Deep Transfer Learning: A Smarter Approach to Wireless Communication Networks

Siling Wang, *Member, IEEE,* Syed Ali Raza Zaidi, *Senior Member, IEEE,* Maryam Hafeez, *Member, IEEE,* Li Zhang, *Senior Member, IEEE*

***Abstract*—Next-generation cellular networks are evolving into more complex and virtualized systems, utilizing machine learning (ML) for enhanced optimization while leveraging higher frequency bands and denser deployments to meet diverse service demands. Although this evolution brings numerous benefits, it also introduces significant challenges, particularly in radio resource management (RRM). In such environments, effective RRM becomes increasingly difficult due to several factors: more intricate interference patterns, the need for rapid decision-making across a growing number of base stations (BSs) and the highly dynamic nature of user mobility. In addition, the requirements of different types of services further complicate resource allocation, necessitating more advanced and adaptive RRM strategies to achieve optimal performance and maintain high quality of service (QoS). To address these challenges, we propose a ML algorithm that predicts the optimal future serving cell using sequential user equipment (UE) measurements. Conventional ML models require retraining for each environmental change, leading to high complexity and energy consumption. Thus, we also introduce the transfer learning (TL) approach to accelerate model adaptation to dynamic networks and evolving channel conditions, significantly reducing retraining time and improving efficiency. Furthermore, it optimizes key network objectives, such as load balancing and energy efficiency through TL techniques. Our framework complies fully with the O-RAN specifications and is designed to be deployable in a Near-Real-Time RAN Intelligent Controller (RIC).**

I. Introduction

Next-generation networks (NGNs) are designed to support a wide variety of cell types, user devices, radio access technologies and communication paradigms. This multi-layered heterogeneity is intended to support numerous use cases and deployment contexts simultaneously. Consequently, mobile network operators (MNOs) should manage and configure network functionalities which operate across different timescales and serve diverse objectives [1].

RRM within Radio Access Networks (RANs) is a complex large-scale control problem that includes a variety of network functions working at multiple timescales, from sub-millisecond to several seconds. The current architecture that handles RRM in modern RANs has evolved incrementally, with new RRM features being continuously integrated to keep pace with system advancements and requirements. RRM can support different functionalities such as admission control, packet scheduling, and link adaptation. In addition, it provides functions related to power allocation, load balancing, beamforming and handover management, etc. The complexity of RRM will increase in NGNs as optimization domains expand and network demands grow stricter [2]. To address these complexities, ML techniques have been widely used for closed loop control, optimization, and automation, further enhancing the network's efficiency and adaptability. Recently, ML with TL has gained more focus because of its capability to adapt effectively to the dynamic nature of RANs such as fast-varying channel conditions or changed network deployment [3]. The authors in [4] designed and evaluated intelligent handover prediction models for 5G networks to ensure zero downtime during user transitions. In [5], the authors proposed an ML-based algorithm for managing and predicting handovers in mobile wireless networks, while also detecting abnormal handovers to neighboring cells. Furthermore, work in [6] introduces a deep transfer reinforcement learning framework based on multi-agent deep Q-network (TL-MADQN) to address beamforming and resource allocation challenges in downlink multicell multi-input single-output OFDMA (MISO-OFDMA) systems. By leveraging knowledge distillation, the TL-MADQN framework utilizes the neural network parameters of pre-trained agents and the experience gathered in new environments, enabling agents to quickly adapt and efficiently train in the new scenario. Experimental results demonstrate that this approach significantly outperforms conventional methods in terms of convergence speed and data rate, demonstrating its effectiveness in dynamic wireless environments.

Recent advances in the open radio access network (O-RAN) architecture, as outlined in [7], facilitate the integration of third-party ML applications within the RAN to improve network automation and management. In the O-RAN architecture, a BS is divided into a Central Unit (CU), a Distributed Unit (DU), and a Radio Unit (RU), which provides flexibility in terms of deployment and operational roles. Additionally, the CU is further split into the CU User Plane (CU-UP) and the CU Control Plane (CU-CP), each responsible for managing user data traffic and control functions, respectively. RICs introduce programmable elements which can support closed-loop control and orchestration of the RAN by applying ML algorithms. Several studies [8][9][10] have investigated the RRM problem within the O-RAN framework. In [8], the authors propose a deep reinforcement learning (DRL) algorithm aimed at optimizing intelligent connection management, with a focus on user–cell association and network load balancing in O-RAN networks. The work in [9] leverages an O-RAN platform to design and evaluate a RRM solution based on reinforcement learning (RL), deployed as an xApp within the O-RAN ecosystem. Their framework periodically collects network status reports from the O-RAN DU and adjusts resource allocation and modulation/coding schemes for each traffic flow, thereby dynamically fulfilling the Key

