

The assessment of Computer Vision Algorithms for the Diagnosis of Prostatic Adenocarcinoma in Surgical Specimens

Running Title: Deep Learning and Prostatic Adenocarcinoma

AUTHORS:

1. Syed Usama Khalid Bukhari
 - a. usama.khalid@cs.uol.edu.pk
 - b. ORCID ID: <https://orcid.org/0000-0003-1581-8609>
 - c. Assistant Professor, Department of Computer Science, The University of Lahore, Islamabad - Pakistan
2. Ubeer Mehtab
 - a. Ubeermehtab456@gmail.com
 - b. House officer, Bahawal Victoria Hospital, Quaid e Azam Medical College, Bahawalpur - Pakistan
3. Syed Shahzad Hussain
 - a. syedshahzadhussain@nutech.edu.pk
 - b. Lecturer, Electrical Engineering Department. National University of Technology (NUTECH), Islamabad, Pakistan.
4. Asmara Syed
 - a. syed_asmara@yahoo.com
 - b. ORCID ID: <http://orcid.org/0000-0002-8895-9069>
 - c. Lecturer, Department of Pathology, Faculty of Medicine, Northern Border University, Arar- Kingdom of Saudi Arabia
5. Syed Umar Armaghan
 - a. Syedumar2345@gmail.com
 - b. Student, Biomedical Engineering, Riphah International University. Islamabad - Pakistan
6. Syed Sajid Hussain Shah
 - a. prof.sajid99@gmail.com
 - b. ORCID ID: <https://orcid.org/0000-0003-3425-6293>
 - c. Professor, Department of Pathology, Faculty of Medicine, Northern Border University, Arar- Kingdom of Saudi Arabia

Corresponding Author:

1. Syed Usama Khalid Bukhari
 - a. usama.khalid@cs.uol.edu.pk
 - b. ORCID ID: <https://orcid.org/0000-0003-1581-8609>
 - c. Assistant Professor, Department of Computer Science, The University of Lahore, Islamabad, Pakistan.

The assessment of Deep Learning Computer Vision Algorithms for the Diagnosis of Prostatic Adenocarcinoma in Surgical Specimens

Abstract:

Introduction: Prostatic malignancy is a major cause of morbidity and fatality among men around the globe. More than a million new cases of prostatic cancer are diagnoses annually. The incidence of prostatic malignancy is rising and it is expected that more than two million new cases of prostatic carcinoma will be diagnosed in 2040. The application of machine learning to assist the histopathologists could be a very valuable adjunct tool for the histological diagnosis of prostatic malignant tumors.

Aim & Objectives: To evaluate the effectiveness of artificial intelligence for the histopathological diagnosis of prostatic carcinoma by analyzing the digitized pathology slides.

Materials & Methods: Eight hundred and two (802) images in total, were obtained from the anonymised slides stained with hematoxylin and eosin which included anonymised 337 images of prostatic adenocarcinoma and 465 anonymised images of nodular hyperplasia of prostate. Eighty percent (80%) of the total digital images were used for training and 20% for testing. Three ResNet architectures ResNet-18, ResNet-34, and ResNet-50 were employed for the analysis of these images.

Results: The evaluation of digital images by ResNet-18, ResNet-34, and ResNet-50 revealed the diagnostic accuracy of 97.1%, 98 % and 99.5 % respectively.

Discussion: The application of artificial intelligence is being considered as a very useful tool which may improve the patient care by improving the diagnostic accuracy and reducing the cost. In radiology, the application of deep learning to interpret radiological images has revealed excellent results. In the present study, the analysis of pathology images by convolutional neural network architecture revealed the diagnostic accuracy of 97.1%, 98 % and 99.5 % with by ResNet-18, ResNet-34, and ResNet-50 respectively. The findings of the present study are in accordance with the other published series, which were carried out to determine the accuracy of machine learning for the diagnosis of cancers of lung, breast and prostate. The application of deep learning for the histological diagnosis of malignant tumors could be quite helpful in improving the patient care.

Conclusion: The findings of the present study suggest that intelligent vision system possibly a worthwhile tool for the histopathological evaluation of prostatic tissue to differentiate between the benign and malignant disorders.

Key Words: Computer Vision, Deep Learning, Prostatic Carcinoma, ResNet.

Introduction:

Prostatic carcinoma is the second most prevalent type of malignant tumor diagnosed in male all over the world (1). In 2018, more than one million (1,276,106) patients have been diagnosed as a case of prostatic cancer around the globe and this malignant tumor caused 358,989 deaths during that time period (1).

