RESEARCH ARTICLE OPEN ACCESS

Continual Learning Systems with Controlled-Complexity using Segmented Reinforcement Machine Learning Algorithms

Hari Krishna Modalavalasa*, Madhavi Latha Makkena**

*(ECE, JNTUH-CEH, Jawaharlal Nehru Technological University-Hyderabad
Email: mhkrishna9s@gmail.com)

** (ECE, JNTUH-CEH, Jawaharlal Nehru Technological University-Hyderabad
Email: mlmakkena@gmail.com)

--************************----------------------------------

Abstract:

Continual Learning is the one of the most emerging research areas in Machine Learning. Continual Learning Models learn adaptively and continuously even after their deployment into real time environment for incremental knowledge about the outside world. Continual Learning plays crucial role in increasing the accuracy of unsupervised and reinforcement learning models. The training complexity or training time period should be minimal in order to meet the timing constraints of real time applications with limited computational resources. Many complex policy gradient and policy optimization reinforcement algorithms predict high accurate solutions with best rewards but they require huge amount of training times. These algorithms cannot meet the time constraints of the continual systems. To overcome this issue two novel Segmented Reinforcement Machine Leaning (SRML) algorithms are proposed in this paper. SRML algorithms use novel search methodologies with controlled complexity for smaller training times. A hybrid continual learning model with Segmented and Recursive Reinforcement Learning algorithms and Segmented and Adaptive Reinforcement Learning algorithm is proposed for proper initial training and quick continual training for lifelong. Wavelet based Image denoising application is considered to validate and compare the performances of existing and proposed algorithms.

Keywords **—Continual Learning, model-free reinforcement learning, Segmented and Recursive Reinforcement Learning algorithms, Segmented and Adaptive Reinforcement Learning algorithm.**

--************************----------------------------------

I. INTRODUCTION

Now a days Artificial Intelligence (AI) plays very crucial role in development of intelligent systems and automation products. Machine Learning (ML) algorithms provide the ability to understand the environment without the need of traditional programming and act as brain of AI systems. ML algorithms dig the data to explore the hidden structures or relationships inside the data and convert those relationships as a ML model. These ML models can understand the future inputs and predict the future outputs or actions. ML algorithms are broadly classified into 3 different categories depending on the availability of datasets [1][2]. First and most power class is Supervised Machine Learning (SML) algorithms. These SML algorithms required huge number of inputs and outputs along with the relationships between them which is known as labelled dataset. Creation of labelled datasets is very complex, time consuming procedure and requires huge manual efforts [1][3]. Unsupervised Machine Learning (UML) is the second category of ML algorithms which is very beneficial in case of system outputs are known prior to training but the relationships between inputs and outputs are unknown. UML algorithms extract the hidden relationships between output values and classify them into different classes [2][4]. In absence of any dataset prior to interaction with real-time environment or statistical model of environment is unknown, then Reinforcement Machine Learning (RML or RL) is the only possible way, which is the third category of ML algorithms. RL algorithms train after the deployment of model into real-time environment by interacting with it. RL algorithms calculate the reward at each possible solution by searching the entire solution space, finally provide the solution with maximum reward as optimal solution [5][6].

Many advanced and complex ML algorithms with high performance are developed for various applications. People all around the world are generating huge datasets with the help of advanced technology and sensors. Training ML models with these datasets is highly time-consuming procedure and requires huge computational and storage resources. Continual training solves this huge storage requirement problem [7]. In continual training, the model keeps the knowledge of previous training during the future training without the need of storage of data used for previous training. This technique rapidly reduces the requirement of storage in SML algorithms. In case of RL algorithms, Continual learning plays very crucial role in improving the reward and decreasing the training time period. Using the continual training technique, the RL models are initially training in simulated or imitated environment before deployment into real environment. In this initial training, model is tuned for long duration with small step size for maximum reward. After initial training the model is deployed into real environment and setup for continual training to understand the real environment. As the model is pre-trained with similar environment, it can easily tune up with real environment and provides the best reward quickly [7][8].

