**Antony Gnanaprakasam DARWIN JOSE RAJU, PhD (corresponding author)**
darwinraju.eee@gmail.com
Department of Electrical and Electronics Engineering, St. Xavier's Catholic College of Engineering, Chunkankadai, Tamil Nadu, India

**Sankaran SOLAI MANOHAR, PhD**
solaimano@gmail.com
Department of Electrical and Electronics Engineering, KCG College of Technology, Karapakkam, Chennai, Tamil Nadu, India

**Tharmaraj Charlet JERMIN JEAUNITA, PhD**
jerminjeaunita@lites.edu.in
Centre for Research in Intelligence and Computing, Rooks and Brooks Technologies Private Limited, Nagercoil, Tamil Nadu, India

# Mitigating Catastrophic Forgetting in Imitation Learning for Embodied AI using Progressive Neural Networks

**Abstract.** *Imitation learning is crucial for agent training to replicate expert behaviour, but it faces issues like catastrophic forgetting, where the model accommodates new tasks but forgets previously learned information. The research aims to develop a Progressive Neural Network and Elastic Weight Consolidation (PNN-EWC) approach that mitigates forgetting while enhancing task retention and adaptation. The PNN-EWC model is experimented with in the Webots simulation environment utilising the Franka Emika Panda robotic arm. This model integrates Progressive Neural Networks, which add new columns to the neural network for each task, and Elastic Weight Consolidation regularises updates to preserve learned tasks. The proposed PNN-EWC model achieved a 95% average success rate and a 6% forgetting rate in preventing catastrophic forgetting, highlighting significant improvements over traditional methods. This approach allows the Franka Emika Panda to continuously learn new tasks without losing previously acquired skills, making it effective for dynamic environments.*

**Keywords**: *catastrophic forgetting, elastic weight consolidation, embodied ai, imitation learning, progressive neural network.*

**JEL Classification:** O0, O1, O4.

| **Received: 9 June 2025** | **Revised: 6 September 2025** | **Accepted: 10 September 2025** |
|---|---|---|

## 1. Introduction

Recent developments in Embodied Artificial Intelligence (EAI) systems have enhanced the capabilities of AI agents by enabling them to learn based on their interactions with the environment (Duan et al. 2022; Banerjee et al. 2024). As more humanoid AI systems become prevalent in today's technologically driven world, the

DOI: 10.24818/18423264/59.3.25.06

role of ensuring cooperative behaviour among agents also increases. These cooperative humanoid AI systems focus on algorithms and strategies that allow multiple agents to learn how to collaborate, adapt, and make decisions within a shared environment. Each agent makes decisions at each time step and coordinates with others to achieve its goals (Tang et al. 2023, Oroojlooy and Hajinezhad (2023)). They gradually collaborate to do tasks and communicate with one another through the radio. However, many scenarios were not anticipated before deployment in real-world applications. Therefore, the robots must make plans based on their experience Orr and Dutta (2023). Imitation learning is a technique that involves action cloning, where agents replicate the state and action trajectory to learn from demonstrations. This method has gained popularity in developing predictive models of agent behaviours, with growing interest in imitation learning (Zare et al., 2024). However, the challenge faced is the ability to generalise across multiple tasks without forgetting the previous one. When new tasks are introduced, traditional models are fine-tuned to adapt, but this process leads to catastrophic forgetting, in which the model loses its understanding of previously learned tasks (Alammar et al., 2024).

Imitation learning-based online trajectory guidance systems utilising expert-like movement trajectories provide novice surgeons with intra-operative trajectory guidance and perform manipulation similarly to experts Chen and Fan (2025). Instead of explicit programming, bilateral control-based imitation learning utilises human demonstrations to achieve human-level motion speeds with environmental adaptation (Yamane et al. 2023). Another innovative approach involves learning a shared latent space representation for communication between humans and robots, allowing robots to generate action based on demonstrations reactively Prasad et al. (2024). While imitation learning has shown a promising solution for enabling humanoid AI systems to mimic expert behaviours, existing methods have struggled with catastrophic forgetting when learning new tasks. The current approach has no solid mechanism to retain older knowledge and facilitate quick adaptability to changing tasks in dynamic environments. To address these limitations, the proposed work introduces an innovative approach to mitigate catastrophic forgetting and enhance performance in dynamic environments by facilitating rapid learning and task retention.

### *1.1 Research Contribution*

The following is a summary of our proposed work's primary contributions:

- The proposed PNN-EWC model combines Progressive Neural Network (PNN) and Elastic Weight Consolidation (EWC) to prevent catastrophic forgetting and allow continuous learning in imitation learning. This enables the system to retain the knowledge gained prior to the new adaptation process without forgetting previous tasks in changing, real-time environments.
- Using PNN, the model dynamically adapts the network by introducing new columns to include new tasks without losing previously learned knowledge. This

enables the agents can handle sequential and changing tasks effectively without performance degradation.

