2021 IEEE International Conference on Image Processing (ICIP) | 978-1-6654-4115-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIP42928.2021.9506805

ADVANCED DEEP NETWORK WITH ATTENTION AND GENETIC-DRIVEN REINFORCEMENT LEARNING LAYER FOR AN EFFICIENT CANCER TREATMENT OUTCOME PREDICTION

Francesco Rundo[1], *Giuseppe Luigi Banna*[2], *Francesca Trenta*[3], *Sebastiano Battiato*[3]

[1] STMicroelectronics, ADG Central R&D, Catania, Italy
Email: francesco.rundo@st.com
[2] Queen Alexandra Hospital, Medical Oncology Department, Lincolnshire, UK
Email: giuseppe.banna@nhs.net
[3] University of Catania - Department of Mathematics and Computer Science, Catania, Italy
Email: francesca.trenta@unict.it, battiato@dmi.unict.it

ABSTRACT

In the last few years, medical researchers have investigated promising approaches for cancer treatment, leading to a major interest in the immunotherapeutic approach. The target of immunotherapy is to boost a subject's immune system in order to fight cancer. However, scientific studies confirmed that not all patients have a positive response to immunotherapy treatment. Medical research has long been engaged in the search for predictive immunotherapeutic-response biomarkers. Based on these considerations, we developed a non-invasive advanced pipeline with a downstream 3D deep classifier with attention and reinforcement learning for early prediction of patients responsive to immunotherapeutic treatment from related chest-abdomen CT-scan imaging. We have tested the proposed pipeline within a clinical trial that recruited patients with metastatic bladder cancer. Our experiment results achieved accuracy close to 93%.

Index Terms— Deep Convolutional Network, Immunotherapy, Imaging.

1. INTRODUCTION

Immunotherapy acts on the body's immune system to stimulate the destruction of cancer cells [1, 2]. The immune system is the main subject body's natural defense barrier for infection and correlated disease. However, the body defense mechanism is not always effective against cancer cells because they are able to implement a whole series of escape strategies[1, 2]. Based on this increasingly in-depth knowledge, researchers have investigated several strategies based on "re-educating" the immune system against the aforementioned cancer escape strategies. In this contribution, we will focus on the "Immune Checkpoint Inhibitors" (ICIs) immunotherapy treatments that act over the so called PD-L1/PD-1 cell receptors [3]. The ICIs immunotherapeutic drugs inhibit the action of PD-L1/PD-1 receptors used by cancer cell to prevent T lymphocytes (subject's immune system) from destroying tumor cells. We analyzed this physiological mechanism in the medical use-case of metastatic bladder cancer [1, 4, 5]. Unfortunately, despite the growing advances of immunotherapy treatments, only about 20% - 30% of patients have a positive response [6, 7]. Therefore, the scientific community has focused its research efforts on delivering innovative predictive biomarkers of response to immunotherapy treatment [8]. In this study, we explored the application of an innovative and less invasive 3D deep pipeline with self-attention and data augmentation, which aims to predict a bladder cancer suffering patient's response to immunotherapy treatment from the analysis of chest-abdomen CT-scan imaging.

2. RELATED WORKS

An extensive research effort (including recent Deep Learning) has been employed by scientific community for identifying efficient predictive bio-markers of response to cancer treatments[9, 10, 11]. Garapati et al. [10] applied some of the most common Machine Learning (ML) algorithms to analyze CT-scan urography of each recruited patient belonging to a clinical study of 84 subjects. The reported evaluation confirmed that Support Vector Machine (SVM) achieved impressive results[10]. Other ML methods have been proposed for bladder cancer disease estimation[11]. Specifically, the method proposed in [11] described an interesting and effective (accuracy close to 80 %) approach for predicting cancer recurrence and survival in recruited treated patients from multi-modal data analysis (imaging, surgical findings, etc.). Further interesting deep pipelines for estimating the response to such cancer treatments based on quantitative data analysis have been proposed in [12, 13, 14]. The Deep Learning architecture proposed in [12] includes a modified version of the AlexNet backbone [13] used to learn visual features from seg-

mented patient CT imaging in order to assess the chemotherapic treatment outcome. The experimental results pointed out the effectiveness of the proposed solution [12]. In [15], the authors introduced a novel deep stack of encoders for predicting immunotherapy outcomes from visual features based on CT-scan imaging. With an accuracy of 86.05%; Specificity of 89.29%, and Sensitivity of 80.00% , the proposed pipeline showed promising results. Further enhancements of that deep architecture have been proposed in [16] and [17]. More in detail, through innovative deep pipelines including self-attention [16] and data augmentation [17] the authors retrieved promising results regarding the prediction of the immunotherapy treatment outcome as confirmed by the performance results that will be later compared in this paper.