II. OBJECTIVES AND GOALS

In Reinforcement Learning, initial learning phase of continual learning is very important to tune the system with simulated environment. After initial training, as the system is deployed into real environment, the training overload should be very low in order to minimize the response time and maximize the throughput. Even though many advanced and complex RL algorithms are developed [9], they are not suitable for continual training of ML models with tight timing constraints. More precisely, to take the full advantage of continual training in RL, different algorithms can be used at different phases of continual training with required benefits. For initial training, complex RL algorithms can be used for better performance and during the further continual training one can go for simple and quick RL algorithms to meet the timing constraints of applications. Existing RL algorithms are not compatible with one other to change during the continual learning.

III. METHODOLOGY

To take the complete advantage of continual learning, new RL algorithms needs to be developed with internal compatibility to transit from one algorithm to another algorithm during the continual training. Establishing this compatibility among existing complex algorithms is very difficult procedure and requires major modifications which will affect the performance of the algorithm. To solve this issue, two new Segmented Reinforcement Machine Learning (SRML) algorithms are proposed in this paper with different complexities. Both have the same hyperparameters and high-level segmentation structure. But their internal learning techniques are different which vary the complexity and training period of the algorithm. First algorithm is Segmented and Recursive Reinforcement Machine Leaning (SRRL) algorithm and second algorithm is Segmented and Adaptive Reinforcement Machine Leaning (SARL) algorithm [10][11]. SRRL algorithm first segments the entire solution space then recursively search for best reward in each segment in multidimensional solution space with gradient step. In SRRL, the step size is decreasing gradient but uniform during each epoch and the number of iterations in each epoch is also constant during the entire training procedure. SARL algorithm has adaptive nature in its internal search policy to break the epoch as well as jump from one branch to another inside the conditional search policy for precise tuning towards the maximum reward towards the maximum reward. In SARL also, the step is decreasing gradient but uniform in each epoch. The main advantage of SRRL algorithms is identification of global maximum in environment with a greater number of local maxima in lesser train period comparing to other existing algorithms. The SARL algorithm provides quick and precise maximum reward but its performance depends on the hyper parameters. A hybrid continual learning methodology is proposed by combining the advantages of both SRRL and SARL algorithms. In this methodology, proposed SRRL algorithm is used for initial training of the RL algorithm for better understanding of environment and identification of all local maxima and global maximum. After initial training, the SRRL is replaced with SARL by keeping the knowledge of initial training in terms of tuned hyper parameters and coefficients. As show in fig.1 this is very simple and straight forward procedure as both SRRL and SARL has similar outer layer architecture and same hyper parameters. After the initial training and during the real time RL training SARL offers very minimal learning overload with its quick training time and least storage requirements.

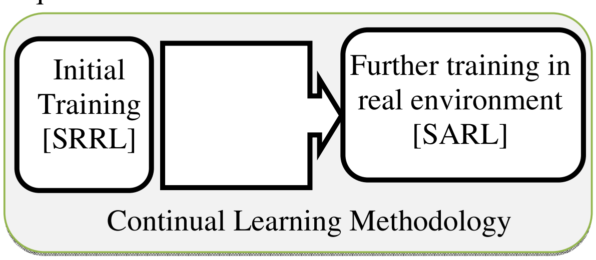

Fig. 1 Proposed continual learning methodology with SRML algorithms

Wavelet based image noise reduction application is considered to train and validate proposed continual learning methodology with proposed SRML and existing Markov Decision Process (MDP) based RL algorithms. This application has three tuneable coefficients in calculation of Hybrid Thresholding Factor (HTF) calculation which effects the performance of noise reduction [12]. RL algorithms are incorporated into HTF calculation to get the maximum performance of noise reduction in terms of PSNR.