- To preserve critical weights stable, the proposed model includes EWC, which is used to regularise weight updates for the learned tasks in earlier columns. This helps to maintain essential knowledge of past tasks while the model learns new tasks.
- The proposed model makes use of an Auction-Based Task Allocation (ABTA) approach to enhance task specialisation and coordination among agents. The approach ensures that agents are effectively assigned tasks based on their abilities, which improves their capability to execute complex roles in multitasking scenarios.

The remaining part of the paper is organised as follows: Section 2 examines the existing research on the imitation learning process, Section 3 outlines the proposed PNN-EWC framework's workflow, Section 4 discusses and analyses the findings, and Section 5 concludes the paper.

## 2. Related Work

This section reviews and analyses previous research and developments related to imitation learning. (Dey et al. 2023) developed a Reinforcement Learning (RL) method for policy transfer in building control using imitation learning. This method enhanced a rule-based policy, reduced training time, and minimised unstable early exploration behaviour. However, there was considerable conflict when it came to complex building environments. To explore learning in complex behaviours, (Wan et al. 2023) employed an innovative end-to-end method utilising a neural network to develop a system for the autonomous flight of multiple drones. The model demonstrated improved performance in terms of single-point failure and scalability, successfully enabling the cooperative motion of UAVs. Even though imitation learning demonstrated successful dynamic adaptation, it could still be improved in more complex and uncertain environments. A similar problem was identified in a study by (Han et al. 2023) which combined RL and imitation learning to enable robots to play beach volleyball in a 3D environment. The robots outperformed the conventional RL approach, achieving a higher score in the Elo rating system. However, the model's high complexity could limit its scalability in more complex environments.

Sun and Kim (2023) introduced a data-driven network traffic simulation framework using Multi-agent Generative Adversarial Imitation Learning, which directly learns traffic behaviours from observed vehicle trajectory data. This model successfully mimicked real-world vehicle movement and identified state change patterns, still enhancing the model's scalability to handle large networks and more complex network traffic scenarios. To overcome this, Sun and Kim (2024) employed multi-agent imitation learning to develop a simulation model for an unsignalised junction utilising two-dimensional data.

They concentrated on employing multi-agent adversarial inverse reinforcement learning to simulate the paths of vehicles. The model demonstrated superior accuracy in generating trajectories for vehicles moving in a straight line, but was less effective for vehicles turning. Similarly, (Li et al. 2024) suggested Generative Adversarial Imitation Learning Policy Gradients (GAILPG) to enhance the utilisation of experience and the capacity for exploration of agents in multi-agent RL. This method outperformed the related advanced policy-based and value-based methods. However, trained discriminators struggled to evaluate and judge the agent's actions properly.

One of the most challenging problems in imitation learning is enabling agents to learn new skills continuously while maintaining an understanding of previously learned ones and adapting to dynamic, real-time environments. This requires balancing the need to learn quickly with the retention of long-term knowledge to deal with sequential and changing task requirements. Existing approaches often fall short in maintaining performance across tasks due to catastrophic forgetting, which limits their scalability in complex settings. To address these challenges, we propose a PNN-EWC model that integrates methods to expand the size of the neural network, allowing it to learn new tasks without compromising its prior capabilities, and retains critical weights to prevent catastrophic forgetting. This method supports rapid learning and task retention, hence enhancing adaptability and performance in changing environments.

## 3. Proposed Methodology

The proposed PNN-EWC framework is illustrated in Figure 1. The proposed methodology is implemented in a simulation environment that creates dynamic conditions. Initially, imitation learning is employed using a behavioural cloning approach, where agents learn from expert demonstration by mapping states to actions. As new tasks are introduced, the agent knows how to perform them efficiently. To prevent forgetting, the framework utilises PNNs that introduce a new column in the neural network for every new task, while maintaining lateral connections to previously learned features. EWC selectively controls the weight updates and retains essential knowledge. The method enables effective learning, adaptability, and long-term performance in humanoid AI in dynamic, challenging environments.

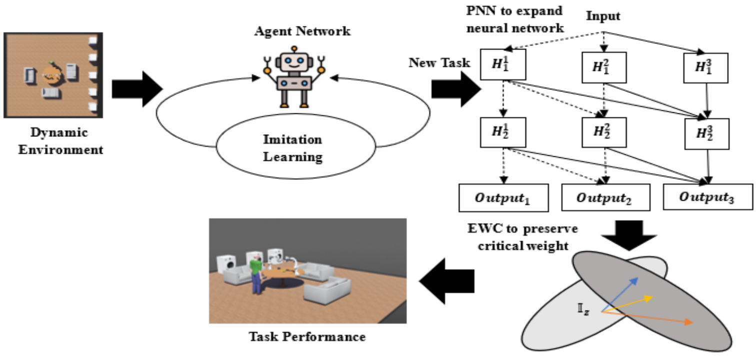



**Figure 1. The Proposed PNN-EWC Framework enhancing task retention and adaptation in imitation learning**
*Source:* Authors' own creation.

### *3.1 System Setup*

The proposed PNN-EWC framework is implemented in a dynamic environment using a Webots simulation environment with a Franka Emika Panda robot. In this environment, an agent needs to learn and adapt to multiple tasks sequentially. The environment is designed to mimic real-world task transitions, requiring the agent to retain prior knowledge while incorporating new learning objectives.