Variation in the prevalence of cancers exists among the different geographic regions (2). Some tumors are more prevalent in one region and less common in other. The prostatic cancer is the most frequent malignant tumor of male population in one hundred and five countries which mainly include the countries from the continents of Europe, Africa, North & South America (3). The geographic variation in the prevalence of prostatic cancer has been attributed to genetic and environmental factors. Higher consumption of meat has been linked to higher risk of development of prostatic carcinoma (4-5).

The incidence of prostatic malignancy rises with increase in the age. The prostatic malignancy is more common in 7th decade with mean age of sixty-six years (3). An increased incidence of prostatic adenocarcinoma has been observed in the younger population. This rising trend in the incidence of cancer of prostate in the younger age group has been attributed to improved diagnostic techniques and environmental factors including the rising incidence of obesity among the adolescent and younger age groups (6).

The prostatic carcinoma causes urinary tract symptoms due to enlargement of prostate. The spread of prostatic adenocarcinoma to vertebral column causes backache. The common symptoms with which these patients presents usually includes urinary obstruction, urgency, hesitancy, discomfort, frequency and hematuria (7-9). The suspected cases of prostatic malignancy are tested for prostate specific antigen. There is an increased suspicion of prostatic carcinoma in case of elevated blood PSA level. But the definitive diagnosis of prostatic malignant tumor is confirmed by the histopathological examination of tissue specimen from the prostate. The rising number of prostatic biopsies for the histopathological examination are going to increase the stress on the histopathologists in future as more than two million cases of prostatic cancer are expected by year 2040 (3).

The application of artificial intelligence and development of Computer aided diagnostic programs may be instrumental in enhancing the laboratory efficiency by reducing the turnaround time, cost and chance of human errors. Artificial intelligence is yielding very good results in the interpretation of digital images by analyzing them with Deep learning architecture.

The main focus of this study is to assess the usefulness of machine learning for the histological diagnosis of prostatic adenocarcinoma.

Materials & Methods:

The research project has been approved from the ethical committee of University of Lahore - Islamabad Campus. Eight hundred and two (802) digital anonymized images were collected, which included 337 images of prostatic adenocarcinoma and 465 images of prostatic hyperplasia. We have used three ResNet convolutional neural network architectures, ResNet -18, ResNet -34 and ResNet-50. Because of dataset limitation, the transfer learning approach has been employed to train the models using the ImageNet data set (10).

A total of 337 images of prostatic adenocarcinoma and 465 images of nodular hyperplasia of prostate are used in the present research project. These images have been reviewed and labeled by two histopathologists.

A random approach of FastAI API is used to load the data into train and test sets, to avoid the chances of images of the same class been selected for a mini-batch input, that can harm the performance of model validation.

Two sets (training set and test set) are made from the whole dataset. The training set is comprised 80% of total dataset, while the Testing set is 20% of total dataset.

For image, we used data augmentation for regulation.

Results:

A total of eight hundred and two (802) images were collected. Three hundred and thirty-seven digital images (337) of adenocarcinoma of prostate and four hundred and sixty-five (465) images of prostatic hyperplasia. Six hundred and forty (640) images randomly chosen from both carcinoma and prostatic hyperplasia images that is 80% of the total data set, and employed for training data set. Where one hundred and sixty-two (162) images randomly chosen from both carcinoma and hyperplasia images that is 20% of the total data set employed in the test data set.

Three ResNet architecture models, ResNet-18, ResNet-34 and ResNet-50 are used to classify the input image. Accuracy of 97.1%, 98 % and 99.5 % respectively was obtained on test data set.