IV. RESULTS AND DISCUSSIONS

The existing MDP based RL and proposed SRML algorithms are developed using the MATLAB-R2020b. The continual learning architecture and image denoising application also simulated using the MATLAB software. The computational resource is a Laptop with Intel core-i7 Processor, 32GB DDR5 RAM, NVIDIA GEFORCE GTX 1660-Ti 8GB GPU card. Standard Lena image is considered image data and 3 different noise models (Gaussian, Speckle and Salt & Pepper) are taken at 10 random variance level (0.2,0.35,0.45,0.1,0.04,0.07,0.15,0.25,0.3,0.2). In each noise model, the RL model is initially trained with first variance level (considered as simulated environment) and the remaining 9 variance cases used for continual training (considered as real environment). Standard MDP based RL algorithm is considered to compare the performance of SRML algorithms in continual learning.Fig.2 shows the optimal tuning points of all three coefficients (C_f - Feedthrough coefficient, C_d – Direct coefficient, C_f – gain coefficient) in the calculation of HTF in Gaussian environment. The tuning using SRML is very smooth compared to tuning from MDPR at all test instances of continual learning. Along with smooth tuning proposed SMRL based hybrid continual Learning provides better reward compared to existing continual learning models with very quick training periods as shown in fig.3.

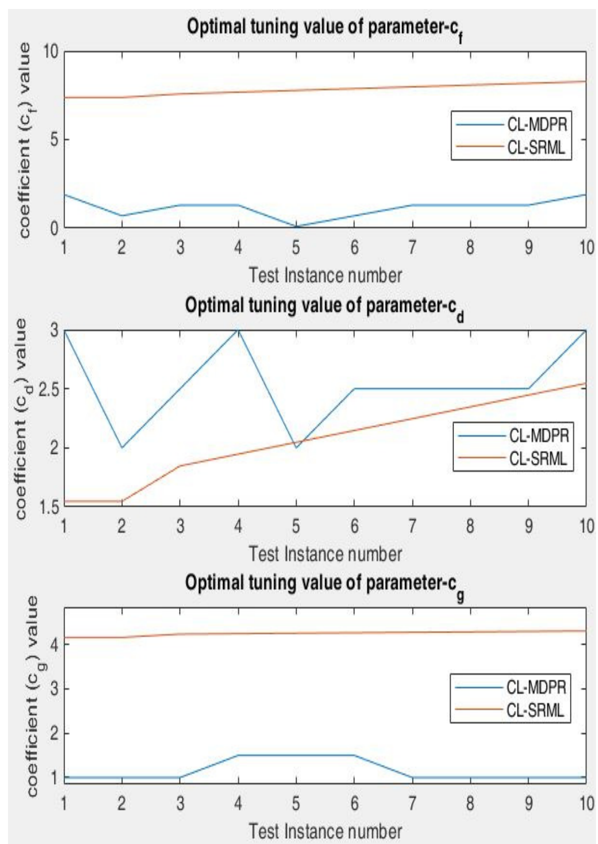

Fig. 2 Optimal Coefficient values instance with best reward at each test in Gaussian noise environment

Along with best reward, the average reward analysis and worst reward analysis are also very important and provide the consistence of the models. The worst-case rewards from proposed model in all test instance are very high compared to rewards from existing RL models shows the consistence of proposed models.

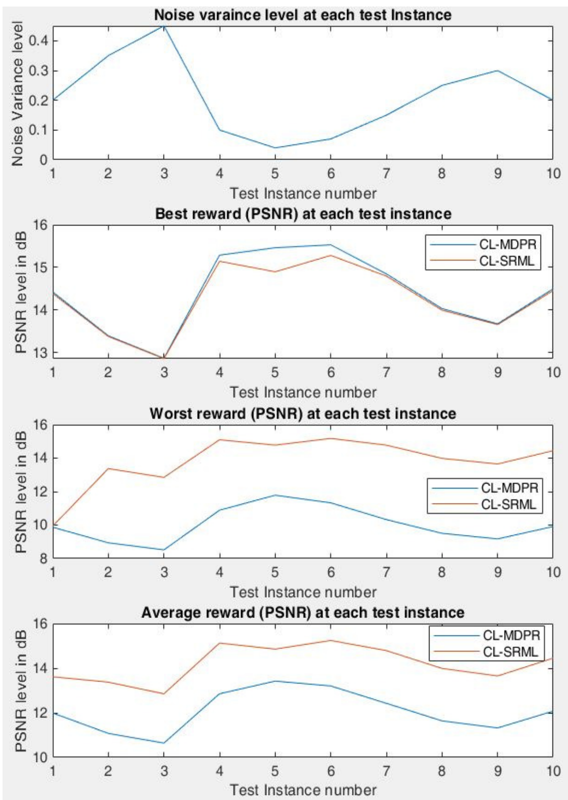

Fig. 3 Noise level, Best reward (PSNR), worst reward (PSNR) and Average reward (PSNR) at each test instances (From top to bottom) in Gaussian noise environment.

