Deep Learning-based Network Slice Recognition

Harun Ur Rashid
Dept. of Information and Communications Engineering
Hankuk University of Foreign Studies (HUFS)
Seoul, Korea
Email: harun@hufs.ac.kr

Seong Ho Jeong
Dept. of Information and Communications Engineering
Hankuk University of Foreign Studies (HUFS)
Seoul, Korea
Email: shjeong@hufs.ac.kr

***Abstract*— In the context of future wireless networks, various emerging applications such as AR/VR/XR, e-Health, live video streaming, and automated vehicles are expected to have the diverse and strict quality of service (QoS) requirements. Meeting these requirements with high reliability, low latency, high security, high throughput, and high speed, will demand mobile service providers to offer programmable solutions for delivering services in multiple ways. Network slicing that can be achieved through the combination of Network-Functions Virtualization (NFV) and Software-Defined Networking (SDN) is one such solution for customized service instances. This paper proposes an artificial intelligence (AI)-based deep learning method using fully connected neural networks to select an appropriate network slice for users from a public dataset. The considered communication slices include enhanced Mobile Broadband (eMBB), large-scale machine-type communication (mMTC), and Ultra-Reliable Low Latency Communications (URLLC). The performance evaluation of the proposed slice recognition algorithm demonstrates its high accuracy for recognizing the best slice for a given service.

Keywords*— *Network slicing, fully connected layer, future wireless networks, deep learning (DL).

I. Introduction

Comparing 4G and 5G networks, 5G networks offer significant improvements in coverage, management, and accounting capabilities [1], which foster the rapid pace of innovation in cellular communication technologies and generate new use cases in the era of the Internet of Things (IoT), autonomous driving, and augmented, virtual and extended reality (AR/VR/XR) services. With the increasing number of end users and connected devices, the volume of generated data has also been growing tremendously, as evidenced by the current 8.1 billion mobile subscriptions and the projected total mobile data traffic of around 288 EB per month by 2027, according to the most recent Ericsson mobility report [2]. Consequently, one of the main goals of 6G systems is to address challenges that were not effectively tackled by 5G, such as demands for higher data rates, higher capacities, lower latency, real-time processing, massive device connectivity, higher reliability, lower cost, improved QoS, and quality of experience (QoE).

To address the challenges posed by the increasing number of diverse applications and their strict requirements, recent efforts have aimed at shifting from traditional "one-size-fits-all" network designs to more flexible and programmable architectures. SDN and NFV have emerged as key technologies that enable the creation of a scalable and flexible network platform capable of managing multiple services with varying performance requirements [3]. In the context of 6G networks, a network slicing approach has been proposed to achieve this goal by leveraging SDN and NFV. Network slicing allows the coexistence of services with different demands in a unified infrastructure. Also, it opens up new possibilities for vertical business segments and services for consumers and enterprises. The adoption of machine learning (ML), AI, feedback-based automation, and advanced analytics in the telecom industry is transforming the networks significantly to accommodate the next generation applications and services. With the vast amount of data available over 5G, the ability to predict the network situation proactively and make a decision with accuracy is vital. To this end, AI will enhance the performance of the future network by enabling data-driven decision-making, which will lead to minimize the impact of the traffic explosion. The objective of this paper is to develop a deep learning model that can detect and eliminate threats via incoming connections, even in the event of network failure, by selecting the most appropriate network slice. Fig.1 represents the idea of achieving the recognition of an appropriate network slice over the future network using the 5G/LTE data.

The remaining sections of the paper are structured in the following way. A literature review on the utilization of deep learning in network slicing allocation is presented in Section II. The system model and its solution are described in Section III. The outcomes of the simulations are demonstrated in Section IV, and finally the paper concludes with Section V.

II. Related Studies

This section includes various studies that explore the use of ML techniques for network slicing allocation. Numerous studies have employed various ML models for network slice allocation. For example, [5] employed a Convolutional Neural Network (CNN) model for network slicing prediction, while [6] utilized both a neural network and a Deep Belief Network model to make joint decisions for network slicing. Another study [7] focused on autonomous vehicles and used a combination of CNN and reinforcement LSTM learning to estimate user needs. In a study [8], network slicing was performed using an LSTM model with user data such as volume and usage time. Finally, a study [9] generated synthetic data on a realistic environment based on 3GPP standards and compared the performances of two different ML methods for network slice assignment based on the data and user requirements.

III. System Model

A. Slice Deployment

The concept of network slicing is a novel and innovative approach to obtain distinct configurable slices from a physical network, enabling the network to meet the diverse requirements of expected services in the 5G and the era of the 6G network. The application of ML algorithms facilitates the analysis of a large volume of data within a short period, efficient learning of operations in diverse environments, and precise prediction of future outcomes. The slice-allocation mechanism assigns the best slice for a given service type based on its requirements, such as latency, reliability, availability, and throughput.

