A Comparison of Deep Reinforcement Learning and Deep learning for Complex Image Analysis

Rishi Khajuria[1*], Abdul Quyoom[1], Abid Sarwar[1]

Abstract

The image analysis is an important and predominant task for classifying the different parts of the image. The analysis of complex image analysis like histopathological define a crucial factor in oncology due to its ability to help pathologists for interpretation of images and therefore various feature extraction techniques have been evolved from time to time for such analysis. Although deep reinforcement learning is a new and emerging technique but very less effort has been made to compare the deep learning and deep reinforcement learning for image analysis. The paper highlights how both techniques differ in feature extraction from complex images and discusses the potential pros and cons. The use of Convolution Neural Network (CNN) in image segmentation, detection and diagnosis of tumour, feature extraction is important but there are several challenges that need to be overcome before Deep Learning can be applied to digital pathology. The one being is the availability of sufficient   training examples for medical image datasets, feature extraction from whole area of the image, ground truth localized annotations, adversarial effects of input representations and extremely large size of the digital pathological slides (in gigabytes).Even though formulating   Histopathological Image Analysis (HIA) as   Multi Instance Learning   (MIL) problem is a remarkable step where histopathological image is divided into high resolution patches to make predictions for the patch and then combining them for overall slide predictions but it suffers from loss of contextual and spatial information. In such cases the deep reinforcement learning techniques can be used to learn feature from the limited data without losing contextual and spatial information.

Key Words: Deep reinforcement learning, Deep learning, Complex images, CNN, DQN.

I. Introduction

Reinforcement learning [1], is a sub domain of artificial intelligence that allows agent to fulfil a given goal while maximizing a numerical reward signal. It was developed to deal with three main strategies. The first is the concept of learning by trial and error [2], discovered during researches undertaken in psychology and neuroscience of animal learning. The second concept is the problem of optimal control developed in the 1950s using a discrete stochastic version of the environment known as Markovian decision processes [3] and adopting a concept of a dynamical system's state and optimal return function (Reward) and defining the "Bellman equation" [4] to optimize the agent behaviour over the time (Dynamic programming [5]). The last concept concerns the temporal-difference methods [6], which become the mainstream, and was boosted by the actor-critic [7] architecture.

Deep learning [8] is a machine learning sub-domain, based on the concept of artificial neural networks that works similar to human brain for processing data and create patterns to be used in decision making. Deep learning allows automatic feature engineering and end-to-end learning through gradient descent [9] and back-propagation [10]. There are different types of Deep learning nets, whose usage depend and the nature of the problem being treated and on the application for which they are used. For time sequences like speech recognition [11], natural language processing [12] recurrent neural network is used.

For extracting visual features, like image classification [13] and object detection [14], convolutional neural network [15] is used. For data pattern recognition like classification and segmentation [16], [50], [51], the feed-forward networks and for some complex tasks like video processing [17], object tracking [18] and image captioning [19] the combinations of the networks are used.

The link between reinforcement learning [1] and deep learning [8] technologies was made, while artificial intelligence researchers were seeking to implement a single agent that can think and act autonomously in the real world, and get rid of any hand-engineered features [20]. In fact, in 2015, Google Deep-mind succeed to combine

reinforcement learning, which is a decision-making framework and deep learning, which is a representation learning [21] framework allowing visual features extraction, to create the first end-to-end artificial agent that achieves human-level performance in several and diverse domains. This new technology named deep reinforcement learning is used now, not only to play ATARI [22] games, but also to design next generation of intelligent self-driving cars like Google with Waymo, uber and tesla.

1.1. Preliminary: Deep Reinforcement Learning

Deep reinforcement learning(DRL) [23] is a newer machine learning technique which became popular in 2015 that combines deep learning architectures and reinforcement learning algorithms to create efficient algorithms that can be applied to solve the problems in the areas of robot engineering, healthcare maintenance, video gaming applications, financing etc. The deep reinforcement learning algorithms uses neural networks algorithms like CNN [15] to learn deep features and reinforcement learning algorithms like Q-Learning [24], actor critic [7] etc. to solve previous unsolvable problems. The traditional reinforcement learning use lookup table to store states and actions, which is too slow, since it learns the value of each state individually, and it is memory consuming, especially when we deal with large or infinite problems, and this is due to what Richard Bellman called the curse of dimensionality [25]. By leveraging deep learning algorithms, especially, convolutional neural networks, it became possible for RL algorithm not only to act but to be totally autonomous and learn to see and act. There are basically three main types of deep reinforcement learning (DRL) algorithms.