Performance Indicator (KPI) requirements. Meanwhile, [10] addresses the slow convergence of DRL-based RRM algorithms in 5G networks by minimizing the exploration phase and expediting policy learning, especially for RAN slicing. However, when traffic patterns differ substantially between source and target domains, the transferred knowledge may become unsuitable for the target environment, leading to suboptimal decision-making.

Therefore, there is still a gap in the literature regarding RRM by using ML with TL scheme based on O-RAN architecture, especially as network environments are becoming increasingly complex and dynamic. Thus, our main contributions in this work are recapped as follows:

- We analyze and summarize the practical challenges faced by deep TL-based RRM and the stochastic nature of NGNs RAN environments. We believe that overcoming these issues is crucial for enabling the adoption of deep TL in real-world applications.
- We provide a clear classification of TL techniques, allowing readers to have a deeper understanding of the principles and definitions of TL. Additionally, it provides valuable insights to help in selecting the most suitable TL method for their specific needs.
- We introduce a case study about user mobility management which can be deployed in O-RAN architecture for real time decision. Our objective is to demonstrate the significance of utilizing safe and accelerated ML with TL approaches in NGNs. Our proposed scheme enables network operators to dynamically update decision-making rules for network optimization and adapt to new network topologies to enhance capacity. This approach improves network coverage while minimizing retraining time and reducing computational costs.

II. A Brief Overview of Transfer Learning

In the field of ML, one of the greatest challenges is training models to perform effectively when faced with entirely new data or environments. Traditional ML typically involves training models from scratch using a large amount of data, which can be both time-consuming and computationally intensive. The model learns patterns, rules, and representations from the dataset and applies them to solve a specific problem. However, when faced with a new, yet somewhat related problem, traditional models often require retraining on entirely new data, resulting in resource inefficiencies.

TL offers a powerful solution to this limitation. It represents an evolution of ML paradigm, where instead of starting from scratch, a model trained on one task is repurposed to solve a different, but related task. In essence, TL is a technique which transfers the learned knowledge from one task to a related task to improve the learning process of the target task [11]. In other words, TL leverages existing learned knowledge to accelerate learning in new target domains. The difference between traditional ML and TL is shown in Fig. 1. In addition, the concept of TL is particularly appealing in situations where data availability is limited or the computational cost of retraining models is high. Instead of requiring massive labeled datasets

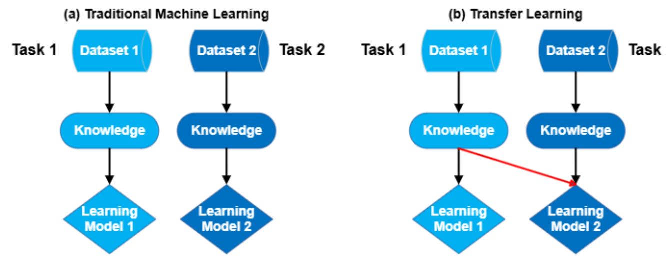


Fig. 1: Differences in the learning process of (a) conventional machine learning, (b) transfer learning.

for every new task, TL allows a model to adapt existing knowledge to perform well in new situations with fewer data. This efficiency makes TL ideal for many practical applications, including computer vision, natural language processing, time series prediction and, as we discuss in this context, RRM issue in wireless communication networks.

Within the scope of next-generation wireless networks, TL is particularly valuable due to the dynamic nature of the environment [12]. For instance, different scenarios such as urban, rural, or dense cities often share similar characteristics, but also have unique distinctions. TL allows a model that has learned to optimize resource allocation in an urban setting to adapt quickly to a rural scenario without requiring exhaustive retraining. This adaptability not only reduces computational costs but also ensures that networks are more responsive to changed channel conditions or changed network deployment. Ultimately, TL expands the utility of ML by allowing pre-trained models to accelerate adaptation and improve efficiency. In the context of next-generation wireless networks, where resource allocation is complex and conditions are constantly changing, TL technique is providing a path toward smarter, more flexible and efficient wireless networks.