### *3.2 Imitation Learning*

The agent initially learns tasks from expert demonstrations using imitation learning. One important learning strategy in autonomous behavioural systems is imitation learning, which aims to replicate human behaviour or an agent's performance in a particular task. Using this method, learning observations are translated into actions. It reduces the effort required to instruct an agent by demonstrating the actions needed to complete a specific task (Dong et al., 2024). The proposed model utilises the behavioural cloning method, a direct mapping of states $\kappa_t$ to the control inputs or actions $\mathrm{A}_t$ as illustrated in Equation (1).

$$\mathrm{A}_t = \pi_{\vartheta_{\mathrm{A}}}(\kappa_t) \tag{1}$$

A supervised learning approach can be used to learn the policy from expert-demonstrated trajectory data, $d = [\kappa_t, \mathrm{A}_t]$. To learn the mapping from the states to the actions, a neural network with a parameter set $\vartheta_{\mathrm{A}}$ is employed.

### *3.3 Progressive Neural Network*

Once the agent successfully learns an initial task, a new task is introduced through additional expert demonstrations. As new tasks are introduced, the agent adapts to handle new tasks while preserving previously learned tasks. Continuous learning is the main goal of an agent system. By constructing a neural network column for each task being solved, progressive networks prevent catastrophic forgetting by incorporating these needs directly into the model architecture. At the same time, lateral connections to previously learned column properties facilitate transfer (Meng et al. 2024). A PNN starts with a single column, whose topology is identical to that of a deep neural network, and consists of $N$ layers, $H_a^1$ hidden activations, $n_a$ units at each layer and convergence parameters $\theta^1$. When moving on to a second task, the parameters $\theta^1$ are frozen, and a new column containing parameters $\theta^2$ is concatenated. In this scenario, the layer $H_a^2$ receives input from both $H_{a-1}^2$ and $H_{a-1}^1$ through lateral connections. This generalisation to $M$ task is shown in Equation (2).

$$H_a^M = \mathcal{F}\left(w_a^M H_{a-1}^M + \sum_{b<M} L_a^{M:b} H_{a-1}^b\right) \tag{2}$$

Where $w_a^M$ represents the weight matrix of the column $M$'s layer $a$ , $L_a^{M:b}$ represents the lateral connection from the column $b$'s layer $a-1$ to column $M$'s layer, $H_0$ is the network input, $\mathcal{F}$ represents an element-wise non-linearity. The PNN architecture with $M=3$. However, previously learned tasks may degrade over time without proper weight regulation. To mitigate these issues, EWC can be integrated with PNN. EWC helps preserve important weights by selectively restricting updates to prevent catastrophic forgetting, while PNN expands the network without any prior understanding to handle new tasks. This improves the model's ability to learn sequential tasks in changing environments.

### *3.4 Elastic Weight Consolidation*

The proposed model utilised EWC, which complements PNN for selective regularisation of neural network parameters to prevent catastrophic forgetting in neural networks during lifelong continuous training. The EWC maintains the network components relevant to the prior task while modifying the entire network for subsequent tasks. The EWC approach retains previously learned knowledge in sequential learning by incorporating a regulariser into the loss function (Aslam et al. 2025). It prevents the most significant weights from deviating substantially from the consolidated values throughout the learning of subsequent tasks, as shown in Equation (3).

$$\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{L}_c} + \frac{\tau}{2}\sum_z \mathbb{I}_z (W_z - W_z^*)^2 \tag{3}$$

Where $\boldsymbol{\mathcal{L}_c}$ represents the loss function of training c, $W_z^*$ and $\mathbb{I}_z$ represents the weight parameter and importance of $z$-th weight of the neural network following training of the previous tasks. The regularising component will provide $-\tau\mathbb{I}_z(W_z - W_z^*)$ to the anti-gradient. The resistance to change each weight will, therefore, be proportional to its relevance and the hyperparameter $\tau$ when training with the gradient method.

The neural network will learn the current task during sequential learning if $\tau$ is small, but it will forget the knowledge from previously learned tasks more quickly. Conversely, the network will have a high resistance to changing weights and will be able to retain the prior knowledge acquired on previous tasks if $\tau$ is too large, but its learning rate on the current task may be inadequate.

## 4. Results and Discussion

The proposed PNN-EWC framework for the dynamic environment was executed in the Webots simulation environment on a machine with Windows 10 Pro, Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40 GHz. The result section includes the performance analysis and comparison analysis of the proposed model to assess its effectiveness.

### *4.1 Environmental setup*

The proposed PNN-EWC framework creates a dynamic environment in the simulation environment using Webots. Webots is an open-source robotics simulation software. The environmental setup is depicted in Figure 2. In the Webots simulator, five distinct tasks were designed for the Franka Emika Panda, which is depicted in Table 1. Franka Emika Panda is a robot, agile as a human arm with a human touch sense, easy to set up, and intuitive to use. The Franka Emika Panda robot has classical stiffness, pose repeatability of 0.1mm, and negligible path deviation even at high velocities of up to 2m/s. An Intel Realsense camera is attached to the robot end-effector. This enables the precise, robust, and rapid execution of processes. This robot is equipped with link-side torque sensors in all 7 axes.