Value optimization [26]: The algorithm optimizes the Value Function V or Q or the advantage function A.

Policy optimization [27]: The algorithm optimizes the policy directly function $\pi(\theta)$ representing the neural network.

Actor-critic [7] which incorporates the advantages of each of the above, by learning value function with implicit policy: It includes Policy gradient component "Actor" which calculates policy gradients. Value function component "Critic" that observes the performance of the agent.

1.2. Preliminary: Deep Learning

Deep learning [8] is a branch of machine learning based on deep (> 2 hidden layers) and wide (many input/hidden neurons) neural networks, that model high-level abstractions in data, based on an architecture composed of multiple non-linear layers of neurons. Each neuron of the hidden layers performs a linear combination of its inputs and applies a non-linear function (Relu, Softmax, Sigmoid, and tanh,) to the result, which allows neurons from the next layer to separate classes with a curve (hypercurve/hyperplane) and no more with a simple line   thus, hidden layers learn hierarchical features. The three main types of neural networks are:

Convolutional Network [15]: A convolutional network assumes special spatial structure in its input. In particular, it assumes that inputs that are close to each other in the original input are semantically related. This assumption makes most sense for images, which is one reason convolutional layers have found wide use in deep architectures for image processing.

Recurrent Neural Network (RNN) [28]: Layers Recurrent neural network layers are primitives which permit neural networks to learn from sequences of inputs. This layer assumes that the input evolves from sequence step to next sequence step following a defined update rule which can be learned from data. This update rule presents a prediction of the next state in the sequence given all the states which have come previously.

II. Analysis of Complex Images

The analysis of histo-patholical images, MRI images, X-ray images in the form of complex images is a daunting and time-consuming task in the domain of image processing [29] and interpretation. Generally these images have indistinguishable parts which are important to be separated to determine the severity of the disease for further medical diagnosis [30]. Technically these images are processed by the system in the binary form using gray scale levels which are difficult to process for these complex images. The following are the difficulties that occurs in complex image analysis.

- The complex images in the form of histopathological images have very minute or indistinguishable differences in terms of important information they carries.

- Traditionally the extraction of the important information in form of features were mainly handcrafted using edge and corner knowledge. The extraction of this information from the images play a vital role for making accurate predictions.

- The histopathological regions of grey scale images are difficult to distinguish as relevant and irrelevant regions may be non-separable. The images needs to be segmented properly for overlapping of cells/scenes before training of the model.

- The images also needs to be denoised, enhanced and restored to enhance its quality before the model is trained.

- The process of image registration is difficult because of morphological distortion and staining variations for histopathological images.

III. Literature Review and Related Work

The following section gives a review of literature with respect to the applications of deep learning techniques and deep reinforcement learning for the year 2018-2019.

Gustavo Carneiro et al. proposed a deep reinforcement-based method for automatic detection of breast cancer [31]. The dynamic contrast enhanced magnetic resonance volumes (DCE-MRI) have been used in past for the study of brain but now its use have been broaden to pathologies of heart ,cerebral afflictions, stroke etc. with the aim of early detection and treatment. The current approaches in lesion detection are mostly dependent on handcrafted features and that too requires exhaustive search procedures to accommodate variations in lesion shape, location and size. Also, these approaches are computationally complex and lacks accuracy of lesion detection. Caicedo and Lazebnik have recently proposed the use of a deep Q-network (DQN) [32] to detect objects efficiently when the amount of data to deal with is limited. Ghesu et al. used DQN to extract consistent patterns from visual classes using fixed small-size regions of the medical images to aid anatomical landmark detection [33]. To overcome these obstacles a new algorithm for lesion detection from DCE-MRI after being inspired by previously proposed DQN is proposed. They replace the fixed small size regions with the bounding box whose size keeps on changing with every successive iteration. The reinforcement learning agent learns a policy to alter the focus of attention using scaling and translation from initially large bounding box to small bounding box inscribing a lesion if exists. The DQN decides the next action (i.e. either to scale or translate the current bounding box or to end the search process).The dataset of 117 patients divided into 57 train and 59 test cases have been used to evaluate the accuracy of proposed methodology and obtained results show same detection accuracy in accordance with the already existent methods but with huge reduction in run times.