A. *Transfer Learning: Terms and Definitions*

The field of TL explores and advances machine learning techniques that utilize knowledge gained from previously addressed source tasks to effectively solve new target tasks. According to [13][14], TL can be categorized by either examining the issues it addresses or the strategies it employs. The categorization based on problems focuses on defining different types of TL depending on the nature of the challenges they solve. More specifically, it considers factors such as the availability of labeled data in the source or target domains, as well as the similarity between the feature spaces of the source and target inputs.

Besides, three core research questions in TL are: What to transfer, How to transfer, and When to transfer.

- What to transfer: this question involves identifying which specific elements of knowledge or information should be shared or transferred across tasks or domains. Not all information is suitable for transfer. Choosing the right information to transfer is essential, as it can enhance the performance of the target domain.
- How to transfer: once the suitable information has been identified, the next step is to determine the learning

algorithms or mechanisms to facilitate the transfer. This addresses the "How to transfer" question, focusing on designing effective methods that allow the transfer of identified knowledge to be successful.
- When to transfer: this question is concerned with determining the right conditions under which knowledge transfer is appropriate. If the source and target domains are significantly different or unrelated, the transfer may not be beneficial and could even lead to a significant failure which is called negative transfer.

In terms of labeled data availability, transfer learning can generally be divided into three distinct categories:

- **Inductive Transfer Learning**: Refers to scenarios where labeled data in the target domain is available[13][14]. Depending on the additional availability of source domain labels, inductive TL can be further divided into: (i) self-taught learning and (ii) multi-task learning[11].
- **Transductive Transfer Learning**: Involves cases where only the labeled data from the source domain is accessible.
- **Unsupervised Transfer Learning**: Involves cases that labeled data are not available at both the target and the source domains. Moreover, the target task is not the same as the source task. Compared to other TL approaches, unsupervised TL has little research work[13][14].

According to the similarity of feature spaces, TL can be divided into two different categories:

- **Homogeneous transfer learning**: Refers to scenarios where source and target feature spaces are identical.
- **Heterogeneous transfer learning**: Refers to scenarios where source and target feature spaces are distinct.

III. Deep Transfer Learning for Radio Resource Management: Challenges and Opportunities

The effective implementation of ML-based TL algorithm for RRM in modern O-RANs can tackle numerous challenges arising from the complex and highly dynamic characteristics of these environments.

A. Challenges

The first challenge is the Complexity of Model Adaptation. For example, when applying deep learning (DL) combined with TL in RRM, adapting pre-trained models to different network scenarios remains challenging due to the highly variable and dynamic nature of wireless environments. Meanwhile, determining which aspects of the pre-trained knowledge are also vital, and which need adjustment, adds further complexity to the model adaptation process.

Another challenge is the Negative Transfer Risk. TL technique poses the risk of negative transfer, where knowledge transferred from a source domain may not be applicable or beneficial in the target domain. This is particularly problematic in radio resource management, where differences between source and target environments (e.g., varying user densities, interference levels, or hardware configurations) can lead to suboptimal performance, ultimately degrading network quality. Additionally, TL generally works well when there is at least some labeled data in the target domain to fine-tune the pre-trained model. In wireless networks, obtaining labeled data is often challenging because the conditions are constantly changing, and the labeling process requires significant resources. Lack of sufficient labeled data in the target domain can make it difficult for the transfer learning model to be accurately adapted.

Lastly, the scale of modern wireless networks is immense, with numerous cells, users, and access points, all requiring continuous management. The fine-tuning process of the transfer model required for adaptation can become computationally expensive, particularly when dealing with thousands of cells and heterogeneous data.

B. Opportunities

Advancements in computer hardware technology, efficient data storage solutions, and more sophisticated machine learning tools provide an opportunity to improve the design of RRM algorithms. The leap in graphical processing units (GPUs) and multi-core central processing units (CPUs) has made large-scale parallel computing widely accessible at relatively low costs. Specifically, GPUs accelerate the convergence speed of artificial intelligence (AI) models by leveraging their immense parallel processing capabilities, enabling faster training and more efficient model updates. This progress allows for leveraging the massive amounts of data continually gathered in radio networks as the foundation for RAN intelligence, from which RRM algorithms can be progressively updated. Modern networks are collecting data related to user mobility, traffic patterns, and user actions (e.g., how, when, and what users do in the network) at significantly higher rates than before, providing even richer datasets to intelligent training model. Moreover, by integrating AI capabilities directly into the RAN infrastructure, networks can achieve more autonomous and intelligent resource management, leading to improved efficiency and adaptability.

