

DeepCare: Deep Learning-Based Smart Healthcare Framework using 5G Assisted Network Slicing

Deborsi Basu, *Graduate Student Member, IEEE*
G. S. Sanyal School of Telecommunications,
Indian Institute of Technology Kharagpur, India
email: d.basu@ieee.org

Vikram Krishnakumar, *Student Member, IEEE*
Dept. of Electronics and Communication Engineering,
SASTRA Deemed Univ., Thanjavur, Tamil Nadu, India
email: vikramkrishna@ieee.org

Uttam Ghosh, *Senior Member, IEEE*
Dept. of Data Science and Computer Science,
School of Applied Computational Sciences, MMC, USA
email: ghosh.uttam@ieee.org

Raja Datta, *Senior Member, IEEE*
Dept. of Electronics and Electrical Comm. Engg.
Indian Institute of Technology Kharagpur, India
email: rajadatta@ece.iitkgp.ac.in

Abstract—5G and beyond communication networks require satisfying very low latency standards, high reliability, high-speed user connectivity, more security, improved capacity and better service demands. Augmenting such a wide range of KPIs (Key Performance Indicators) needs a smart, intelligent and programmable solution for TSPs (Telecommunication Service Providers). Resource availability and quality sustainability are challenging parameters in a heterogeneous 5G environment. Programmable Dynamic Network Slicing (PDNS) is a key technology enabling parameter that can allow multiple tenants to bring their versatile applications simultaneously over shared physical infrastructure. Latest emerging technologies like virtualized Software-Defined Networks (vSDN) and Artificial Intelligence (AI) play a pivotal supporting role in solving the above-mentioned constraints. Using the PDNS framework, we have proposed a novel slice backup algorithm leveraging Deep Learning (DL) neural network to orchestrate network latency and load efficiently. Our model has been trained using the available KPIs and incoming traffic is analyzed. The proposed solution performs stable load balancing between shared slices even if certain extreme conditions (slice unavailability) through intelligent resource allocation. The framework withstands service outage and always select the most suitable slice as a backup. Our results show latency-aware resource distribution for better network stability.

Index Terms—vSDN, NS, DL, PDNS, 5G, URLLC, load balancing

I. INTRODUCTION

Digitization of technologies and modernization of networking systems drive users to use multiple devices simultaneously for the optimum experience. Next-generation of wireless communication technology (e.g. 5G and beyond) also supports all such advanced developments [1]. Future communication networks are becoming more complex and complicated due to all such exponential device aggregations. Information exchange between those connected devices needs to be maintained thoroughly for stable service flows. Gradually, the devices are becoming intelligent and fewer human interventions are required to perform critical tasks as well as general applications. Human life has become more reliant on mobile communication. Over the last two decades, the number of mobile devices has increased at an exponential rate, with additional services and apps acting as a catalyst. This

shift has necessitated not just increased network capacity and speed, but also the close integration of many technologies. However, for diverse wireless networks, smooth operations and administration have always been a problem, but many service providers have found solutions to fulfil client requests. The 5G network, which is an enhancement of the existing 4G-LTE network, is transforming the cellular sector by opening up new business prospects, allowing for new services, and introducing innovation [2]. 5G networks are viewed as multi-service networks with a diversified range of activities embedded in various performances and services, necessitating a larger device ecosystem. They will offer a richer mobile experience in congested network environments, whether it's mobilising media and entertainment, immersive experiences, high-speed mobility, augmented reality, or linked cars. In order to make appropriate judgments in 5G networks, our research incorporates Deep Learning algorithms to comprehend resource requirements and allocations.

The advanced emerging technologies have ignited several revolutions in the telecom sector, providing end-users with a fresh experience by enabling new business models. Programmable technologies like Network Function Virtualization (NFV) and Software-Defined Networking (SDN) change the way the traditional network works [3]. Next-generation technologies like smart vehicular networks, AR-VR, smart healthcare, industrial automation, enterprise business models, smart cities, and many more important services are being encapsulated by 5G networks. According to 3GPP (Third Generation Partnership Project), NS plays a pivotal role in incorporating all such technologies in 5G. NS allows multiple TNOs to run several applications on the common shared physical infrastructure for better quality-of-service (QoS) while supporting new business services, use cases, and external services.