Recently a paper titled "Deep recurrent attention models for histopathological image analysis" by Alexendre Momeni et al. discusses the analysis of histopathological images using deep recurrent models [34]. The authors proposed a Deep Recurrent Attention Model (DRAM) (Mnih et. al) to mimic pathologist process [35]. The model is inspired from the manner how human recognise visual tasks thereby attending to most informative regions of the image. Also similar to CNN it is also translation invariant but is independent of the input image size. The model is trained with the reinforcement learning to attend the most informative regions of the large image patches. The model is tested for histological and molecular subtype classification of the gliomas (a type of adult brain tumour) using data from The Cancer Genome Atlas (TCGA).The results depict that DRAM has a comparable performance to already existent methods of CNN despite processing only a selected number of image patches.

Issa Ali et al. performed a lung node detection using deep reinforcement learning [36]. The authors proposed a method with an objective to develop and validate deep reinforcement learning model based on deep artificial neural networks for early detection of lung nodules in thoracic CT images. The model is inspired from AlphaGo system which takes a raw CT image as input and views it as a collection of states, and output a classification of whether a nodule is present or not. The LIDC/IDRI database hosted by the lung nodule analysis (LUNA) challenge is used to train a model. 888 CT scans with annotations based on agreement from at least three out of four radiologists were considered for experimentation. There were 590 individuals having one or more nodules, and 298 having none. The training results yielded an overall accuracy of 99.1% [sensitivity 99.2%, specificity 99.1%, positive predictive value (PPV) 99.1%, and a negative predictive value (NPV) 99.2%]. In our test, the results yielded an overall accuracy of 64.4% (sensitivity 58.9%, specificity 55.3%, PPV 54.2%, and NPV 60.0%). These early results show an inclination towards solving the major issue of false positives in CT screening of lung nodules, and can be used to save unnecessary follow-up tests and expenditures.

C. Martinez et al. proposed a deep reinforcement learning approach for early classification of time series [37]. The time series data when classified early give rise to many applications in real life ranging from predictive maintenance to personalized medicine. This task is being addressed by a novel approach based on reinforcement learning by introducing an early classifier agent , an end-to-end reinforcement learning agent i.e. (deep Q-network, DQN) [38] to perform early classification in an efficient manner).The early classification problem is formulated in a reinforcement learning framework by

introducing a suitable set of states and actions along with define a specific reward function that aims at finding a settlement between earliness and classification accuracy. Even though there were many solutions already existent but they do not take time into account for making a final decision. The provided solution allows the user to set this trade-off in a more flexible way. Specifically the experiments were performed on datasets from the UCR time series archive [39] which showed the agent's ability in continuously adapting to the behaviour with explicit human intervention and gradually learned to maintain a balance between accurate and fast predictions.

Adriana Dinis et al. presented a self-developing system for medical data analysis [40]. In this paper authors present a concept project for a self-developing system based on agents built for a hospital. The system monitors patients during and after being released from hospitalization, with the aim of understanding patterns and predicting future problems. Due to its complexity and dynamism the agents must be automatically generated and also they need to cooperate and compete with each other in order to get good results. By combining meta-heuristic algorithms with reinforcement and clustering techniques they targeted a large degree of autonomy in decision making. Zi Wang et al. employed the reinforcement learning technique for studying the cell movements in early stage of Caenorhabditis elegans embryogenesis [41]. The proposed work captures the complexity of cell movement patterns in the embryo and overcomes the local optimization problem encountered by traditional rule- based, agent-based modelling that uses greedy algorithms. The data was collected using 3-D time lapse microscopy images to explore the cell migration paths in the process of embryogenesis. The modelled framework uses the individual cell as an agent that stores a variety of information regarding its size, fate, division time and group information. The results indicate that deep reinforcement learning based agent system has gained a remarkable success to model regulatory mechanisms of cell movements.