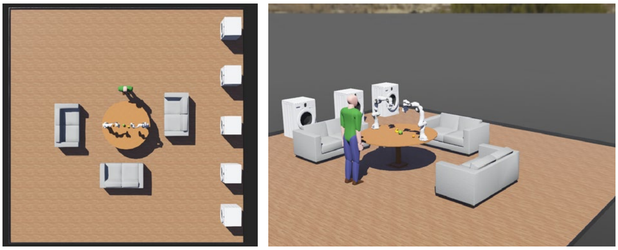


**Figure 2. Visualisation of the dynamic environment**
*Source:* Authors' own creation.

**Table 1. Task Performed by Franka Emika Panda**

| Task | Action |
|---|---|
| Task A | Pick light object |
| Task B | Stack Boxes |
| Task C | Screw Bolt |
| Task D | Arrange Part |
| Task F | Pick heavy object |

*Source:* Authors' processing.

### *4.2 Evaluation Metrics*

Performance evaluation of the PNN-EWC framework requires several evaluation metrics to assess its performance. These metrics ensure the generalisability, robustness, and dependability of the model. Performance metrics include success rate, retention score, convergence time, and forgetting rate. These metrics offer distinct insights into the model's strengths and weaknesses.

Success Rate (SR): Measures the percentage of completed tasks.

$$SR = \left(\frac{Number\ or\ successful\ task\ completions}{Total\ tasks\ attempted}\right) \times 100 \quad (4)$$

Retention Score (RS): It quantifies how well the model retains knowledge from previously learned tasks.

$$RS = \left(\frac{\sum_{i=1}^{N} A_{new}(i)}{\sum_{i=1}^{N} A_{old}(i)}\right) \times 100 \quad (5)$$

Convergence Time (CT): Measures the time taken for the model to reach a stable learning state. It is often computed based on the number of iterations required to achieve an accuracy threshold.

$$CT = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{E_t} \quad (6)$$

Forgetting Rate (FR): Measures the extent of performance degradation in previous tasks upon the acquisition of new ones.

$$FR = \left(\frac{\sum_{i=1}^{N}(A_{old}(i) - A_{new}(i))}{N}\right) \times 100 \quad (7)$$

Where $A_{new}(i)$ is the accuracy of the task $i$ after learning a new task, $A_{old}(i)$ is the accuracy of the task $i$ immediately after its initial learning, $N$ represents the total number of previous tasks, $T$ is the total number of tasks, and $E_t$ represents the number of epochs needed for the task $t$ to reach accuracy.

### *4.3 Evaluation of the Proposed Model's Performance*

Table 2 illustrates the performance of the proposed PNN-EWC approach in mitigating catastrophic forgetting and improving task adaptation. The proposed model demonstrates outstanding performance, achieving a success rate of 95% in most tasks. A 78% retention score indicates strong knowledge preservation allowing the robot to maintain prior understanding while adapting to new tasks, and a 92% convergence time reflects fast adaptation in a dynamic environment. Furthermore, the forgetting rate is only 6%, which helps prevent knowledge degradation, ensuring stable and continuous learning.

**Table 2. Performance Metrics of the Proposed PNN-EWC Approach**

| Metrics | Value (%) |
|---|---|
| Success Rate | 95 |
| Retention Score | 78 |
| Convergence Time | 92 |
| Forgetting Rate | 6 |

*Source:* Authors' processing.

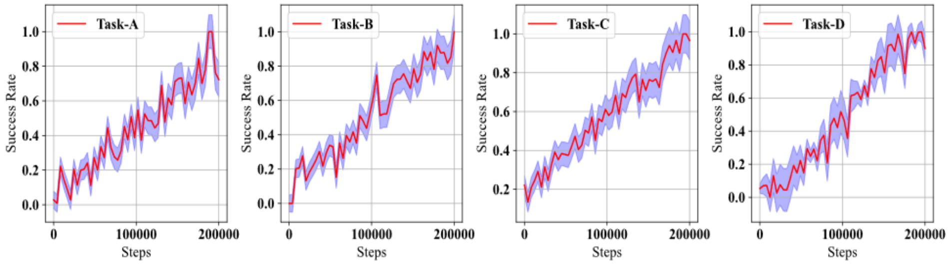



**Figure 3. Learning Curve for a Single Task**
*Source:* Authors' own creation.

To evaluate the performance of the PNN-EWC framework on complex tasks, we selected four manipulation tasks in the dynamic environment. Figure 3 illustrates the learning curve to determine the success rate for each task over time steps. Initially, the success rate is low, as the model has not fully adapted to the task, but as training progresses, the success rate steadily increases, indicating an effective learning process. This shows the proposed PNN-EWC framework, which enables robots to learn and train tasks without experiencing significant performance degradation.