Furthermore, wireless networks are highly heterogeneous, involving multiple types of cells, frequency bands, and user equipment. TL offers the opportunity to share knowledge across different emerging use cases such as ultra-reliable low-latency communications (URLLC), massive IoT, and vehicular communication (V2X) which can help enhance the efficiency of RRM algorithms and reduce redundant learning processes.

IV. Deep Transfer Learning-Based Radio Resource Management in Wireless Networks

In this section, we present a case study which investigate the aforementioned ML-based TL challenges in the NGNs field. We believe that the proposed scheme can assist fellow researchers in further addressing the discussed issues systematically. Our focus is on DL, a subset of ML which uses neural networks with multiple layers to learn complex patterns in data. This type of learning involves training the model using a large set of examples consisting of input-output pairs, also known as labeled samples. During training, the model uses forward propagation to generate predictions and compares them with the actual output to calculate errors. These errors are then used in back propagation, a process that helps refine the model by adjusting its internal parameters using gradient descent, ultimately reducing the overall error. Then, we evaluate the trained model's performance using the test set to assess its ability to make accurate predictions on future new data. After, when the network topology or channel condition is changing, TL will used to reconfigure the pre-trained model in the new scenarios.

A. System Architecture

We consider a heterogeneous dense cell network deployed for urban scenario, as depicted in Fig. 2. The network consists

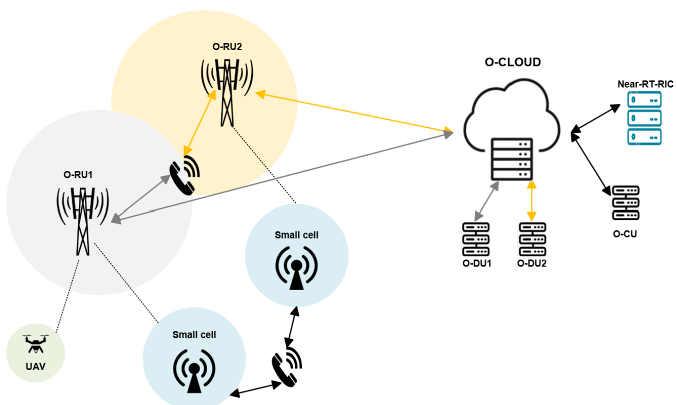


Fig. 2: System model.

of two layers. The first layer utilizes macro BS operating at a carrier frequency of 3.8 GHz to ensure uninterrupted coverage. The second layer employs small BS operating at Millimeter Wave (mm-Wave) frequency of 24 GHz, which offers high-speed data services to UEs thanks to their abundant bandwidth resources. Thus, the mm-Wave technology can support a massive number of connected devices simultaneously, especially in dense urban environment. Additionally, based on the disaggregated RAN architecture, 3GPP has introduced new types of handover [15] considering gNB-CU and gNB-DU implementations. Intra-gNB-DU handover occurs within the same gNB-DU, while an inter-gNB-DU and intra-gNB-CU handover involve transitions between different gNB-DUs under the same gNB-CU. Inter-gNB-CU handovers involve transfers between different gNB-CUs, executed via Xn or N2 interface. In Fig. 2, it illustrates a scenario of an inter-gNB-DU and intra-gNB-CU handover, where cell 1 (O-RU1) is connected to O-DU1, and cell 2 (O-RU2) is connected to O-DU2, where both of the cells are managed by the same O-CU. This figure illustrates the process where the UE transitions from RU1 to RU2.

B. Deep Learning Algorithm

To provide a concrete solution to the challenges of developing RRM algorithms in a radio environment, we propose a general learning framework consisting of two main components: a Long Short-Term Memory (LSTM) learning algorithm and transfer learning. Our approach involves using offline experience to train the model in the Non-RT RIC Layer before deploying it in a live network setting. This offline training utilizes recurrent neural networks (RNNs) that leverage created datasets based on a simulated network environment which takes into account real-world scenarios.