Leo Celi et al. proposed a deep reinforcement learning model for treatment of sepsis [42]. Sepsis is a life-threatening illness caused to multiple infections inside the body which is a leading cause of mortality among the patients in Europe. The authors use a continuous state space models and deep reinforcement learning to deduce the treatment policy for Sepsis. The data is collected from Multiparameter Intelligent Monitoring in Intensive Care (MIMIC-III v1.4) database. The results give a good learning policy that aid clinicians to make medical decisions and improve the likelihood of patient survival. Similarly Ning Liu et al. designed a deep reinforcement learning framework to estimate the optimal dynamic

treatment regimens from observational medical data [43]. The Centre for International Bone Marrow Transplant Research (CIBMTR) registry database was used to construct a dataset. The results gave a promising accuracy in predicting man expert's decisions along with the high expected reward function using Deep Reinforcement Learning (DRL) based dynamic treatment regimens on medical registry data. Alexander Rakhlin et al. used deep convolutional neural network for breast cancer histology image analysis [44]. The work proposes a computational approach based on deep convolution neural networks for breast cancer histology image classification. The ICIAR 2018 Grand Challenge on Breast Cancer Histology Images were used to construct a dataset from hematoxylin and eosin stained breast cancer histology images. The experimental results comprises of 87.2% accuracy for 4-class classification task and 93.8% accuracy for 2-class classification task to detect carcinoma. The approach outperforms other common methods in automated histopathological image classification.

Riku Turkki et al. employed breast cancer outcome prediction with tumour tissue images using deep learning [45]. The tissue microarray samples from 1299 patients of breast cancer were taken nationwide. The findings demonstrate the feasibility of learning prognostic signals in tumour tissue images without domain knowledge. Although further validation is needed, our study suggests that machine learning algorithms can extract prognostically relevant information from tumour histology complementing the currently used prognostic factors in breast cancer.

Dmitrii Bychkov et al. proposed a deep learning based tissue analysis to predict outcome in colorectal cancer [46].A novel approach for directly predicting patient outcome from digitized haematoxylin-eosin-stained tumour tissue microarray (TMA) samples from 420 cancer patients was used using a combination of convolutional and recurrent architectures. The experimental results depicts that deep learning-based outcome prediction using only a portion of tissue areas as input outperforms visual histological assessment performed by human experts.

Parampal S. Grewal et al. reviewed the use of deep learning in ophthalmology [47]. Ophthalmology is a branch of science that is used in the screening, diagnosis and management of eye disease. Deep Learning is an emerging technology with many potential applications in ophthalmology. The deep learning tools has been applied to different diagnostic purposes that includes digital photographs, optical coherence tomography and visual fields. These tools help in assessment of disease processes like cataract, diabetic retinopathy, glaucoma, age related macular degeneration etc. Nowadays the deep learning

techniques are evolving rapidly and are being integrated into ophthalmic care. The paper discuses not only the current evidence for deep learning in ophthalmology but discusses the future applications and drawbacks.

Azam Hamidinekoo et al. applied deep learning in mammography and breast histology [48]. The recent advancements in techniques of biomedical image analysis using deep learning have enhanced the performance of computer aided diagnosis (CAD) systems. The authors have proposed an overview of the state of the art overview of deep learning based applications for breast mammography histopathology images analysis. The study also reveals the relationship between mammography and histopathology phenotypes from the biological perspective.

The following table shows the publication of papers for the year 2018 and 2019 using deep reinforcement learning and deep learning and also compare the pros and cons of the techniques or algorithms used.

Table. 1 shows author, year, methods and pros/cons of the methods used for model training for DRL and DL.

IV. Conclusion and Future Work

From the survey of literature carried in section above it was found that deep learning is technique which has mainly and extensively used for classification and feature extraction from complex images using deep layers of neural networks. The powerful algorithms of deep learning like CNN [15] and RNN [28] are the most popular and successful ones. Deep reinforcement learning is an emerging and niche technique which combines the inherent properties of deep learning and policy optimization techniques using reward enhancement. The DeepQ network [38], Actor Critic [7], Sarsa [49], Temporal Difference [6] are the popular algorithms in the domain of deep reinforcement learning. The two techniques differ in the manner in which they can be used for image segmentation. The applications of image segmentation have a significant influence in the domain of medical diagnosis, image, tissue volume analysis, computer guided surgery, pathological localizations, tomography, object detection etc. The separation of image into some disjoint partitions homogenous to certain features such as colour intensity or texture is termed as image segmentation.

The competency of image segmentation methods is a debatable problem and a lot of research have been carried out in this direction. The solutions to image segmentations are generally problem based and because of resemblance in grey levels amongst the regions of interest in medical images the image segmentation can give produce significant errors. The other disadvantage that pops up is

the lack of sufficient training examples and intervention of expert systems in preparing these samples. Considering these limitations in mind the authors have given a reinforcement learning based method of image segmentation. The techniques of deep learning requires a large amount of data for model training and suffers from the problems of overfitting whereas deep reinforcement learning algorithms can be applied to train a model with limited data without causing much overfitting. The undistinguishable regions of gray scale histopathological or medical images can be efficiently classified with deep reinforcement learning techniques.