With the use of machine learning, artificial intelligence, feedback-based automation, and sophisticated analytics to handle next-generation applications and services, the telecom sector is undergoing a huge digital revolution. The methods employed by Deep Learning and Machine Learning are al-

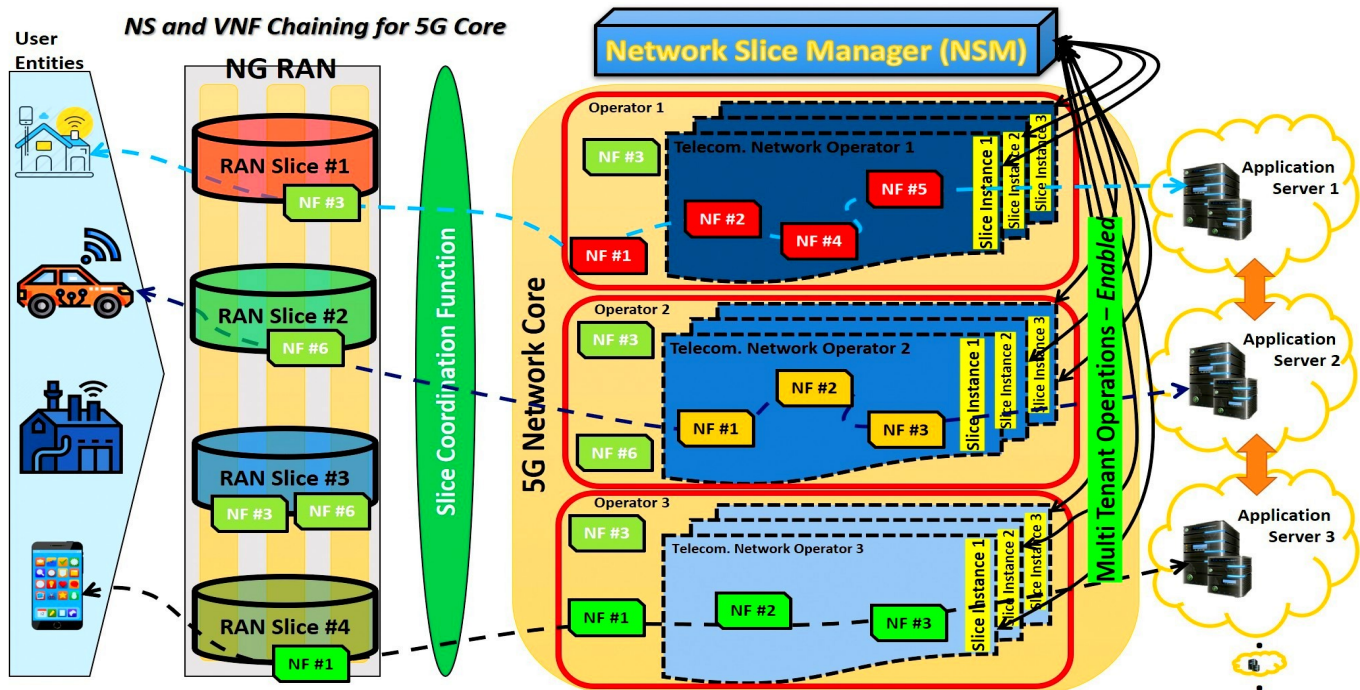


Fig. 1: System model architecture with programmable dynamic network slicing

ready being used in different businesses and technological sectors, therefore AI principles are not new. With the massive amount of data generated by 5G, the ability to forecast data proactively, quickly, and accurately has never been more vital. By enhancing network performance and boosting Quality of Experience (QoE), network functions are primarily driven by AI to offer extended ULL (Ultra Low Latency), reliability, and throughput [4], [5].