In this work, we specifically address the problem of predicting the optimal target BS for users as it approaches the intersection of edge BSs. This problem is formulated as a multi-class classification task. To predict the most suitable target BS, we utilize sequential channel quality measurements, such as Reference Signal Received Power (RSRP) and Signal-to-Interference-plus-Noise Ratio (SINR), which are collected from the UE over a predefined time window size W. These measurements are reported to the serving BS and used to predict the target BS for the subsequent time window W. Additionally, we take into consideration other parameters, such as available Physical Resource Blocks (PRBs) and the number of users in each BS, to ensure effective load balancing and energy efficiency.

In order to solve the multi-class classification problem and incorporate dynamic decision-making rules, we propose a scheme that employs both DL and TL techniques. DL is particularly effective for identifying and learning complex patterns in sequential data, which is crucial for predicting cell handovers in telecommunications—an area where traditional machine learning models often fall short. DL models are highly capable of learning and handling long-term dependencies, which is essential for forecasting network behavior and predicting cell transitions. These models autonomously discover useful representations for feature selection, significantly improving the accuracy of handover predictions. Moreover, DL models can adapt to emerging patterns in dynamic network environments, thereby enhancing the decision-making process during user mobility.

C. Transfer Learning Algorithm

We employ TL because it enables rapid adaptation to new or evolving tasks and environments by utilizing knowledge from previously trained models. By leveraging this prior knowledge and TL weights from the original model, we can significantly reduce training time and computational costs while maintaining a high level of prediction accuracy.

In our system model, TL is particularly useful when new control parameters are introduced or when new BSs are deployed. Our algorithm is UE-centric and is designed to be deployed as an xApp in a near-RT RIC for mobility management. The near-RT RIC collects and pre-processes data from the E2 nodes (such as O-DU and O-CU) and forwards this data through internal interfaces to the xApp. For near-real-time decision making, we assume that the O-DU nodes and the near-RT RIC are deployed in an edge cloud, while the O-CU nodes are deployed in a regional cloud. Fig. 3 illustrates the deployment of our mobility management xApp within the near-RT RIC, and shows the process for making handover decisions. More information about our implemented algorithm can be found in Algorithm 1.

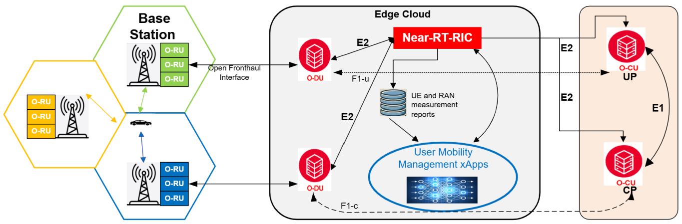


Fig. 3: Framework of our proposed mobility management algorithm.

D. Evaluation Study Cases

For training our model, we created datasets based on a simulated network environment, as described in Fig. 2. We generate a dataset in which user mobility decisions are not only made based on channel quality measurements reported by the UE but also by the availability of radio resources at the target BSs, following Multi-Criteria Decision Making (MCDM) algorithm. we collect a dataset of size 5000 samples in this scenario, from which 4000 are used for training purpose (80%), and 1000 samples for test purpose (20%).

The performance of our proposed scheme was evaluated by comparing it with two different ML algorithms: Support Vector Machine (SVM) and Convolutional Neural Network (CNN). Initially, the performance comparison of our model utilized UE and cell-specific measurements in scenarios without dynamic events.

Fig. 4(a) and 4(b) illustrate the accuracy of algorithms and training losses before a transfer learning event occurs, respectively. Specifically, Fig. 4(a) and 4(b) display the accuracy and training losses when our model is initially trained using UE channel quality features and cell load feature. We assume that a transfer learning event is initiated, i.e., a dynamic UAV BS is added to the network. Initially, the training accuracy decrease to a relatively low value, but eventually, the model converges near to the optimal accuracy and training loss like the model was trained initially with all the features and the retraining process only requires less epochs, saving the retraining time.