3. THE PROPOSED PIPELINE: DESCRIPTION

The flowchart of the proposed pipeline is shown in Fig. 1. We trained a deep architecture to predict the outcome of the immunotherapy treatment from CT-scan imaging of the analyzed patients. The network explored in this work consists of an innovative 3D Densely Connected Convolutional Network embedding self-attention combined with an innovative genetic-driven reinforcement learning layer able to generate hierarchical augmented spatio-temporal features with the target of minimizing the loss function of the overall pipeline. The use of 3D deep convolutional layers was enabled by the results achieved by our previous works [16, 17]. As reported in Fig. 1, the output of the proposed pipeline is the predicted immunotherapy treatment outcome i.e. a patient treatment responsiveness classification as complete (CR as Complete Response), partial (PR as Partial disease Regression) or stable (SD as Stable Disease) – Class 1- or, conversely, a predictive classification (Class 2) of disease progression (PD as Progressive disease). In medical field, the neoplastic disease evolution is monitored through a set of medical guidelines known as Response Evaluation Criteria in Solid Tumors (RECIST) which were developed by the scientific community to objectively identify the visual lesions found in CT-scan/MRI imaging that characterize the disease [18]. In this work, only RECIST (ver. 1.1) CT-scan compliant lesions will be processed. As described in Fig. 1, the involved physicians select the RECIST 1.1 compliant lesion by using supplied CT-scan device imaging software, obtaining the lesion to be analyzed, i.e. the Region of Interest (ROI) $I_s(x, y)$. The segmented $I_s(x, y)$ will be processed by a Genetic-driven Reinforcement Learning block in order to generate augmented spatio-temporal features arranged as a Volume of Interest (VOI) of $T_D \times M_D \times N_D$ to be fed to the downstream 3D Densely Connected Classifier. In this work, we defined $16 \times 64 \times 64$ as a dimension of each VOI. To improve the overall performance of the proposed pipeline, we have extended 3D convolution architectures with an implicit attention mechanism based on the usage of Non-Local Blocks [19]. An overview of the implemented deep architecture is introduced. The backbone of the proposed 3D deep classifier is a sequence of dense blocks. The model processes a batch of $16 \times 64 \times 64$ augmented spatio-temporal VOI generated through the Genetic-driven Reinforcement Learning block applied to the segmented RECIST 1.1 compliant input CT-scan image lesion. This VOI is first fed to a 3D convolutional layer with a kernel size of $3 \times 3 \times 3$, providing an output of 32 features depth. These feature maps will be processed by six dense blocks composed by [6, 8, 8, 8, 8, 6] 3D layers, respectively, with the same kernel size, followed by ReLU non-linear activations. Each dense block is preceded by [0, 1, 2, 3, 4, 5] Embedded Gaussian Non-Local blocks [19] respectively and each dense block is followed by a transition-down layer with $2 \times 2 \times 2$ max pooling. Thus, the input VOI will be processed by the described blocks generating the feature maps which will gradually decrease (in dimension) until it becomes a one-dimensional vector having a length of 736×1. The resulting feature map traverses two Fully Connected (FC) layers followed by RELU, except the last one that, instead, uses a SoftMax layer for the final binary classification. Classical negative log-likelihood is used as loss.

3.1. The CT-scan pre-processing block: CT-scan ROI Segmentation

As introduced and reported in Fig. 1, the oncologists / radiologists pre-process the chest-abdomen CT-scan of the patient through the classical software usually supplied with the CT scanner device [16, 17, 18]. For each patient, the oncologist / radiologist will select the RECIST 1.1 compliant CT-scan image lesion. In order to minimize the subjective impact of the physician-driven manual choice of the RECIST 1.1. compliant lesion, the pipeline validation phase is carried out by means of a k-fold cross-validation session selecting, from time to time, all RECIST 1.1 compliant lesions of each patient. In this way, we have evaluated the performance of the pipeline in each possible setup of the physician's choice of the input CT-scan lesion to be processed. The output of this block is the ROI $I_s(x, y)$ which will be normalized (bi-cubic interpolation) in size and fed to the Genetic-driven Reinforcement Learning block.