Initially we investigated the number of devices that communicate over the 5G network, including smartphones, IoT devices, etc. Various attributes of different devices or users, such as user device type, duration, packet loss ratio, packet delay, delay rate, were gathered. Subsequently, the collected attributes were normalized using Minmax scaler to a scale of 0 to 1 to reduce redundancy in the data.

B. Feature Description

The following are the various features that are utilized for network slicing in 5G networks.

- **User equipment category:** It is a collection of characteristics that identify a type of device and various parts, characters of a type of device. Examples: Healthcare, Smartphone, and Smart Transportation.
- **Duration:** It explains how long something endures from start to finish. Examples: 10 and 80 s.
- **Packet loss ratio:** It refers to the percentage of data packets transmitted over a network that do not reach their intended destination due to various factors such as network congestion, errors, or data corruption. Examples: 0.01 and 0.000001.
- **Packet delay:** It refers to the time taken for a data packet to travel from its source to its destination in a network. Examples: 5 and 21 s.
- **Connection type:** It implies the type of network connection, either LTE or 5G.

C. Objective Model

The proposed model represents the development and training of a fully connected deep learning model using the Keras API, which is based on the TensorFlow backend. The model is designed as a sequential neural network consisting of four layers. Firstly, a Flatten layer is employed to take input with a shape and to flatten it to a 1D array. Subsequently, two Dense layers are utilized with 32 and 16 nodes, respectively, and ReLU activation function. Finally, a Dense layer with 3 nodes, which is equivalent to the number of output classes, is used. The model is compiled by configuring the Adam optimizer, Sparse Categorical Cross Entropy loss function, and accuracy metric. It is then trained on the training data, for 50 epochs with a batch size of 16. A validation split of 0.2 is also set, which denotes that 20% of the training data will be utilized for validation purposes.

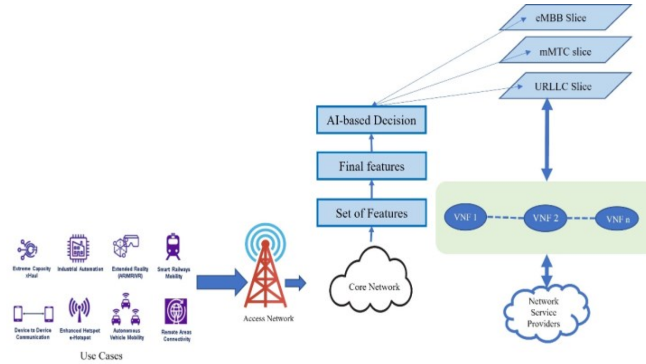

Fig. 1. Architecture of the proposed network slicing model using DL

D. Performance Analysis Metrics

The primary aim of the proposed deep learning techniques for network slicing is to achieve optimal classification performance. To achieve this, a fully connected deep learning model is utilized. The formula used to calculate the model accuracy is as follows:

$$Accuracy = (TP+TN)/(TP+TN+FP+FN) \tag{1}$$

where TP, TN, FP, FN define true positives, true negatives, false positives and false negatives, respectively.

IV. Results and Discussion

The algorithms we proposed were evaluated through simulations for testing their ability to recognize network slices at various split ratios of the dataset. To conduct the simulations, we utilized a computing system with an Intel® Core™ i7-8700 CPU @3.20GHz ×12 Processor, 16GB RAM, and NVIDIA GeForce RTX20270 GPU.

The training data contains 19,958 samples, and validation data consists of 2,218 samples. During the training, the model's loss and accuracy metrics are calculated and printed for each epoch. In Fig. 2(a), the blue line represents the training loss, and the red line represents the validation loss. We can see that the training loss decreases steadily over time as the network learns to fit the training data better. In Fig. 2(b), the blue line represents the training accuracy, and the red line represents the validation accuracy. We can see that both accuracy metrics increase over time, with the training accuracy reaching close to 97% and the validation accuracy reaching close to 97%. Finally, we can say that based on the validation loss and validation accuracy, the model is not overfitting the data. In other words, the model's performance on the validation dataset is not significantly lower than its performance on the training dataset. So, the model has a relatively high accuracy, with a maximum of 97.1%, and a low loss value, indicating that the model is performing well on the dataset.

From Fig. 3, we can see that the model performed very well in class URLLC and mMTC. However, it struggled with class eMBB, misclassifying 65 instances as class eMBB.

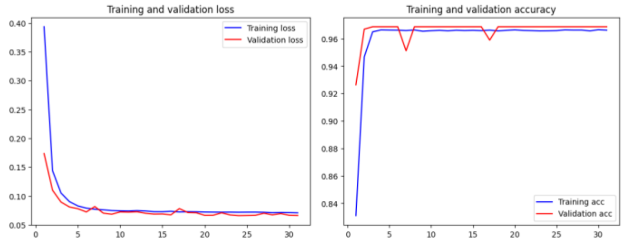

Fig. 2. (a) Proposed DL model loss (b) Proposed DL model accuracy

A confusion matrix is a table used to evaluate the performance of a classification model on a test dataset.