Both existing and proposed continual learning models are again employed in Salt and Pepper noise environment. The coefficient tuning is shown in fig.4 and rewards are compared in fig.5. The proposed model consistently providing reward near optimal point and provided best reward in high noise condition. The worst-case rewards and average rewards from proposed models are higher than MDPR based continual model. Both the models are trained and validated in speckle environment and coefficient tuning is shown in fig.6. Best case performance from proposed model are near (~1dB) optimal solutions from the existing model but average and worst-case rewards from proposed model are better than the existing model as shown in fig.7.

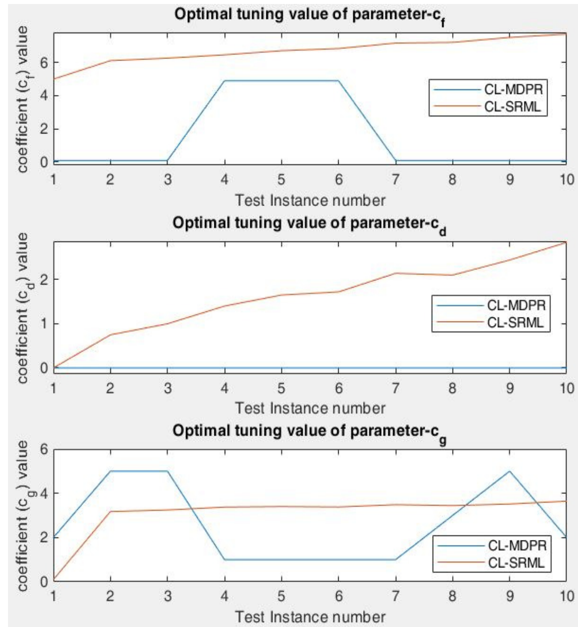

Fig. 4Optimal Coefficient values instance with best reward at each test in Salt & Pepper noise environment

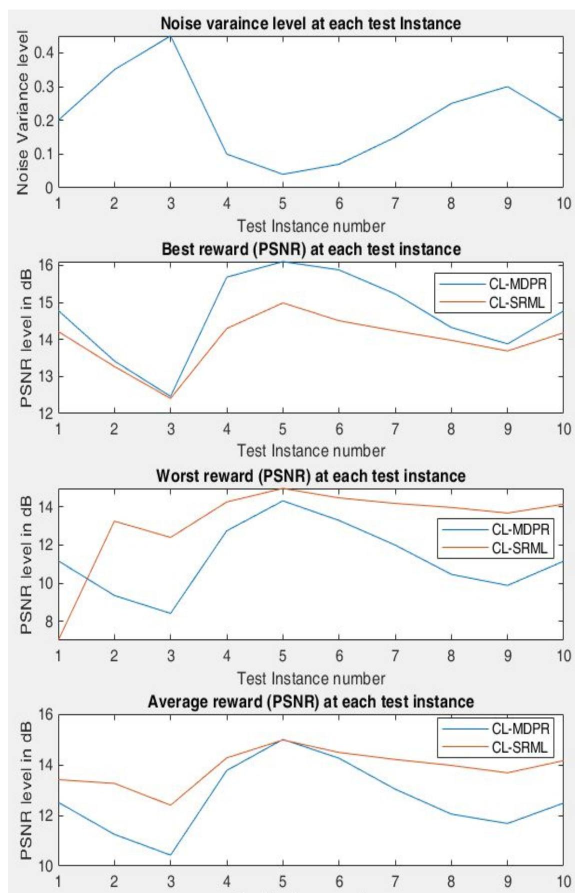

Fig. 5 Noise level, Best reward (PSNR), worst reward (PSNR) and Average reward (PSNR) at each test instances (From top to bottom) in Salt & Pepper noise environment.