Figure 4 shows the robotic system's success rate across five tasks of increasing mechanical complexity ranging from simple object picking to high-precision circuit assembly. Despite the rising complexity, the proposed PNN-EWC framework maintains a high success rate, dropping marginally from 95% to 80%. This proves the strength of the framework and its ability to generalise the learned behavior to more complex engineering processes, like torque-controlled insertion or path constraint execution, which are crucial in robotic engineering and assembly lines.

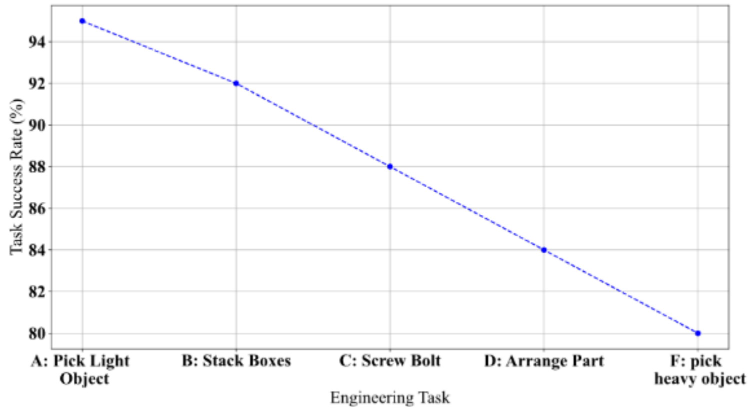



**Figure 4. Performance of the Engineering Task Complexity**
*Source:* Authors' own creation.

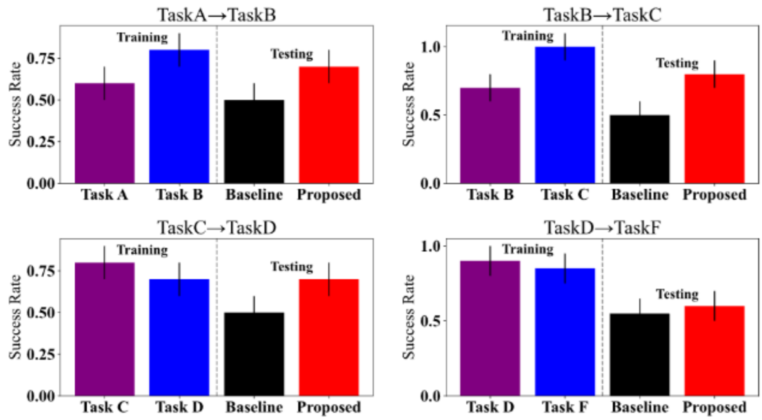



**Figure 5. Performance During Training and Testing**
*Source:* Authors' own creation.

Figure 5 illustrates the performance during training and testing as the Franka Emika Panda robot sequentially learns two manipulation tasks. The operational tasks include Task A, Task B, Task C, Task D, and Task F. Initially, two tasks are trained, the first task is shown with a purple bar, and the second task with a blue bar. During testing, the proposed model consistently outperforms the baseline in previously learned tasks. Across two-task learning sequences, the proposed model achieves success rates exceeding the baseline.

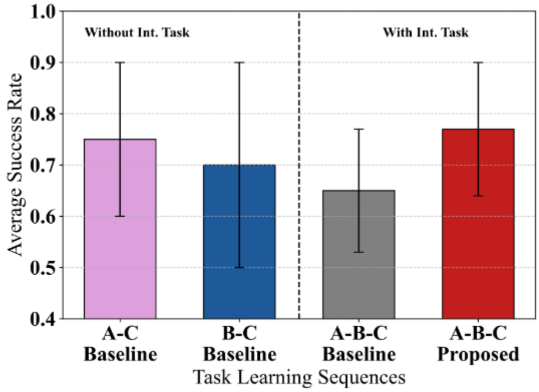



**Figure 6. Effect of Forgetting on Forward Transfer**
*Source:* Authors' own creation.

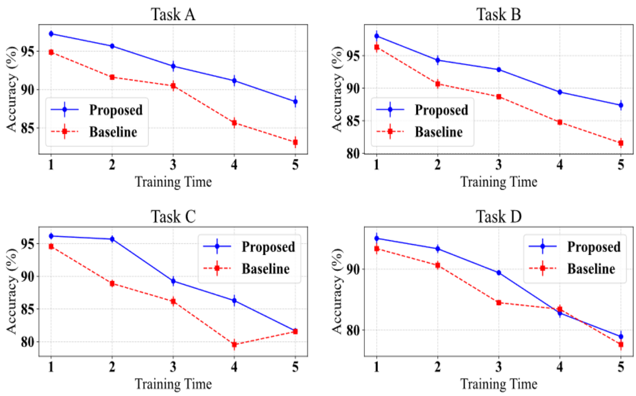



**Figure 7. Evolution Performance for Each Task in Terms of Accuracy**
*Source:* Authors' own creation.

Figure 6 illustrates the impact of forgetting on forward transfer. In the absence of forgetting, the robot arm, after continual learning on the sequence Task A →Task B→Task C should perform well as on the sequence Task A→Task C. However, the intermediate Task B diminishes this transfer. In this graph, A-C (Baseline), the average success rate is 0.76, whereas A-B-C, it drops to 0.64. This denotes a decrease in the average success rate due to the presence of Task B. Conversely, the proposed approach effectively mitigates the impact of forgetting on task transfer by achieving an average success rate of 0.78.