The integration of a new UAV BS into the existing network topology introduces dynamic changes that require updating the LSTM-based model to adapt to the altered configuration. Retraining the model from scratch is computationally expensive and time-consuming. Therefore, transfer learning is employed to leverage the knowledge embedded in a pre-trained model and adapt it efficiently to the new scenario. In our scheme, the process begins by loading the pre-trained LSTM model, which has already been trained on the original network topology. In our experiment, we remove its final three layers: the fully connected layer, the softmax layer, and the classification layer. These layers are specific to the original topology, and their removal allows the remaining layers to act as a feature extractor for the new topology. Next, new layers are appended to the truncated model. A new LSTM layer with 50 units is added to provide flexibility for adapting to the unique characteristics of the updated topology, such as dynamic UAV interactions. Additionally, a new fully connected layer is introduced to output probabilities for the updated set of BSs such as UAV. The architecture is completed with a softmax layer for normalizing the output and a classification layer for final predictions. The modified model is then trained on the updated dataset. The training process reuses the weights of the pre-trained layers as a starting point, focusing on refining the knowledge for the new topology instead of relearning general

---

Algorithm 1 Cell Prediction using Transfer Learning

Inputs:

- θ: Initial parameters of the model
- N: Number of neighboring cells
- $X_{W,M}$: Sequence of size W with M features at each time step

Output: $C_{W,N}$: Cell prediction over the next W window

Algorithm:

1: Initialize model parameters θ. Define LSTM network structure:
- Input layer: $X_{W,M}$
- LSTM layer: Extracts features from the sequence, outputs the hidden state of the last time step h_{last}
- Fully connected layer: Maps h_{last} to N neighboring cells
- Softmax layer: Converts the outputs to probability distributions
- Classification layer: Handles supervised learning

2: **Training:**
- For $i = 1$ to K:
  1) Preprocess the input $X_{i:W,M}$: $X_{\text{scaled}} \leftarrow \text{Scaler}(X_{i:W,M})$.
  2) Extract features using LSTM: $h_{\text{last}} \leftarrow \text{LSTM}(X_{\text{scaled}}, h, c)$.
  3) Compute cell predictions: $\hat{C}_{(i+1):W,N} \leftarrow \text{Softmax}(\text{FC}(h_{\text{last}}))$.
  4) Optimize the cross-entropy loss $L(\theta)$ and update parameters: $\theta \leftarrow \theta - \eta \nabla_\theta L$.

3: **Prediction:**
- If no dynamic features or events are detected:
  - Predict the next window:

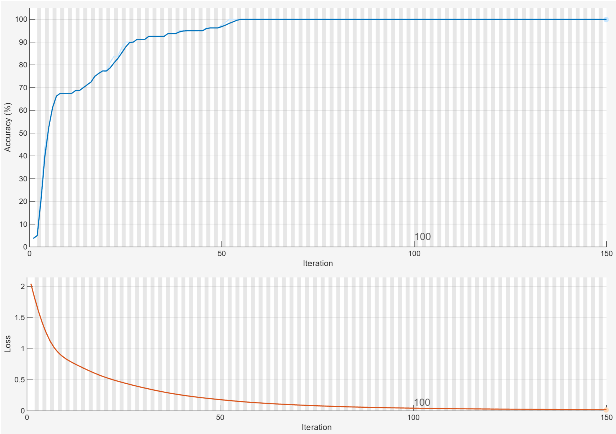
$$\hat{C}_{\text{next},W,N} \leftarrow \text{Softmax}(\text{FC}(\text{LSTM}(X_{\text{previous},W,M})))$$

- If dynamic features or events are detected:
  - Reconfigure the model by adding new randomized parameters:

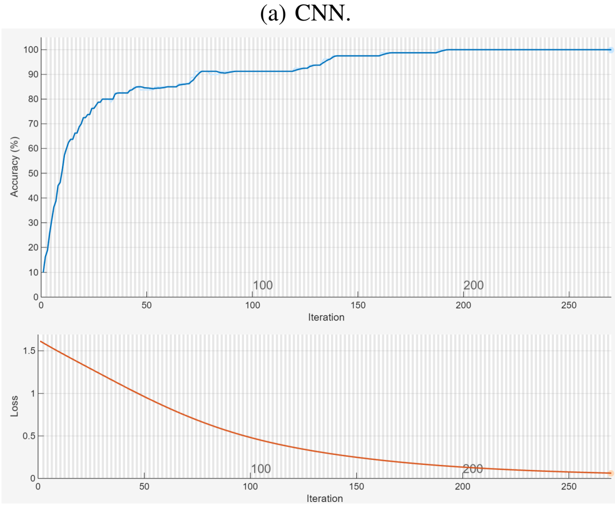
$$\theta' \leftarrow \theta \cup \{\text{randomized new parameters}\}$$

  - Retrain the model.

4: Output the prediction $C_{W,N}$.



(a) CNN.



(b) LSTM.

Fig. 4: Training process of (a) CNN, (b) LSTM.