The implementation of deep reinforcement algorithms area is new and emerging technique that was first of all designed by Google Deepmind for testing the Atari games. The use of deep reinforcement learning with recurrent attention vision for complex images which may be in the form of hand written MNIST dataset was done by Volodymyr Mnih et.al [52].The future work can be carried for image analysis which may be static for an image dataset or dynamic for video formats. This requires building a robust machine learning models that can be trained with limited data and should be efficiently reinforced to achieve a desired goal.

REFERENCES

[1] Kaelbling, Leslie Pack, Michael L. Littman, and Andrew W. Moore. "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237-285, 1996.

[2] Saunders, William, et al. "Trial without error: Towards safe reinforcement learning via human intervention," in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, pp. 2067-2069, 2018.

[3] Bellman, Richard. "A Markovian decision process." *Journal of mathematics and mechanics,* pp. 679-684, 1957.

[4] Beard, Randal W., George N. Saridis, and John T. Wen. "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159-2177, 1997.

[5] Busoniu, Lucian et al., *Reinforcement learning and dynamic programming using function approximators*, CRC press, 2017.

[6] Precup, Doina, Richard S. Sutton, and Sanjoy Dasgupta. "Off-policy temporal-difference learning with function approximation," in *Proceedings of ICML*, pp. 417-424, 2001.

Table 1. Model training for DRL and DL.

References	Algorithm/Methods for training machine learning model	Pros and Cons of the algorithm
Gustavo Carneiro et al. [31] (2019) Alexendre Momeni et al. [34] (2019) Issa Ali et al. [36] (2018) C. Martinez et al. [37] (2018) Adriana Dinis et al. [40]. (2018) Zi Wang et al. [41] (2018) Leo Celi et al. [42] (2017) Ning Liu et al. [43] (2017)	**Deep Q Learning:** -The deep Q-learning algorithm is a deep reinforcement learning based algorithm that combines deep learning based neural network architecture with reinforcement learning based Q-learning. -The algorithm was developed by google deep-mind in 2015 for testing ATARI gaming platform. It showed that deep Q-network based agent on given the pixels and game score in input form was able to outperform against all the already known algorithms and gave a similar accuracy of a human game tester when tested against a set of 49 games. -In deep Q-learning, we use a neural network to approximate the Q-value function. The state is given as the input and the Q-value of all possible actions is generated as the output.	**Pros**: The Deep Q algorithm is a powerful algorithm that combines the feature extraction capability of CNN and reward enhancement capability of reinforcement learning to achieve a desired objective function. DQ learning agent enable an autonomous decision making with high accuracies and precision. **Cons:** DQN requires a large amount of time to train a reinforcement learning agent. The rewards and environment must be effectively defined.
Alexander Rakhlin et al. [44] (2018) Riku Turkki et al .[45] (2019) Dmitrii Bychkov et al. [46] (2018) Parampal S. Grewal et al. [47] (2018) Azam Hamidinekoo et al. [48] (2018)	**CNN (Convolutional Neural Network):** -A CNN stands for convolutional neural network. -CNN is a combination of different layers. The four important ones are: input layer, convolutional layer, activation layer and pool layer. It is an important algorithm for image classification. 1. Input Layer: This layer holds the raw input of image with width 32, height 32 and depth 3. 2. Convolution Layer: This layer computes the output volume by computing dot product between all filters and image patch. Suppose we use total 12 filters for this layer we'll get output volume of dimension 32 x 32 x 12. 3. Activation Function Layer: This layer will apply element wise activation function to the output of convolution layer. 4. Pool Layer: This layer is periodically inserted in the covnets and its main function is to reduce the size of volume which makes the computation fast reduces memory and also prevents from overfitting.	**Pros** -CNN is a very powerful algorithm which is widely used for image classification and object detection. The hierarchical structure and powerful feature extraction capabilities from an image makes CNN a very robust algorithm for various image and object recognition tasks. **Cons**: -CNN requires a large amount of data to train the model. -The CNN output changes with the adversarial effects of input representations .i.e. the output changes as the input is changed.

[7] Konda, Vijay R., and John N. Tsitsiklis, "Onactor-critic algorithms," *SIAM journal on Control and Optimization,* vol. 42, no. 4, pp. 1143-1166, 2003.