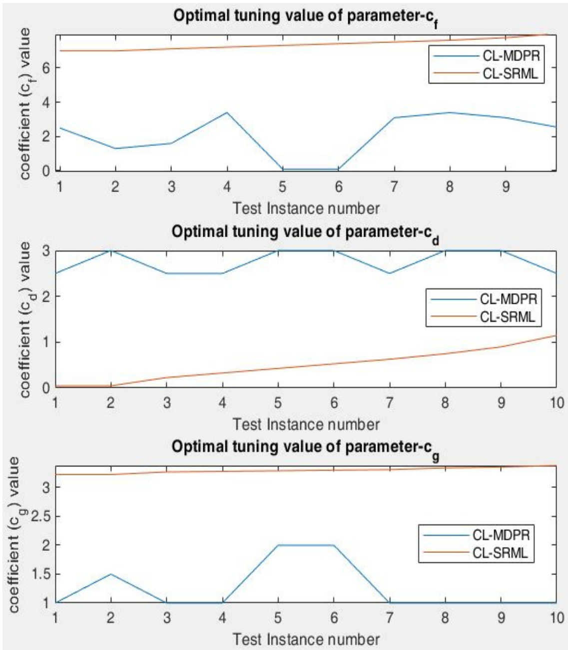

Fig. 6Optimal Coefficient values instance with best reward at each test in Speckle noise environment

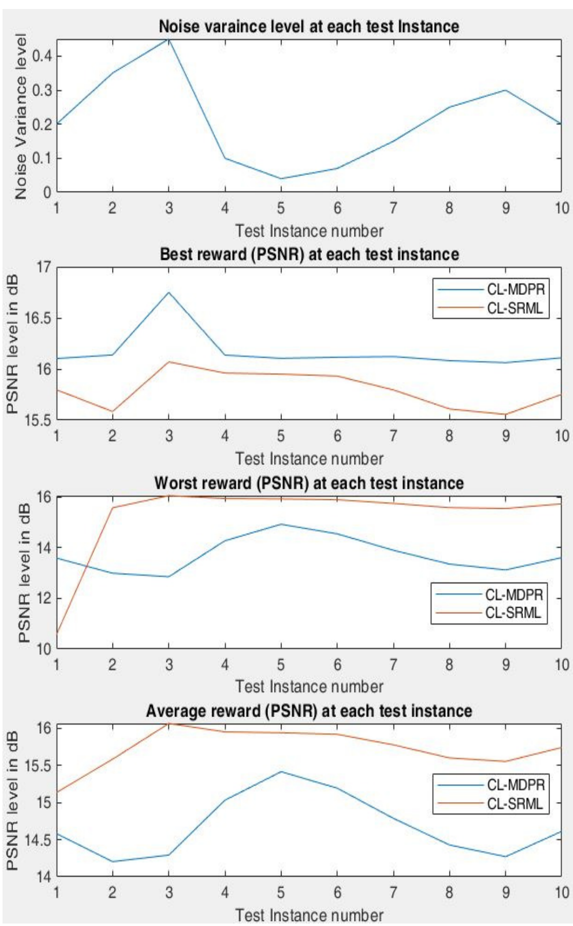

Fig. 7 Noise level, Best reward (PSNR), worst reward (PSNR) and Average reward (PSNR) at each test instances (From top to bottom) in Speckle noise environment.

TABLE I
COMPARISON IN TERMS OF COMPLEXITY AND REWARD

Test Instance No.	Complexity and reward Values					
	Iterations		Time		Reward (PSNR)	
	CL - MDPR	CL - SRML	CL - MDPR	CL - SRML	CL - MDPR	CL - SRML
1	1071	124	39.53	4.538	14.78	14.81
2	1071	024	39.90	0.896	13.41	13.26
3	1071	024	40.11	0.882	12.46	12.48
4	1071	024	40.78	0.998	15.68	14.29
5	1071	024	40.54	0.902	16.10	14.98
6	1071	024	44.80	0.961	15.87	16.50
7	1071	024	43.14	1.002	15.23	15.23
8	1071	024	41.61	0.888	14.32	14.97
9	1071	024	41.98	0.876	13.88	13.98
10	1071	024	41.69	0.960	14.77	14.78