TABLE I. CLASSIFICATION REPORT OF TESTING DATA

Predicted class	Precision	Recall	F1-score
eMBB	1.00	0.92	0.96
mMTC	0.92	1.00	0.96
URLLC	1.00	1.00	1.00
Weighted Avg.	0.97	0.97	0.97

The table presents the evaluation results of the proposed fully connected deep learning model on a test dataset. Table I is represented a classification report, generated by the proposed fully connected deep learning model that has been tested on a dataset of 2,218 instances. The model has predicted the class labels of each instance and compared them to the true class labels in the dataset to calculate several performance metrics, including precision, recall, and F1-score.

The precision, recall, and F1-score values of the model have been evaluated for three classes. The precision value of 1.00 for eMBB and URLLC classes and 0.92 for mMTC class indicates that the model predicted almost all instances of these classes accurately. The recall values of 0.92 for eMBB and 1.00 for mMTC and URLLC classes demonstrate the model's ability to identify almost all instances of these classes correctly. The F1-score values of 0.96 for eMBB and mMTC and 1.00 for URLLC confirm the model's high overall performance. Additionally, the model's weighted average F1-score of 0.97 implies that it has an excellent level of overall performance. Overall, the proposed fully connected deep learning model has demonstrated high precision, recall, and F1-score values, resulting in a highly accurate classification of instances in the network slicing recognition dataset.

V. CONCLUSIONS

In this paper, we proposed a fully connected deep learning technique for efficient network slicing recognition in future networks. The technique accurately classified network slices such as "eMBB, mMTC, and URLLC" for each device based on given attributes. Simulation results demonstrated the effectiveness of the proposed model. The training and validation set accuracy were nearly identical, indicating that the network was not overfitted. Thus, we concluded that the proposed model was efficient for network slicing in 5G and future networks, although further improvement is needed to address more complex problems. Future studies will test the algorithm on other datasets and tune the algorithm parameters for better performance.

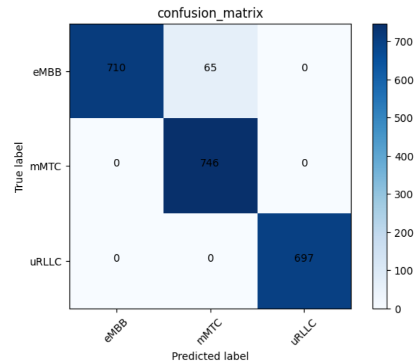

Fig. 3. Result of confusion matrix for testing data

Additionally, edge computing could be implemented to reduce latency further in processing large amounts of data in the cloud and enable real-time analytics.

ACKNOWLEDGMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-RS-2022-00156353) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation). This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00484).

REFERENCES

[1] M. Agiwal, A. Roy, and N. Saxena, “Next Generation 5G Wireless Networks: A Comprehensive Survey,” *IEEE Commun. Surv. Tutor.*, vol. 18, no. 3, pp. 1617–1655, 2016, doi: 10.1109/COMST.2016.2532458.

[2] “Ericsson Mobility Report November 2022,” 2022.

[3] X. Li *et al.*, “Network Slicing for 5G: Challenges and Opportunities,” *IEEE Internet Comput.*, vol. 21, no. 5, pp. 20–27, 2017, doi: 10.1109/MIC.2017.3481355.

[4] V. K. Choyi, A. Abdel-Hamid, Y. Shah, S. Ferdi, and A. Brusilovsky, “Network slice selection, assignment and routing within 5G Networks,” in *2016 IEEE Conference on Standards for Communications and Networking (CSCN)*, Oct. 2016, pp. 1–7. doi: 10.1109/CSCN.2016.7784887.

[5] A. Thantharate, R. Paropkari, V. Walunj, and C. Beard, “DeepSlice: A Deep Learning Approach towards an Efficient and Reliable Network Slicing in 5G Networks,” in *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, Oct. 2019, pp. 0762–0767. doi: 10.1109/UEMCON47517.2019.8993066.

[6] M. H. Abidi *et al.*, “Optimal 5G network slicing using machine learning and deep learning concepts,” *Comput. Stand. Interfaces*, vol. 76, p. 103518, Jun. 2021, doi: 10.1016/j.csi.2021.103518.

[7] J. Mei, X. Wang, and K. Zheng, “Intelligent Network Slicing for V2X Services Toward 5G,” *IEEE Netw.*, vol. 33, no. 6, pp. 196–204, Nov. 2019, doi: 10.1109/MNET.001.1800528.

[8] M. Toscano, F. Grunwald, M. Richart, J. Baliosian, E. Grampín, and A. Castro, “Machine Learning Aided Network Slicing,” in *2019 21st International Conference on Transparent Optical Networks (ICTON)*, Jul. 2019, pp. 1–4. doi: 10.1109/ICTON.2019.8840141.

[9] A. Endes and B. Yuksekkaya, “5G Network Slicing using Machine Learning Techniques,” in *2022 30th Signal Processing and Communications Applications Conference (SIU)*, May 2022, pp. 1–4. doi: 10.1109/SIU55565.2022.9864770.