[8] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature* 521.7553, pp. 436-444, 2015.

[9] Bottou, Léon. "Stochastic gradient learning in neural networks," in *Proceedings of Neuro-Nımes* 91.8, vol. 12, 1991.

[10] Benvenuto, Nevio, and Francesco Piazza, "On the complex backpropagation algorithm," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 967-969, 1992.

[11] Graves, Alex, and Navdeep Jaitly, "Towards end-to-end speech recognition with recurrent neural networks," in *Proceedings of International conference on machine learning*, pp. 1764-1772, 2014.

[12] Cho Kyunghyun et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

[13] McCoppin, Ryan, and Mateen Rizki, "Deep learning for image classification," *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR V*, vol. 9079. International Society for Optics and Photonics, pp. 946401-1, 2014.

[14] Price, Micah, et al., "Object detection using image classification models," U.S. Patent No. 10,223,611. 5 Mar. 2019.

[15] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[16] Naylor Peter et al., "Nuclei segmentation in histopathology images using deep neural networks," in *Proceedings of IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 933-936, 2017.

[17] Ullah, Amin, et al., "Action recognition in video sequences using deep bi-directional LSTM with CNN features," *IEEE Access*, vol. 6, pp. 1155-1166, 2017.

[18] Wang, Naiyan, and Dit-Yan Yeung, "Learning a deep compact image representation for visual tracking," *Advances in neural information processing systems*, pp. 809-817, 2013.

[19] You, Quanzeng et al., "Image captioning with semantic attention," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4651-4659, 2016.

[20] Le, Quoc V., "Building high-level features using large scale unsupervised learning," in *Proceedings of the IEEE international conference on acoustics, speech and signal processing*, pp. 8595-8598, 2013.

[21] Radford, Alec, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2015.

[22] Mnih, Volodymyr et al., "Playing atari with deep reinforcement learning," arXiv preprint arXiv:1312.5602, 2013.

[23] Mnih, Volodymyr et al., "Human-level control through deep reinforcement learning," *Nature* 518.7540-529, 2015.

[24] Watkins, Christopher JCH, and Peter Dayan, "Q-learning." *Machine learning*, vol. 8, no. 3-4, pp. 279-292, 1992.

[25] Rust, John, "Using randomization to break the curse of dimensionality," *Econometrica: Journal of the Econometric Society*, pp. 487-516, 1997.

[26] Nachum, Ofir, et al., "Bridging the gap between value and policy based reinforcement learning," *Advances in Neural Information Processing Systems*, pp. 2775-2785, 2017.

[27] Grudic, Gregory Z., Vijay Kumar, and Lyle Ungar, "Using policy gradient reinforcement learning on autonomous robot controllers," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)* (Cat. No. 03CH37453). vol. 1, pp. 406-411, 2003.

[28] Mikolov, Tomáš, et al., "Recurrent neural network based language model," in *Proceedings of Eleventh annual conference of the international speech communication association*, 2010

[29] Nixon, Mark, and Alberto Aguado, *Feature extraction and image processing for computer vision*. Academic press, 2019

[30] Bakator, Mihalj, and Dragica Radosav, "Deep learning and medical diagnosis: A review of literature," *Multimodal Technologies and Interaction*, vol. 2, no. 3 2018.

[31] Maicas, Gabriel, et al., "Deep reinforcement learning for active breast lesion detection from DCE-MRI," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, pp. 665-673, 2017.

[32] Caicedo, Juan C., and Svetlana Lazebnik, "Active object localization with deep reinforcement learning," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2488-2496, 2015.

[33] Ghesu, Florin C. et al., "An artificial agent for anatomical landmark detection in medical images," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, pp. 229-237, 2016.

[34] Momeni, Alexandre, Marc Thibault, and Olivier Gevaert, "Deep Recurrent Attention Models for

Histopathological Image Analysis," BioRxiv: 438341, 2018.

[35] Mnih, Volodymyr, Nicolas Heess, and Alex Graves, "Recurrent models of visual attention," Advances in neural information processing systems, pp. 2204-2212, 2014.

[36] Ali, Issa et al., "Lung nodule detection via deep reinforcement learning," *Frontiers in oncology*, vol. 8, 2008.

[37] Martinez Coralie et al., "A deep reinforcement learning approach for early classification of time series," in *Proceedings of the 26th European Signal Processing Conference (EUSIPCO)*, pp. 2030-2034, 2018.