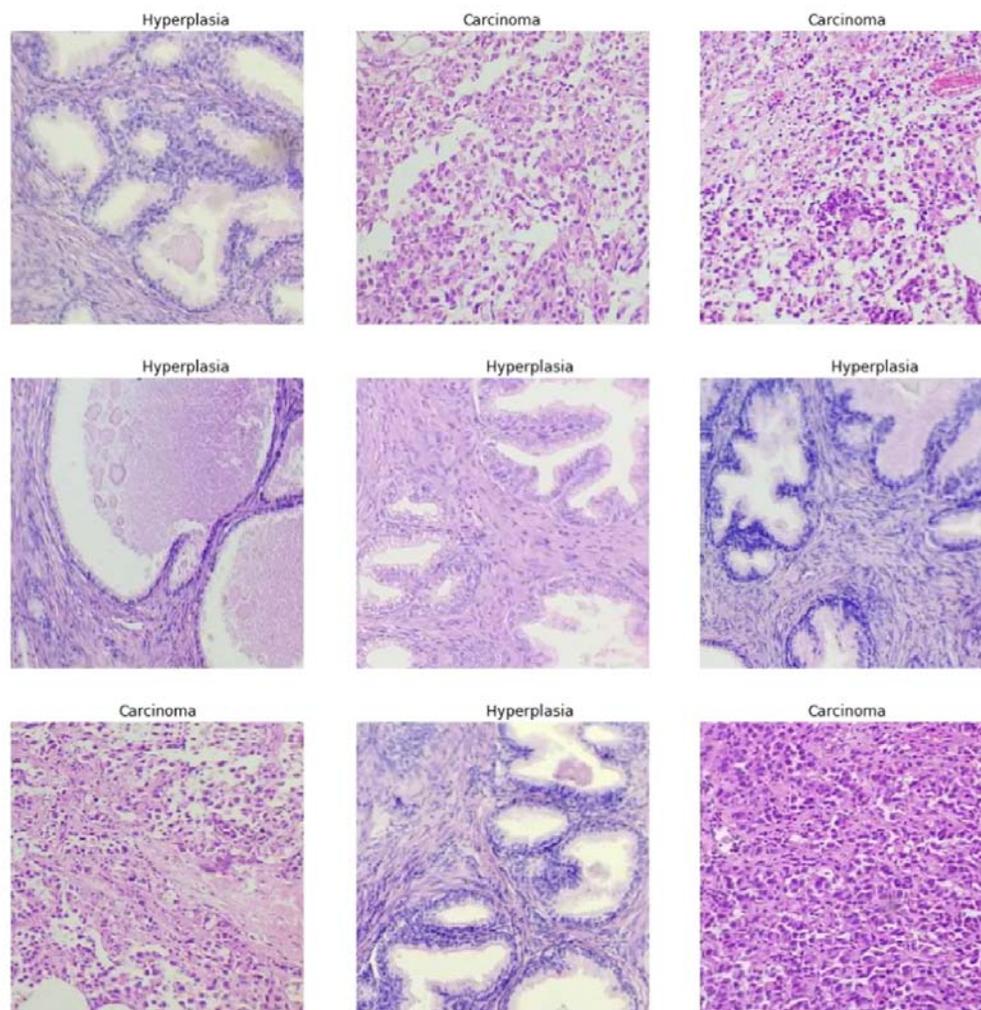


Figure 1: Example of Data and labels loaded using FastAI (randomly)