Main motivation behind the work: Network disruptions are pretty common nowadays. But sudden service outage creates severe challenges for critical applications like the smart healthcare system. Remote surgeries, teleconsultation, critical healthcare monitoring, and general healthcare observation need an uninterrupted flow of service requests through the network. Any negligence can cause a massive disaster that may further result in loss of life. Though slicing techniques are used in healthcare for a while now, the non-existence of algorithms for consistent slice survivability during certain service windows is still a challenge. The following constraints motivate us to develop a unique slice backup framework using deep learning techniques. In this work, we have considered our area of focus as healthcare, but a similar approach can easily be adapted for other related fields.

Main Contributions: We have drafted a brief outline of our overall contributions mentioning the existing research gaps and scopes for further works. The following points further explain our main contributions:

- We have introduced a *PDNS model with DL framework*. We have primarily targeted building a reliable smart healthcare network architecture to prevent service unavailability due to resource outages.

- Next, we proposed a slice backup algorithm using the DL technique. The aim is to support the healthcare slice in case it goes off due to any unknown disturbances.
- We have analysed and allocated the resources based on load and latency criteria. The service mapping from guest to host slice is done keeping the service flow intact.
- Finally, the load transformation is done using a case study. It is shown that adjacent slices can back the process up in real-time during any critical situations.

Paper Organization: The paper is summarized as follows: Section II describes programmable dynamic slicing and DL for 5G briefly, some related works are shown in Section III, we have explained our PDNS model in Section IV and in Section V result analysis and discussion are done for our use cases for unavailable slice resources, load balancing and network service outage scenario. Finally, Section VI concludes the work with some open future directions.

II. PROGRAMMABLE DYNAMIC NETWORK SLICING AND DEEP LEARNING FOR 5G

The existing LTE design has a fixed conventional framework that makes it difficult to adapt to a variety of use cases. When it comes to supplying any specialised company requirements or meeting unique business expectations, it frequently lacks customisation. Business expectations for better throughput and faster connectivity cannot be met by today's 4G LTE network due to increased mobile data and consumer demands. As mentioned in the earlier section, in 5G, SDN and NFV allow the dynamic up-gradation of the slice clusters without any direct physical interventions with the hardware counterpart [6].