3.2. The Genetic-driven Reinforcement Learning Block

This block is designed to process the segmented ROIs in order to generate high-level spatio-temporal discriminant features. To perform this augmentation task, we extended the classical 2D Cellular Non-linear Networks (2D-CNN) [20] firstly introduced by Chua and Yang [21]. The enhanced 2D-CNN embeds a state-controlled template with a transient-response stage [17, 18, 19, 20]. Let us introduce the proposed enhanced 2D-CNN model:

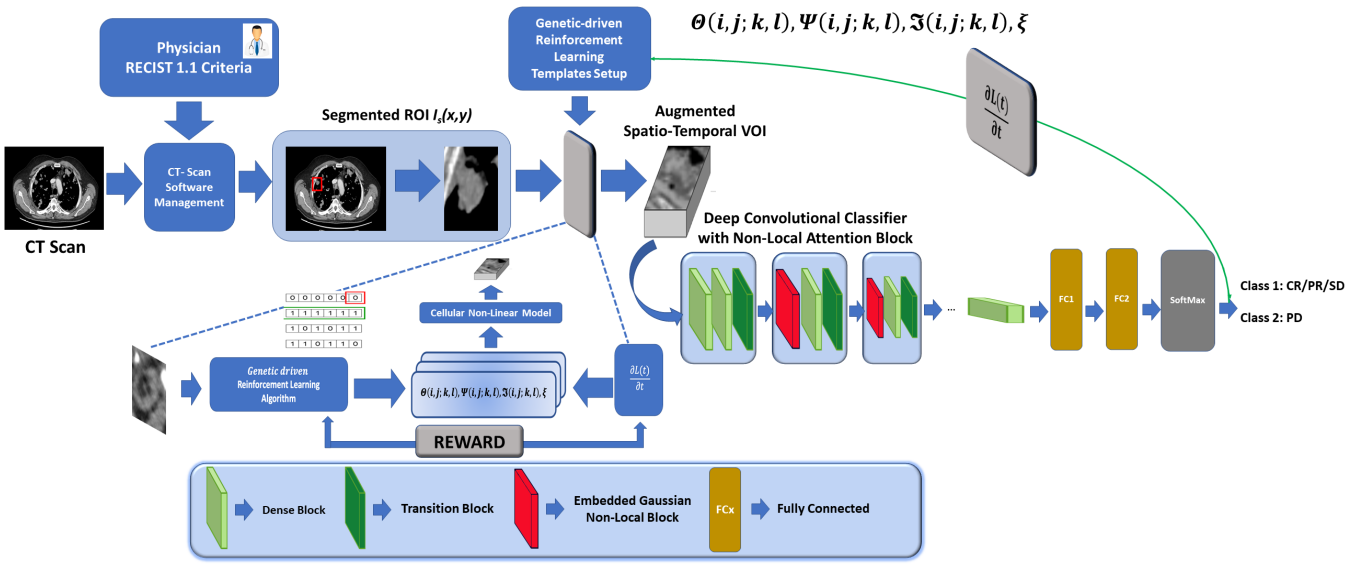

Fig. 1. The proposed 3D Densely Connected Network with Non-Local Block and Unsupervised Genetic-Driven Reinforcement Learning Layer.

$$
\begin{aligned}
C\frac{dx_{ij}(t)}{dt} = & -\frac{1}{R_x}x_{ij}+ \\
& + \sum_{C(k,l)\in N_r(i,j)} \Theta(i,j;k,l)y_{kl}(t)+ \\
& + \sum_{C(k,l)\in N_r(i,j)} \Psi(i,j;k,l)u_{kl}(t)+ \\
& + \sum_{C(k,l)\in N_r(i,j)} \Im(i,j;k,l)x_{kl}(t)+ \\
& + \xi \\
& 1 \le i \le M, 1 \le j \le N; t \le t_k
\end{aligned} \tag{1}
$$

$$
y_{ij}(t) = \frac{1}{2}(|x_{ij}(t)+1| - |x_{ij}(t)-1|) \tag{2}
$$

$$
\begin{aligned}
N_r(i,j) = \{C_r(k,l); (max(|k-i|,|i-j|) \le r)\} \\
(1 \le k \le M, 1 \le l \le N))
\end{aligned} \tag{3}
$$