[38] Volodymyr Mnih et al., "Human-level control through deep reinforcement learning," *Nature* 518.7540, pp. 529-533, 2015.

[39] Chen Yanping et al., "The ucr time series classification archive," *http://www. cs. ucr. edu/~ eamonn/time_series_data*, 2015.

[40] Dinis Adriana, Todor Ivascu, and Viorel Negru, "A Self Developing System for Medical Data Analysis," in *Proceedings of the 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, pp. 335-339, 2018.

[41] Wang Zi et al., "Deep reinforcement learning of cell movement in the early stage of C. elegans embryogenesis," *Bioinformatics*, vol. 34, no. 18, pp. 3169-3177, 2018.

[42] Raghu, Aniruddh et al., "Deep reinforcement learning for sepsis treatment," arXiv preprint arXiv:1711.09602, 2017.

[43] Liu Ying et al., "Deep reinforcement learning for dynamic treatment regimens on medical registry data," in *Proceedings of IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 380-385, 2017.

[44] Rakhlin Alexander et al., "Deep convolutional neural networks for breast cancer histology image analysis," in *Proceedings of International Conference Image Analysis and Recognition,* Springer, Cham, pp. 737-744, 2018.

[45] Turkki Riku et al., "Breast cancer outcome prediction with tumour tissue images and machine learning," *Breast cancer research and treatment*, pp. 1-12, 2019.

[46] Bychkov Dmitrii et al., "Deep learning based tissue analysis predicts outcome in colorectal cancer," *Scientific Reports*, vol. 8, no.1: 3395, 2018.

[47] Grewal Parampal S. et al., "Deep learning in ophthalmology: a review," *Canadian Journal of Ophthalmology*, vol. 53, no. 4, pp. 309-313, 2018.

[48] Hamidinekoo Azam et al., "Deep learning in mammography and breast histology, an overview and future trends," *Medical image analysis*, vol. 47, pp. 45-67, 2018.

[49] Zhao Dongbin et al., "Deep reinforcement learning with experience replay based on SARSA," in *Proceedings of IEEE Symposium Series on Computational Intelligence (SSCI)*, pp 1-6, 2016.

[50] Ji-Hae Kim, Gwang-Soo Hong, Byung-Gyu Kim, Debi P. Dogra, "deepGesture: Deep Learning-based Gesture Recognition Scheme using Motion Sensors," *Displays*, vol. 55, pp. 38-45, 2018.

[51] Ji-Hae Kim, Byung-Gyu Kim, Partha Pratim Roy, Da-Mi Jeong, "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure," *IEEE Access*, vol. 7, pp. 41273-41285, 2019.

[52] Mnih, Volodymyr, Nicolas Heess, and Alex Graves, "Recurrent models of visual attention," *Advances in neural information processing systems*, pp 2204-2212, 2014.

Authors

Rishi Khujria has received Master in Computer Application in 2017 from Department of Computer Sc. & IT, University of Jammu and now Pursuing Ph.D. from University of Jammu . He is working on Reinforcement Learning on Medical Disorders. He is JRF research fellow in department of Computer Sc. & IT, University of Jammu. He has interdisciplinary research interests spanning the area of applications of Artificial Intelligence in medical diagnosis.

Abdul Quyoom received B.Tech from BGSBU Rajouri, J&K in 2013 and he is awarded M.Tech from Central University of Rajasthan specialized in Information security in 2015 . He has five Publications and attended various National, international conferences and workshops. He has worked as an Assistant Professor in YCET Jammu and the department of Computer Science in BGSBU Rajouri. Currently, he is pursuing Ph.D. from Department of Computer Sc. & IT, University of Jammu and working on Artificial Intelligence and its application in medical domain. His research interest includes medical image processing, intelligent image segmentation, and deep learning.

Dr. Abid Sarwar has been working in the field of application of Artificial Intelligence in Medicine (especially in cervical cancer and diabetes) for the last 8 years. He did Masters in Computer Applications from Department of Computer Sc. & IT, University of Jammu in 2009. He obtained PhD degree from Department of Computer Sc. & IT, University of Jammu in 2017. Besides He has published more than 15 research articles in leading journals, conference proceedings, he has created a database of 8,091 digitally calibrated cervical cells, which is the only research database available to work on cervical cancer based on Bethesda system of classification. His research interest includes medical image processing, intelligent image segmentation, and deep learning.