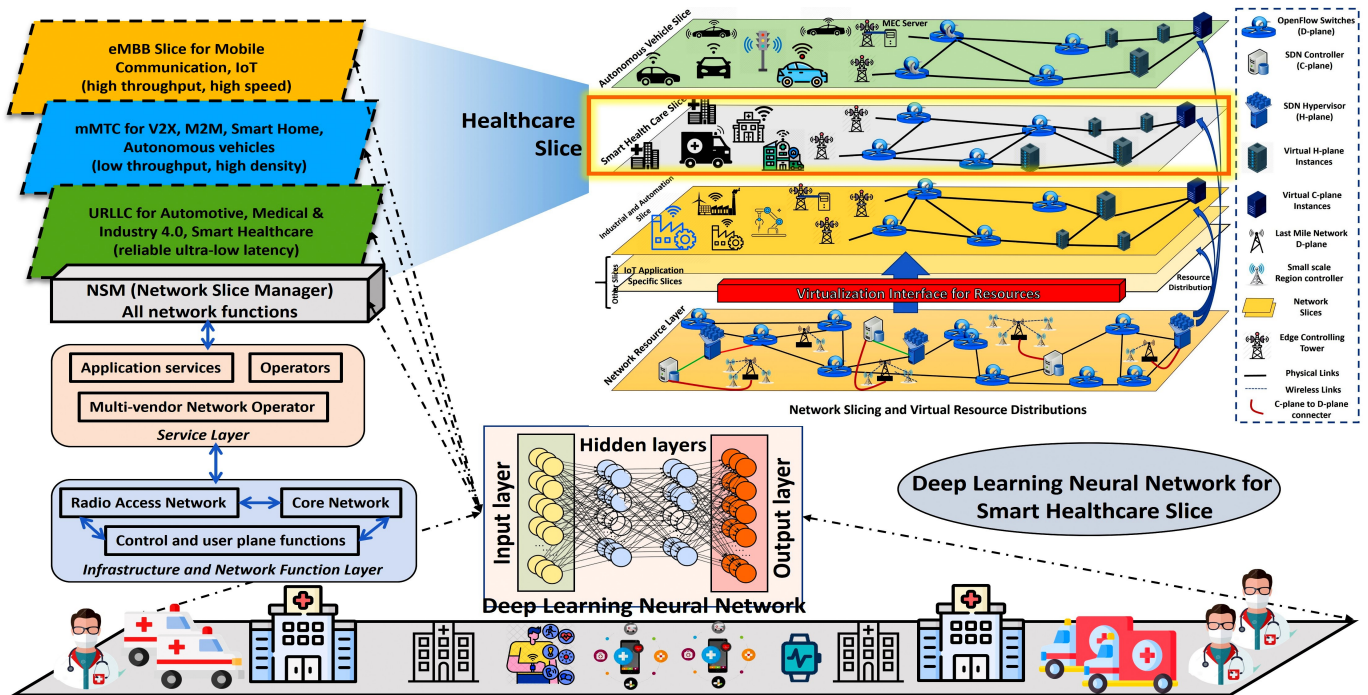


Fig. 2: 5G driven DLNN for Smart Healthcare Framework

The exponential surge in incoming network traffic due to rapid device addition inside the 5G network puts a humongous load on the network operators in managing the system. This network traffic can be classified into different segments based on characteristic evaluations. NS plays a major role here. Individual traffic flows are managed by selected network slices which makes things much more convenient for the operators to imply the business plan (tariffs) accordingly [7], [8]. Thus, business profitability and customer satisfaction are balanced nicely. Moreover, the concept of dynamic network slicing provides additional features by adding, updating, and releasing network functions (virtual network functions) without disrupting the ongoing traffic flows. The concept of NS comprehensively supports the 'One size fits all' approach. Furthermore, TNOs have realized the importance of network learning processes. ML and DL are included to further study the infinite traffic flows and take necessary actions by analyzing unprecedented network behaviour. The correcting measures efficiently modify the slice's operational characteristics making it more user-friendly. Intelligent slicing is introduced through service automation and AI techniques. DNN acts as a catalyst here by enabling dynamic programming features on slices. Resource allocation and reallocation need decision-making skills and DL helps in performing related actions without human interventions. DL techniques consider all possible factors while filtering the appropriate service-specific decision in the 5G network.

The operational standards are followed for individual slices inside a 5G network. A baseline for performance is built via real-time analysis using DL. The analysis results in detecting abnormal service addition in the slices or any resource outage

phenomenon. A single slice instance can be further separated into sub-slices if a sudden overshoot in demand happens inside the network. The network automation system can redistribute the capacity to adjust the sudden traffic turbulence. The roll-back of slices is done once the demand gets satisfied for a particular application. Thus, good resource utilization efficiency is achieved. The dynamic resource flow, allocation, release, reallocation, and deletion are controlled by the NSM (Network Slice Manager). Initially, implementation of new slices was difficult as one needs to manually operate multiple hardware devices which is time-consuming. Programmability solves this issue by providing a flexible and elastic network architecture. The service can be generated on demand. Users can request mobility, isolation, data rate, latency, power constraints, etc. as per their requirements. Accordingly, NSM provides suitable function chains to satisfy the corresponding demand. In case of unavailability of resources on the host slice, associated slices are requested considering the ongoing operations in those slices must remain unaltered [9].

NS is typically a collection of multiple VNFs (virtual network functions) connected via logical links. Tenants manage the VNFs in their respective slices. Each TNO proposes specified SLAs (service level agreements) by governing the factors like service availability, latency, connectivity, data rate, speed, capacity, etc. A slice type is defined by several parameters like power, isolation, mobility, reliability, bandwidth, throughput, etc. In 5G, big data analytics is included to control the vast data explosion across the globe. ML techniques are widely used in cellular communication networks for optimizing mobile tower operations, autonomous decision-making, predictive maintenance, real-time data analytics, anomaly detection, rapid

wireless channel adoption, etc. Applications with differential pricing follow separate resource allocation techniques. ML is used in those cases for real-time prediction with higher accuracy. Big data processing needs high computational power. Hardware compatibility is a must in running such complex applications even though the existence of slicing is there.