Figure 7 illustrates the performance evolution for each task in terms of accuracy over time, comparing the proposed PNN-EWC model with the baseline approach. The result indicates that PNN-EWC maintains higher accuracy across all functions, with a slower decline in performance as new tasks are introduced. This demonstrates how effectively the proposed model retains the knowledge gained from previous tasks.

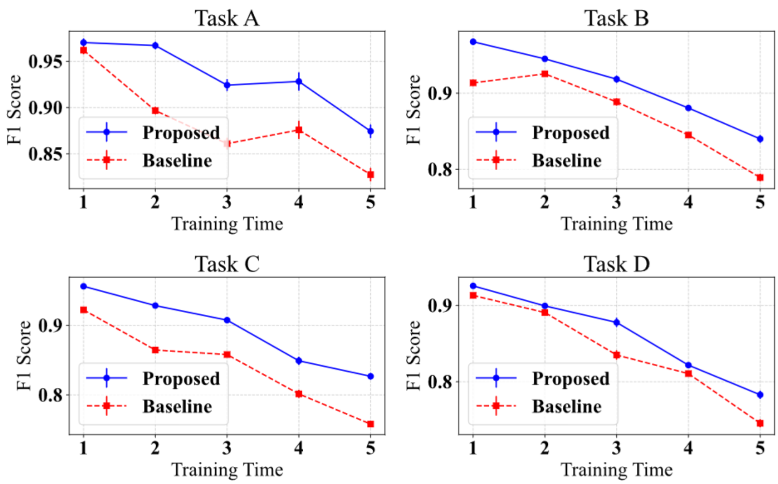



**Figure 8. Evolution Performance for Each Task in Terms of F1 Score**
*Source:* Authors' own creation.

Figure 8 illustrates the performance evolution for each task in terms of the F1 Score, which combines the precision and recall scores of each task, comparing the proposed PNN-EWC model with the baseline approach. The result indicates that the proposed PNN-EWC maintains a higher F1 Score across all tasks, confirming better knowledge retention and stability across sequential tasks.

Figure 9 illustrates the model's ability to retain information for previously learned tasks. The result shows that PNN-EWC achieves a significantly higher retention score than the baseline approach. As the robotic arm learns new tasks, PNN-EWC prevents the loss of previously acquired knowledge, while the baseline model struggles to remember previously learned tasks. This highlights the robust learning capability of PNN-EWC, making it suitable for long-term robotic learning scenarios.

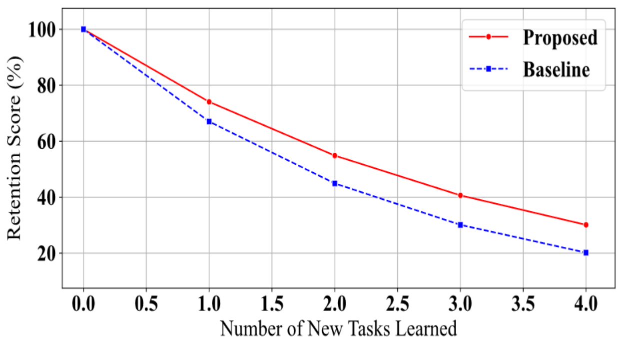
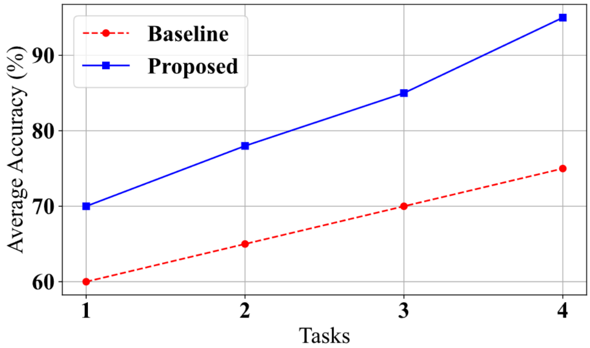
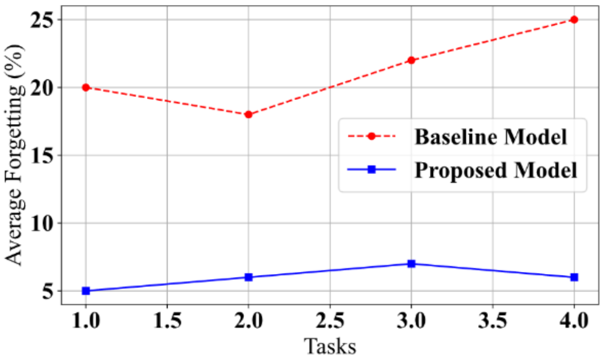



**Figure 9. Retention Score Evaluation on Task Learned**
*Source:* Authors' own creation.




**(a) Average Accuracy**




**(b) Average Forgetting Rate**

**Figure 10. Incremental Performance Evaluated on all tasks observed during Continual Learning in terms of (a) Average Accuracy, (b) Average Forgetting Rate**

*Source:* Authors' own creation.