Table 1 Confusion matrix of train and test datasets

Architectures	Train Data Set	Test Data Set																														
ResNet-18	<p>Confusion matrix</p> <table border="1"> <tr> <td rowspan="2">Actual</td> <td>Carcinoma</td> <td>273</td> <td>6</td> </tr> <tr> <td>Hyperplasia</td> <td>14</td> <td>347</td> </tr> <tr> <td></td> <td></td> <td>Carcinoma</td> <td>Hyperplasia</td> </tr> <tr> <td></td> <td></td> <td colspan="2">Predicted</td> </tr> </table>	Actual	Carcinoma	273	6	Hyperplasia	14	347			Carcinoma	Hyperplasia			Predicted		<p>Confusion matrix</p> <table border="1"> <tr> <td rowspan="2">Actual</td> <td>Carcinoma</td> <td>58</td> <td>1</td> </tr> <tr> <td>Hyperplasia</td> <td>2</td> <td>101</td> </tr> <tr> <td></td> <td></td> <td>Carcinoma</td> <td>Hyperplasia</td> </tr> <tr> <td></td> <td></td> <td colspan="2">Predicted</td> </tr> </table>	Actual	Carcinoma	58	1	Hyperplasia	2	101			Carcinoma	Hyperplasia			Predicted	
Actual	Carcinoma		273	6																												
	Hyperplasia	14	347																													
		Carcinoma	Hyperplasia																													
		Predicted																														
Actual	Carcinoma	58	1																													
	Hyperplasia	2	101																													
		Carcinoma	Hyperplasia																													
		Predicted																														
ResNet-34	<p>Confusion matrix</p> <table border="1"> <tr> <td rowspan="2">Actual</td> <td>Carcinoma</td> <td>268</td> <td>3</td> </tr> <tr> <td>Hyperplasia</td> <td>13</td> <td>356</td> </tr> <tr> <td></td> <td></td> <td>Carcinoma</td> <td>Hyperplasia</td> </tr> <tr> <td></td> <td></td> <td colspan="2">Predicted</td> </tr> </table>	Actual	Carcinoma	268	3	Hyperplasia	13	356			Carcinoma	Hyperplasia			Predicted		<p>Confusion matrix</p> <table border="1"> <tr> <td rowspan="2">Actual</td> <td>Carcinoma</td> <td>64</td> <td>0</td> </tr> <tr> <td>Hyperplasia</td> <td>0</td> <td>98</td> </tr> <tr> <td></td> <td></td> <td>Carcinoma</td> <td>Hyperplasia</td> </tr> <tr> <td></td> <td></td> <td colspan="2">Predicted</td> </tr> </table>	Actual	Carcinoma	64	0	Hyperplasia	0	98			Carcinoma	Hyperplasia			Predicted	
Actual	Carcinoma		268	3																												
	Hyperplasia	13	356																													
		Carcinoma	Hyperplasia																													
		Predicted																														
Actual	Carcinoma	64	0																													
	Hyperplasia	0	98																													
		Carcinoma	Hyperplasia																													
		Predicted																														
ResNet-50	<p>Confusion matrix</p> <table border="1"> <tr> <td rowspan="2">Actual</td> <td>Carcinoma</td> <td>262</td> <td>3</td> </tr> <tr> <td>Hyperplasia</td> <td>1</td> <td>374</td> </tr> <tr> <td></td> <td></td> <td>Carcinoma</td> <td>Hyperplasia</td> </tr> <tr> <td></td> <td></td> <td colspan="2">Predicted</td> </tr> </table>	Actual	Carcinoma	262	3	Hyperplasia	1	374			Carcinoma	Hyperplasia			Predicted		<p>Confusion matrix</p> <table border="1"> <tr> <td rowspan="2">Actual</td> <td>Carcinoma</td> <td>72</td> <td>0</td> </tr> <tr> <td>Hyperplasia</td> <td>0</td> <td>90</td> </tr> <tr> <td></td> <td></td> <td>Carcinoma</td> <td>Hyperplasia</td> </tr> <tr> <td></td> <td></td> <td colspan="2">Predicted</td> </tr> </table>	Actual	Carcinoma	72	0	Hyperplasia	0	90			Carcinoma	Hyperplasia			Predicted	
Actual	Carcinoma		262	3																												
	Hyperplasia	1	374																													
		Carcinoma	Hyperplasia																													
		Predicted																														
Actual	Carcinoma	72	0																													
	Hyperplasia	0	90																													
		Carcinoma	Hyperplasia																													
		Predicted																														

Table 2 Detail results Test data set

ARCHITECTURES	PRECISION	RECALL	F-1 SCORE	ACCURACY
RESNET-18	0.966	0.983	0.974	97.1%
RESNET-34	1.0	1.0	1.0	98.0 %
RESNET-50	1.0	1.0	1.0	99.5 %

Discussion:

Adenocarcinoma of prostate is a quite prevalent tumor in the majority of the countries of the world. More than one million cases of prostatic cancer are diagnosed every year around the globe. Approximately two million persons may become victim of prostatic malignancy every year by year 2040 (3). The rising number of prostatic biopsies for the purpose of histopathological diagnosis is going to mount the burden on the diagnostic laboratories and will increase the stress on the histopathologists. In order to overcome this situation, the application of smart technology such as artificial intelligence may be very helpful.

The rapid advancement in field of artificial intelligence in the recent past has drawn the attention of health care providers to apply this technology for the solution of complex medical problems.

The artificial intelligence is characterized by development of machines which have the capability to do cognitive assignments for the achievement of desirable output based on given input. The application of artificial intelligence has been expected to revolutionize the patient care by improving the diagnostic accuracy and reducing the cost with better choices for the management of diseases (11). An important sub group of artificial intelligence is Deep learning. The application of deep learning revealed very encouraging results in the field of radiology for image interpretation. A study conducted by Pasa F et al showed very encouraging results of Convolutional neural network (CNN) for the image analysis of the X-rays. They used convolutional neural network to diagnose tuberculosis (12).

The histopathological diagnosis of malignant tumors is based upon certain specific microscopic features such as pleomorphism, hyperchromasia, high nucleus to cytoplasmic ratio, atypical mitosis, necrosis, multinucleated giant cells and pattern of arrangement of neoplastic cells. The assessment of digital pathology image with machine learning may provide valuable assistance in the interpretation and differentiation of neoplastic and non-neoplastic lesions.