From the above training and validation analysis of both existing and proposed models in different analysis, it is evident that the proposed continual learning model with SRML algorithms consistently provides optimal or near optimal rewards. The main advantage of proposed model is its quick training time with controlled complexity. For each test case the existing MDPR based continual training model takes 40 seconds to provide optimal solution whereas the initial training of SRML based continual learning model with SRRL algorithm takes less than 5 seconds and further continual learning with SARL algorithm takes less than 1 second to provide its best reward. The savings in training times is due to complexity-controlled structure of SRML algorithms

V. CONCLUSIONS

The existing and proposed continual models are trained for parameter tuning and detailed comparative analysis is performed on their rewards and training periodsin three different environments. In all the scenarios the proposed SRML based continual training model provided optimal or near optimal reward with 87.5% savings in initial training and 97.5% savings in further training in continual learning. The proposed model exhibits high degree of consistency and stability with higher worst-case and average reward in all test scenarios and noise environments. MDP based RL algorithms search the specified solution space in 1071 iteration and tune parameters with one decimal point precision where as SRRL searches the same solution space within 124 iterations.

REFERENCES

[1] John D. Kelleher; Brendan Tierney,“MACHINE LEARNING 101, in Data Science”, MITP,pp..97-150, 2018.

[2] E. Alpaydin, “Introduction to Machine Learning”, in MIT Press, Cambridge, Mass, USA, 2nd edition, 2010.

[3] Osisanwo F.Y., Akinsola J.E.T., Awodele O., Hinmikaiye J. O., Olakanmi O., Akinjobi J. "Supervised Machine Learning Algorithms: Classification and Comparison", in International Journal of Computer Trends and Technology (IJCTT) V48(3):128-138, June 2017.

[4] Alloghani, Mohamed & Al-Jumeily Obe, Dhiya & Mustafina, Jamila & Hussain, Abir & Aljaaf, Ahmed,“A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science”, 10.1007/978-3-030-22475-2_1, 2020.

[5] Lalit, Reema & Sharma, Nitin,“A Review on Methods of Reinforcement Machine Learning”, in International Journal of Advanced Science and Technology. 29. 1826-1831, 2021.

[6] M. A. -M. Khan et al., "A Systematic Review on Reinforcement Learning-Based Robotics Within the Last Decade," in IEEE Access, vol. 8, pp. 176598-176623, doi: 10.1109/ACCESS.2020.3027152, 2020.

[7] S. Jia, L. Shen and H. Xue, "Continuous-time Markov decision process with average reward: Using reinforcement learning method,"34th Chinese Control Conference (CCC), Hangzhou, China, pp. 3097-3100, doi: 10.1109/ChiCC.2015.7260117, 2015.

[8] H. Li, P. Barnaghi, S. Enshaeifar and F. Ganz, "Continual Learning Using Bayesian Neural Networks," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2020.3017292, 2020.

[9] E. M. Perez, A. Majumdar, P. Benavidez and M. Jamshidi, "Implementation of Reinforcement Learning Simulated Madel on Physical UGV Using Robot Operating System for Continual Learning," 14th Annual Conference System of Systems Engineering (SoSE), Anchorage, AK, USA, pp. 212-217, doi: 10.1109/SYSOSE.2019.8753801, 2019.

[10] Modalavalasa Hari Krishna, Dr.Makkena Madhavi Latha "Intelligent Parameter Tuning Using Segmented Recursive Reinforcement Learning Algorithm" International Journal of Engineering Trends and Technology vol.68, issue:6,pp:1-8, 2020.

[11] Modalavalasa, Hari Krishna and Madhavi Latha Makkena. “Intelligent Parameter Tuning Using Segmented Adaptive Reinforcement Learning Algorithm.” 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (2020): 252-256., 2020.

[12] M. Hari Krishna, G. Sateesh Kumar (2015), A Noise Based Hybrid Thresholding For Enhanced Noise Reduction Using Double Density Dual Tree Complex Wavelet Transform, 13th International Conference on Electromagnetic Interference and Compatibility (INCEMIC), doi : 978-1-5090-5350-6/15, 2020.