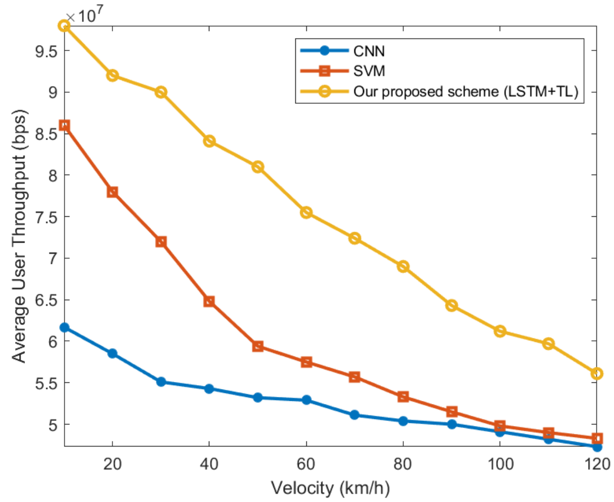


Fig. 5: Average User Throughput.

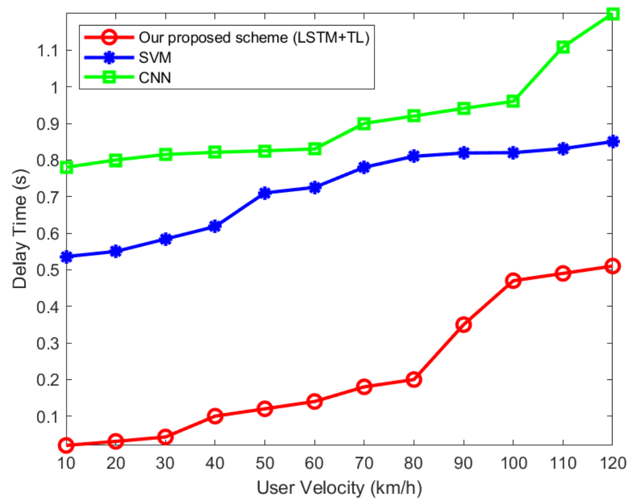


Fig. 6: Delay Time.

sequential features. This significantly reduces training time and computational requirements.

With the increase of user velocity, the throughput of users for all methods is decreased as illustrated in Fig. 5. CNN-based method has the worst throughput performance when compared with other ML algorithms. In addition, the LSTM-based TL scheme clearly outperforms the other classification learning algorithms. Specifically, when the speed of users is 60 km/h, the throughput of the proposed scheme outperforms by about 23.8% and 29.9% respectively in comparison to SVM and CNN algorithms. Moreover, the increased number of HO can reduce the throughput when the user velocity increases. Since frequent handovers can lead to temporary disruptions (handover time) in data transmission. Each handover process involves control signaling and potential latency, which can lower the effective desired data throughput.

In Fig. 6, the abscissa is the user velocity. It is clear that the proposed scheme has the lowest delay time and also achieves the best user QoS when compared with other ML algorithms such as SVM and CNN.

Additionally, the performance of our proposed algorithm, SVM and CNN are evaluated by using different metrics such as Precision, Recall, accuracy and F1 score. The performance metrics obtained during the test phase are summarized in Table I. The trained model is assessed on the test set of the dataset. Table I shows that our proposed algorithm and CNN can achieve better accuracy and outperform in general when compared with SVM classification methods.

TABLE I: Performance metrics obtained on the test set of dataset

Algorithm	Accuracy	Precision	Recall	F1 score
SVM	82.3%	0.85	0.71	0.72
CNN	94.7%	0.94	0.932	0.935
Our Model	98.4%	0.978	0.96	0.97

V. CONCLUSION

This paper has investigated the user mobility problem in the next generation wireless communication systems. We propose a case study about mobility management scheme based on ML with TL techniques to predict the optimal target BS. Here, we approach the prediction task as a multi-classification problem by applying the ML algorithm on available information in RAN environment. At the same time, the proposed predictive algorithm is deployed as an xApp in the near-RT RIC. This paper establishes an effective classifier model for optimal user mobility management by considering multiple criteria within the wireless network. In addition, when the network deployment or channel condition is changing, our experiment highlights the potential of using TL to guide the reconfiguration of model structure. As demonstrated by the simulation results, the proposed scheme outperforms other ML-based algorithms in terms of average user throughput and delay time. In addition, we also use various performance metrics to evaluate their accuracy. The results show that the LSTM-based TL and CNN models can achieve higher accuracy which are found to be 98.4% and 94.7%. However, during the actual implementation process (user mobility decision making process), our proposed scheme and SVM can perform better in terms of the average user throughput and delay time. Thus, the hybrid approach (LSTM-based TL) can be considered as an example of speeding up the training process when the network environment is altering and the proposed prediction framework holds great promise and deserves further investigation in real-world scenarios.