Figure 10 illustrates the average incremental performance evaluated in all tasks after each task is completed. (a) Compare the accuracy at the end of all tasks and observe that the proposed PNN-EWC approach achieves higher accuracy than the baseline approach. (b) Assess the ability to prevent forgetting, which measures performance degradation in subsequent tasks. The results show that the proposed approach has a lower forgetting rate than the baseline approach.

Figure 11 shows the time required by the robotic agent to adapt to new tasks in a sequential learning scenario. Here, the adaptation time increases when transitioning to more complex tasks. However, the observed increase is gradual and within practical limits, confirming that the model enables real-time task switching without re-training. This makes the faster deployment of robotic systems in dynamic manufacturing environments with variable task assignments.

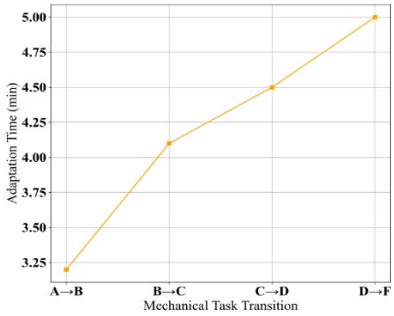



**Figure 11. Adaptation Time Across Mechanical Task Transitions**
*Source:* Authors' own creation.

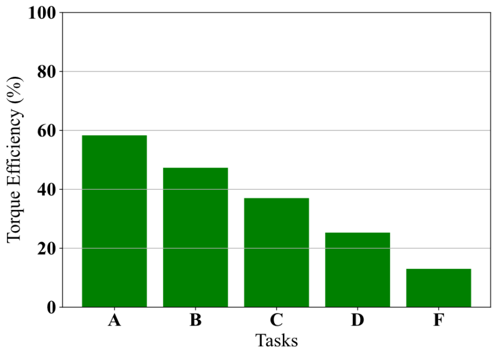



**Figure 12. Joint Torque Efficiency Across Tasks**
*Source:* Authors' own creation.

Figure 12 measures the torque efficiency of the robot arm for each task. Efficiency is computed relative to the robot's maximum torque capacity. The results show high torque efficiency (>70%) across all tasks, even as task complexity increases. This suggests that the control policies generated through the PNN-EWC model are energy-efficient and mechanically optimal, which minimises strain on the robot's joints.

### *4.4 Comparative Analysis*

In this section, we analyse the proposed model's effectiveness compared to traditional imitation learning and compare the success rates of the proposed and traditional imitation learning approaches.

Table 3 compares the proposed PNN-EWC framework's success rate against the existing Imitation learning methods, such as IL-RL, Visual hindsight self-imitation learning, and Visual imitation learning. The existing approaches achieve success rates of 94.4%, 92.1%, and 80%, respectively, while the proposed PNN-EWC framework achieves a success rate of 95%. This demonstrates the effective performance of the proposed model compared to existing imitation learning approaches.

**Table 3. Comparative analysis of the proposed PNN-EWC with existing approaches**

| Reference | Methods | Success rate (%) |
|---|---|---|
| Han et al. (2023) | IL-RL | 94.4 |
| Kim et al. (2024) | Visual hindsight self-imitation learning | 92.1 |
| Jonnavittula et al. (2025) | Visual imitation learning | 80 |
| Proposed Model | PNN-EWC | 95 |

*Source:* Authors' processing.

### *4.5 Discussion*

In this study, the performance of the proposed PNN-EWC framework is evaluated within the Webots simulation environment, demonstrating its effectiveness in mitigating catastrophic forgetting in imitation learning. The outcomes indicate that the proposed model achieves high accuracy and F1 Scores for multiple tasks, outperforming traditional imitation learning models, particularly in terms of a slower decline in performance as new tasks are introduced. The analysis of the learning curve reveals that PNN-EWC enables stable task adaptation with minimal performance loss even as new tasks are introduced. Expanding the network progressively can ensure that knowledge is retained, and weight disruption by EWC regularisation ensures a smooth transition from one task to another. Additionally, compared to the baseline approach, PNN-EWC exhibits superior forward knowledge transfer and a high retention score, making it well-suited for continuous learning. These findings confirm the potential of PNN-EWC for real-world robotic applications, in which sequential learning is crucial.

## 5. Conclusions

The proposed PNN-EWC framework successfully addressed the catastrophic forgetting problem in imitation learning, ensuring that the Franka Emika Panda robotic arm retains previously learned tasks while adapting to new tasks. The PNN dynamically expands the neural network by adding a new column for each task while maintaining lateral connections, thus preserving past knowledge. Meanwhile, the EWC enhances the stability of the network by regularising weight updates, preventing a significant deviation in crucial parameters. The experimental findings validate the performance of the proposed framework, achieving a 95% success rate, a 78% retention score, and a 92% convergence time, while maintaining a low 6% forgetting rate, which significantly outperforms baseline models. These findings demonstrate that the proposed PNN-EWC framework enhances learning stability, adaptability, and efficiency in continual learning scenarios. Although the current implementation is simulation-based, future work could explore real-world deployment and adaptive strategies for task recognition that do not rely on predefined task boundaries.