In our pipeline, $R_x = 1$ while each pixel of the segmented $I_s(x,y)$ will be fed to the model as input u_{kl} and state x_{kl}. In Eqs. (1)-(3) the $Nr(i,j)$ represents the neighborhood of each 2D-CNN neuron-cell $C(i,j)$ with radius r while $y_{ij}(t)$ represents the generated hierarchical feature. The dynamic evolution of the 2D-CNN model occurs in a reduced transient t_k. We defined 16 different setup of the 2D-CNN transient processing. Specifically, the 3×3 cloning templates $\Theta(i,j;k,l)$, $\Psi(i,j;k,l)$, $\Im(i,j;k,l)$ and bias ξ for each of the defined 16 setup are randomly initialized together with a 3×3 binary mask $\Theta^{B^v}(i,j;k,l)$ for the $v-th$ template matrix $\Theta^v(i,j;k,l)$ with $v = 1, 2...16$. During the training phase of the whole pipeline, the temporal dynamics of the overall loss $L(t)$ will be retro-propagated to this block which will use it to configure the elements of the matrix $\Theta^v(i,j;k,l)$ using a Reinforcement Learning (RL) algorithm. The other cloning templates remain constant. More in detail, we investigated the optimal policy P_o that optimize the cumulative discount reward R:

$$
P_0 = argmax_{P_0}\ E[\sum_{t\ge0} \gamma^t R(.|s_t,a_t)|P_0] \tag{4}
$$

Where γ is a proper discounted coefficient in (0,1). In order to evaluate the state s_t (specific setup of the $v-th$ cloning templates), we defined the Value function $V^{P_0}(s_t)$ and the Q-value function $Q^{P_0}(s_t,a_t)$ respectively:

$$
V^{P_0}(s_t) = E[\sum_{t\ge0} \gamma^t R(.|s_t)|P_0] \tag{5}
$$

$$
Q^{P_0}(s_t,a_t) = E[\sum_{t\ge0} \gamma^t R(.|s_t,a_t)|P_0\ ] \tag{6}
$$

while the reward function is so defined:

$$
R = -(\frac{\partial L(\Theta(\cdot),\Psi(\cdot),\Im(\cdot),\xi,\Theta^{B^v}(\cdot),B_v,v,t)}{\partial t})^2 \tag{7}
$$

Where $L(\cdot)$ is the loss of the overall pipeline which depends on the state s_t ($\Theta(i,j;k,l)$, $\Psi(i,j;k,l)$, $\Im(i,j;k,l)$ and bias ξ) and the actions a_t and policy P_0 defined by the update of the mask $\Theta^{B^v}(i,j;k,l)$ and B_v. For each

Table 1. Experimental performance benchmarking (Mean ± Standard Deviation (STD))

Model	Metrics					
	Accuracy		***Sensitivity***		***Specificity***	
	Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***
2D ResNet-50	$0,827$	$0,040$	$0,817$	$0,048$	$0,837$	$0,070$
2D DenseNet-201	$0,830$	$0,036$	$0,847$	$0,069$	$0,813$	$0,067$
2D ResNet-101	$0,827$	$0,040$	$0,817$	$0,048$	$0,837$	$0,070$
2D VGG-19	$0,775$	$0,071$	$0,827$	$0,082$	$0,723$	$0,115$
ResNet-18 + Aug. [17]	$0,918$	$0,0439$	$0,917$	$0,064$	$0,921$	$0,045$
3D Resnet-101	$0,857$	$0,0487$	$0,847$	$0,044$	$0,867$	$0,065$
3D DenseNet-201	$0,840$	$0,047$	$0,833$	$0,051$	$0,847$	$0,065$
3D DenseNet-NLB [16]	$0,913$	$0,035$	$0,923$	$0,062$	$0,904$	$0,033$
Proposed	**0,932**	**0,038**	**0,933**	**0,061**	**0,931**	**0,050**
Proposed w/o NLB	$0,913$	$0,033$	$0,916$	$0,047$	$0,911$	$0,048$
Proposed w/o Gen-driv RL	$0,901$	$0,038$	$0,903$	$0,050$	$0,900$	$0,550$

training iteration, a Genetic-driven approach through classical $crossover$ and $mutation$ operations [22] applied to the binary mask B_v , selects the $v - th$ cloning templates setup to modify (among to the 16 defined setup) and always with a set of $crossover$ and $mutation$ operations, it will change the binary mask $\Theta^{B^v}(i, j; k, l)$ of the selected $v - th$ setup thus identifying the coefficient of the cloning template $\Theta^{B^v}(i, j; k, l)$ which will be randomly updated (action a_t) generating a new configuration of the 2D-CNN model. Only setup that produce a decrease in the overall loss dynamic will be accepted while the others will be discarded.