The convolutional neural network has been applied in this research project to diagnose prostatic adenocarcinoma by analyzing the digital pathology images which revealed excellent results. With the application of ResNet-50 architecture, a

diagnostic accuracy of 99.5 % has been achieved while ResNet-34 and ResNet-18 yielded diagnostic accuracy of 98 % and 97.1%, respectively. The findings of the present study are in accordance with the other published series which were carried out to assess the accuracy of deep learning for the diagnosis of cancers of lung, breast and prostate (13-16).

The histopathological assessment is quite subjective matter which may be a possible reason for the inter-observer discrepancies. The advancement in the computational image technology may be able to make the histopathology more objective and quantitative (17).

The application of computer based diagnostic system in histopathology will be a great help to the histopathologists in the current scenario of rising trend in the incidence of neoplastic lesions leading to increase in the number biopsies.

Conclusion: The findings of the present research project raises the potential application of computer vision-based systems as a valuable tool for the histopathological evaluation of prostatic tissue in differentiating the benign and malignant disorders. The use of intelligent vision system in the field of histopathology for the diagnosis of malignant disorders may also, lessen the risk of human errors and it will decrease the workload on histopathologists along with provision of a second opinion to them.

Acknowledgement:

The authors are grateful to Farhan Mazher for his valuable assistance.

Funding: None

Conflict of interest: None

References:

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
2. Idrees R, Fatima S, Abdul-Ghafar J, Raheem A, Ahmad Z. Cancer prevalence in Pakistan: meta-analysis of various published studies to determine variation in cancer figures resulting from marked population heterogeneity in different parts of the country. *World Journal of Surgical Oncology* (2018) 16:129. <https://doi.org/10.1186/s12957-018-1429-z>
3. Rawla P. Epidemiology of Prostate Cancer. *World J Oncol*. 2019;10(2):63-89. doi:10.14740/wjon1191.
4. Gibson TM, Ferrucci LM, Tangrea JA, Schatzkin A. Epidemiological and clinical studies of nutrition. *Semin Oncol*. 2010;37(3):282-296.
5. Rohrmann S, Platz EA, Kavanaugh CJ, Thuita L, Hoffman SC, Helzlsouer KJ. Meat and dairy consumption and subsequent risk of prostate cancer in a US cohort study. *Cancer Causes Control*. 2007;18(1):41-50.
6. Bleyer A, Spreafico F, Barr R. Prostate cancer in young men: An emerging young adult and older adolescent challenge. *Cancer* 2020;126:46-57. <https://doi.org/10.1002/ncr.32498>
7. Estebanez J, Teyrouz A, Gutierrez MA, Linazasoro I, Belloso J, Cano C, et al. Natural history of prostate cancer. *Arch Esp Urol*. 2014;67(5):383-387.
8. Hamilton W, Sharp D. Symptomatic diagnosis of prostate cancer in primary care: a structured review. *Br J Gen Pract*. 2004;54(505):617-621.
9. Tabassum T, Imtiaz A, Jabeen R. A Study to Evaluate α -Methylacyl Co-A Racemase Expression in Hyperplasia and Different Grades of Adenocarcinoma of Prostate. *Pakistan Journal of Medical and Health Sciences*. 2017;11(3):909-914.
10. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis*. 2015: 115:211-252. <http://doi.org/10.1007/s11263-015-0816-y>.
11. Goldenberg S, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. *Nat Rev Urol* 16, 391-403 (2019). <https://doi.org/10.1038/s41585-019-0193-3>
12. Pasa F, Golkov V, Pfeiffer F, Cremers D, Pfeiffer D. Efficient deep network architectures for fast chest X-ray tuberculosis screening and visualization. *Scientific Reports*. 2019; 9:6268. DOI: 10.1038/s41598-019-42557-4.
13. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 2017;318(22):2199-2210. doi: 10.1001/jama.2017.14585
14. Campanella G, Hanna MG, Geneslaw L, Mirafior A, Silva VWK, Busam KJ et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 2019; 25: 1301-1309. doi: 10.1038/s41591-019-0508-1.
15. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018; 24: 1559-67. doi: 10.1038/s41591-018-0177-5.
16. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep*. 2016;6:26286. doi: 10.1038/srep26286.

17. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Med Image Anal.* 2016; 33:170–175. doi:10.1016/j.media.2016.06.037