REFERENCES

[1] Francesco Davide Calabrese, Li Wang, Euhanna Ghadimi, Gunnar Peters, Lajos Hanzo, and Pablo Soldati. Learning radio resource management in rans: Framework, opportunities, and challenges. *IEEE Communications Magazine*, 56(9):138–145, 2018.

[2] Bharat Agarwal, Mohammed Amine Togou, Marco Marco, and Gabriel-Miro Muntean. A comprehensive survey on radio resource management in 5g hetnets: Current solutions, future trends and open issues. *IEEE Communications Surveys Tutorials*, 24(4):2495–2534, 2022.

[3] Shuteng Niu, Yongxin Liu, Jian Wang, and Houbing Song. A decade survey of transfer learning (2010–2020). *IEEE Transactions on Artificial Intelligence*, 1(2):151–166, 2020.

[4] Navdeep Uniyal, Anderson Bravalheri, Xenofon Vasilakos, Reza Nejabati, Dimitra Simeonidou, Walter Featherstone, Shangbin Wu, and Daniel Warren. Intelligent mobile handover prediction for zero downtime edge application mobility. In *2021 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2021.

[5] Li-Ping Tung, Bao-Shuh Paul Lin, et al. Big data and machine learning driven handover management and forecasting. In *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*, pages 214–219. IEEE, 2017.

[6] Xiaoming Wang, Gaoxiang Sun, Yuanxue Xin, Ting Liu, and Youyun Xu. Deep transfer reinforcement learning for beamforming and resource allocation in multi-cell miso-ofdma systems. *IEEE Transactions on Signal and Information Processing over Networks*, 8:815–829, 2022.

[7] Aly S Abdalla, Pratheek S Upadhyaya, Vijay K Shah, and Vuk Marojevic. Toward next generation open radio access networks: What o-ran can and cannot do! *IEEE Network*, 36(6):206–213, 2022.

[8] Oner Orhan, Vasuki Narasimha Swamy, Thomas Tetzlaff, Marcel Nassar, Hosein Nikopour, and Shilpa Talwar. Connection management xapp for o-ran ric: A graph neural network and reinforcement learning approach. In *2021 20th IEEE international conference on machine learning and applications (ICMLA)*, pages 936–941. IEEE, 2021.

[9] Federico Mungari. An rl approach for radio resource management in the o-ran architecture. In *2021 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pages 1–2, 2021.

[10] Ahmad M Nagib, Hatem Abou-Zeid, and Hossam S Hassanein. Transfer learning-based accelerated deep reinforcement learning for 5g ran slicing. In *2021 IEEE 46th Conference on Local Computer Networks (LCN)*, pages 249–256. IEEE, 2021.

[11] Cong T Nguyen, Nguyen Van Huynh, Nam H Chu, Yuris Mulya Saputra, Dinh Thai Hoang, Diep N Nguyen, Quoc-Viet Pham, Dusit Niyato, Eryk Dutkiewicz, and Won-Joo Hwang. Transfer learning for wireless networks: A comprehensive survey. *Proceedings of the IEEE*, 110(8):1073–1115, 2022.

[12] Meiyu Wang, Yun Lin, Qiao Tian, and Guangzhen Si. Transfer learning promotes 6g wireless communications: Recent advances and future challenges. *IEEE Transactions on Reliability*, 70(2):790–807, 2021.

[13] Zhuangdi Zhu, Kaixiang Lin, Anil K Jain, and Jiayu Zhou. Transfer learning in deep reinforcement learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[14] Asmaul Hosna, Ethel Merry, Jigmey Gyalmo, Zulfikar Alom, Zeyar Aung, and Mohammad Abdul Azim. Transfer learning: a friendly introduction. *Journal of Big Data*, 9(1):102, 2022.

[15] O-ran working group 1. (2021, july). o-ran architecture description 5.00 (oran.wg1.o-ran-architecture-description-v05.00 technical specification). o-ran alliance.