3.3. Self-Attention through Non-Local Blocks

As introduced, the proposed downstream classifier consists of a 3D DenseNet (3D-DCNN) embedding separable densely connected layers (both depth-wise and point-wise) [16]. The output of each dense block is then passed to Non-Local Blocks (NLBs). Non-local blocks have been recently introduced [19], as a very promising approach for capturing space-time long-range dependencies on feature maps, resulting in a sort of "self-attention" mechanism. The mathematical processing of Non-Local Block is introduced. Given a generic deep network as well as a general Non-Local Block input data x (feature map), the employed non-local operation computes the corresponding response y_i at a i location in the input data as a weighted sum of the input data at all positions $j \neq i$:

$$y_i = \frac{1}{\psi'(x)} \sum_{\forall j} \zeta(x_i,\ x_j)\beta(x_j) \tag{8}$$

With $\zeta(\cdot)$ being a pairwise potential describing the affinity between data positions at index i and j respectively. $\beta(\cdot)$ is, instead, a unary potential modulating ζ according to input data. The sum is then normalized by a factor $\psi'(x)$. The parameters of potentials $\zeta(\cdot)$ are learned during model's training and defined as follows:

$$\zeta(x_i,\ x_j) = e^{\Theta'(x_i)^T\ \Phi(x_j)} \tag{9}$$

Where Θ' and Φ are two linear transformations of the input data x with learnable weights $W_{\Theta'}$ and W_{Φ} [19]. For the $\beta(\cdot)$ function, a common linear embedding (classical 1x1x1 convolution) with learnable weights W_β is defined. In Eqs (8)-(9) an Embedded Gaussian setup is reported [19]. The so processed features will be fed into the final block of the 3D-DCNN composed by a stack of two FC layers FC_1 , FC_2 having 350 and 250 neurons respectively. A SoftMax layer for binary classification close the 3D deep classifier.

3.4. Dataset: Recruitment and data pre-processing

A dataset of 106 histologically confirmed bladder cancer CT-scan image lesions were retrospectively analyzed within a clinical trial under the IRB "Catania 1 Ethical Committee" Nr. D4191C00068 and MO29983 approval. Each recruited patients were treated with a PD-1/PD-L1 immunotherapy agent. The used CT scanner: GE multi-slice (64 slices) with a thickness of 2.5 mm. More in detail, 43 target lesions (of the 106) are associated with patients who experienced positive response or a disease stabilization (Class 1), while 63 lesions are associated to patients experienced disease progression (Class 2). We divided the dataset into a training/validation set of 76 target lesions (28 of Class 1 and 48 of Class 2) and a test-set including 30 CT images (15 of Class 1 and 15 of Class 2). For each lesion, an augmented $16 \times 64 \times 64$ VOI was generated. The overall deep backbone was trained with an initial learning rate of $3e^{-4}$ and a mini-batch size of 10. A k-fold cross-validation session has been applied.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Table I reports comparison results in terms of accuracy, sensitivity and specificity. As expected, the collected results highlight the out-performance of 3D architectures compared to 2D ones. The augmentation pipeline developed by the authors in [17] allows bringing the performance of the ResNet-18-based deep backbone to higher values even than 3D architectures. A similar analysis can be applied to self-attention techniques[16]. Therefore, the high performance of our proposed pipeline (Accuracy: $0,932 \pm 0,038$ - Sensitivity: $0,933 \pm 0,061$ - Specificity: $0,931 \pm 0,050$) appears in line with the mentioned scientific evidence as it combines both self-attention and augmentation operations. Although a larger and more balanced dataset is desirable, in medical practice this is not always feasible due to long, expensive and complex clinical studies. The reported results confirmed the effectiveness of the proposed approach even in the presence of limited cases.

