lu F, Shi Y, Li Y. A transfer learning approach for malignant prostate lesion detection on multi-parametric MRI. Technol Cancer Res Treat, 2019, 18:15330 33819858363. https://doi.org/10.1177/1533033819858363 PMID: 31221034 PMCID: PMC6589968

Corresponding author
Liliana Streba, Teaching Assistant, MD, PhD, Department of Medical Oncology, University of Medicine and Pharmacy of Craiova, 2 Petru Rareş Street, 200349 Craiova, Romania; Phone +40770–220 398, e-mail: liliana.streba@umfcv.ro