**References**


[1] Alammar, Z., Alzubaidi, L., Zhang, J., Li, Y., Gupta, A., Gu, Y. (2024), *Generalisable deep learning framework to overcome catastrophic forgetting. Intelligent Systems with Applications*, 23, 200415.

[2] Aslam, S., Rasool, A., Li, X., Wu, H. (2025), *Cel: A continual learning model for disease outbreak prediction by leveraging domain adaptation via elastic weight consolidation. Interdisciplinary Sciences: Computational Life Sciences*, 1-19.

[3] Banerjee, S., Paul, S., Roychoudhury, R., Bhattacharya, A., Sarkar, C., Sau, A., ..., Bhowmick, B. (2024), *Teledrive: An Embodied AI Based Telepresence System. Journal of Intelligent & Robotic Systems*, 110(3), 96.

[4] Chen, Z., Fan, K. (2025), *An online trajectory guidance framework via imitation learning and interactive feedback in robot-assisted surgery. Neural Networks*, 107197.

[5] Dey, S., Marzullo, T., Zhang, X., Henze, G. (2023), *Reinforcement learning building control approach harnessing imitation learning. Energy and AI*, 14, 100255.

[6] Dong, W., Zhang, F., Li, M., Fang, X., Yang, Q. (2024), *Imitation Learning Based Real-Time Decision-Making of Microgrid Economic Dispatch Under Multiple Uncertainties. Journal of Modern Power Systems and Clean Energy*, 12(4), 1183-1193.

[7] Duan, J., Yu, S., Tan, H. L., Zhu, H., Tan, C. (2022), *A survey of embodied ai: From simulators to research tasks. IEEE Transactions on Emerging Topics in Computational Intelligence*, 6(2), 230-244.

[8] Han, Z., Liang, Y., Ohkura, K. (2023), *Developing multi-agent adversarial environment using reinforcement learning and imitation learning. Artificial Life and Robotics*, 28(4), 703-709.

[9] Jonnavittula, A., Parekh, S., Losey, D. (2025), *View: Visual imitation learning with waypoints. Autonomous Robots*, 49(1), 1-26.

[10] Kim, K., Lee, M., Lee, M. W., Shin, K., Lee, M., Zhang, B.T. (2024), *Visual hindsight self-imitation learning for interactive navigation. IEEE Access*.

[11] Li, W., Huang, S., Qiu, Z., Song, A. (2024), *GAILPG: multi-agent policy gradient with generative adversarial imitation learning. IEEE Transactions on Games*.

[12] Meng, W., Ju, H., Ai, T., Gomez, R., Nichols, E., Li, G. (2024*), Transferring meta-policy from simulation to reality via progressive neural network. IEEE Robotics and Automation Letters.*

[13] Oroojlooy, A., Hajinezhad, D. (2023), *A review of cooperative multi-agent deep reinforcement learning. Applied Intelligence*, 53(11), 13677-13722.

[14] Orr, J., Dutta, A. (2023), *Multi-agent deep reinforcement learning for multi-robot applications: A survey. Sensors*, 23(7), 3625.

[15] Prasad, V., Kshirsagar, A., Koert, D., Stock-Homburg, R., Peters, J., Chalvatzaki, G. (2024), *Moveint: Mixture of variational experts for learning human–robot interactions from demonstrations. IEEE Robotics and Automation Letters*, 9(7), 6043-6050.

[16] Sun, J., Kim, J. (2023), *Toward data-driven simulation of network-wide traffic: A multi-agent imitation learning approach using urban vehicle trajectory data. IEEE Transactions on Intelligent Transportation Systems*, 25(7), 6645-6657.

[17] Sun, J., Kim, J. (2024), *Modelling two-dimensional driving behaviours at unsignalised intersection using multi-agent imitation learning. Transportation Research Part C: Emerging Technologies*, 165, 104702.

[18] Tang, F., Wang, H., Zhang, L., Xu, N., Ahmad, A.M. (2023), *Adaptive optimized consensus control for a class of nonlinear multi-agent systems with asymmetric input saturation constraints and hybrid faults. Communications in Nonlinear Science and Numerical Simulation*, 126, 107446.

[19] Wan, Y., Tang, J., Zhao, Z. (2023), *Imitation Learning of Complex Behaviors for Multiple Drones with Limited Vision. Drones*, 7(12), 704.

[20] Yamane, K., Saigusa, Y., Sakaino, S., Tsuji, T. (2023), *Soft and rigid object grasping with cross-structure hand using bilateral control-based imitation learning. IEEE Robotics and Automation Letters*, 9(2), 1198-1205.

[21] Zare, M., Kebria, P. M., Khosravi, A., Nahavandi, S. (2024), *A survey of imitation learning: Algorithms, recent developments, and challenges. IEEE Transactions on Cybernetics.*