5. REFERENCES

[1] Hashem O Alsaab et al., "Pd-1 and pd-l1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome," *Frontiers in pharmacology*, vol. 8, pp. 561, 2017.

[2] Kristen R Spencer et al., "Biomarkers for immunotherapy: current developments and challenges," *American Society of Clinical Oncology educational book*, vol. 36, pp. e493–e503, 2016.

[3] Xiangli Ding et al., "Clinicopathological and prognostic value of pd-l1 in urothelial carcinoma: a meta-analysis," *Cancer management and research*, vol. 11, pp. 4171, 2019.

[4] Tian C Zhou et al., "A review of the pd-1/pd-l1 checkpoint in bladder cancer: From mediator of immune escape to target for treatment," in *Urologic Oncology: Seminars and Original Investigations*. Elsevier, 2017, vol. 35, pp. 14–20.

[5] Jacques Ferlay et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012," *International journal of cancer*, vol. 136, no. 5, pp. E359–E386, 2015.

[6] Padmanee Sharma et al., "Efficacy and safety of nivolumab monotherapy in metastatic urothelial cancer (muc): Results from the phase i/ii checkmate 032 study.," 2016.

[7] Christophe Massard et al., "Safety and efficacy of durvalumab (medi4736), an anti–programmed cell death ligand-1 immune checkpoint inhibitor, in patients with advanced urothelial bladder cancer," *Journal of Clinical Oncology*, vol. 34, no. 26, pp. 3119, 2016.

[8] Giuseppe Luigi Banna et al., "The promise of digital biopsy for the prediction of tumor molecular features and clinical outcomes associated with immunotherapy," *Frontiers in medicine*, vol. 6, pp. 172, 2019.

[9] Guanjin Wang et al., "Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques," vol. 63, pp. 124–132, 2015.

[10] Sankeerth S Garapati et al., "Urinary bladder cancer staging in ct urography using machine learning," *Medical physics*, vol. 44, no. 11, pp. 5814–5823, 2017.

[11] Zaki Hasnain et al., "Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients," *PloS one*, vol. 14, no. 2, 2019.

[12] Eric Wu, , et al., "Deep learning approach for assessment of bladder cancer treatment response," *Tomography*, vol. 5, no. 1, pp. 201, 2019.

[13] Alex Krizhevsky et al., "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[14] Kenny H Cha et al., "Bladder cancer treatment response assessment in ct using radiomics with deep-learning," *Scientific reports*, vol. 7, no. 1, pp. 1–12, 2017.

[15] Francesco Rundo et al., "Advanced deep learning embedded motion radiomics pipeline for predicting anti-pd-1/pd-l1 immunotherapy response in the treatment of bladder cancer: Preliminary results," *Electronics*, vol. 8, no. 10, pp. 1134, 2019.

[16] Francesco Rundo et al., "3d non-local neural network: A non-invasive biomarker for immunotherapy treatment outcome prediction. case-study: Metastatic urothelial carcinoma," *Journal of Imaging*, vol. 6, no. 12, pp. 133, 2020.

[17] Francesco Rundo et al., "Advanced non-linear generative model with a deep classifier for immunotherapy outcome prediction: A bladder cancer case study," in *ICPR Workshop 2020 - Artificial Intelligence for Healthcare Applications (AHIA 2020)*, 2021, accepted.

[18] Elizabeth A Eisenhauer et al., "New response evaluation criteria in solid tumours: revised recist guideline (version 1.1)," *European journal of cancer*, vol. 45, no. 2, pp. 228–247, 2009.

[19] Xiaolong Wang et al., "Non-local neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7794–7803.

[20] Tamas Roska and Leon O Chua, "Cellular neural networks with non-linear and delay-type template elements and non-uniform grids," *International Journal of Circuit Theory and Applications*, vol. 20, no. 5, pp. 469–481, 1992.

[21] Leon O Chua and Lin Yang, "Cellular neural networks: Theory," *IEEE Transactions on circuits and systems*, vol. 35, no. 10, pp. 1257–1272, 1988.

[22] Adarsh Sehgal et al., "Deep reinforcement learning using genetic algorithm for parameter optimization," in *2019 Third IEEE International Conference on Robotic Computing (IRC)*. IEEE, 2019, pp. 596–601.