III. RELATED WORKS

Inclusion of SDN and NFV is done comprehensively inside IoMT (Internet of Medical Things) [10]. Network softwarization and virtualization are the key two parameters which further drive the notion of network slicing for next-generation communication networks. NS enable the multi-tenant environment considering parallel operations. Multi-tenancy and vendor Independence are explored by Konstantinos Samdanis *et al.* There the authors have shown how the behaviour of the Multi-Vendor Network Operating system varies depending on the transmit power and user capacity [11]. Ping *et al.* have proposed a radio spectrum scheduling algorithm for 5G RAN using Deep Learning methods [12]. NFV and SDN based 5G core (as shown in figure (1)) is shown in [13]. Using a priority management SDN method, the authors of [14] present a system for prioritising network traffic for smart cities. Taiwan began working on network slicing early on and covers standardisation, network slice selection, and identifying slice-independent operations before proposing a slicing architecture and the RRC frame [15].

Apart from the works stated above, we are unaware of any other work that addresses the simple but challenging problem of providing slice backups during resource outages. The approach of applying deep learning is also a new approach in this regard. Reliable, fast, accurate and informative decision-making skills are adapted to the work using proper learning models. An FDOP (Fade Duration Outage Problem) is addressed in [16] as compared with traditional handover (SINR based) in cellular networks. Authors in [17] have explored a separate SDN and NFV driven 5G model for flexible service orchestration for 5G NR (5G New Radio) air interfaces. Multiple related studies have been published as white papers and survey reports showing the increasing growth of mobile devices and their corresponding usage in the coming years. A public safety handover module is proposed by R. A. Paropkari *et al.* in [18] where they have explored the matrix exponential distribution for developing an accurate handover decision provisioning considering all parameters of the decision-making process.

In healthcare, the importance of a survivability framework is non-negotiable [10]. Leveraging the multi-vendor virtualization process, authors in [19] & [20] have built a survivability architecture for 5G networks. Cost-aware slice allocation is done in [21] where authors have shown how efficiently network operators can allocate the resources dynamically considering the diverse service requests of the end-users. However, none of them has considered the resource or slice availability which may alter due to server failures or other relevant problems. The slice backup technique along with load

Input Type	Packet Delay Budget	Duration (sec.)	Packet Loss Rate	Predicted Slice
Industry 4.0	10 / 50	180	$10^{-3} / 10^{-6}$	URLLC / mMTC
Public Safety E911	10	300	10^{-6}	URLLC
IoT Devices	10	60	10^{-2}	mMTC
Smart city / home	50 / 300	120	10^{-2}	mMTC
Smartphone	60 / 75 / 100 / 150 / 300	300	$10^{-2} / 10^{-3} / 10^{-6}$	eMBB
Smart Transportation	10	60	10^{-6}	URLLC
AR / VR / Gaming	10 / 50	600	10^{-3}	eMBB
Smart Healthcare	10	180	10^{-6}	URLLC

TABLE I: The feature highlights of the proposed model

and latency-aware resource distribution is a unique approach in this work. The following sections will elaborate on the system model in detail.

IV. SYSTEM MODEL FOR S-HEALTHCARE FRAMEWORK

The industry is producing super-computing facilities which allow TNOs to incorporate neural networks to access massive data inside the network. Leveraging the concepts of such intelligent techniques, prediction with high accuracy becomes possible. The proposed work uses an ML model and later realizes a DLNN to decide slice availability for resource transfer. The proposed framework is used to detect slice failure (healthcare-related services in this case) and further select suitable available slices as the temporary backup. The statistical ML model follows the conventional Random Forest (RF) approach and the CNN (convolutional neural network) classifier is introduced inside the architecture. RF and CNN are very well-known models mostly used in their respective domains. A dataset consisting of 65,000 unique entries as inputs are used for both ML and DLNN models.

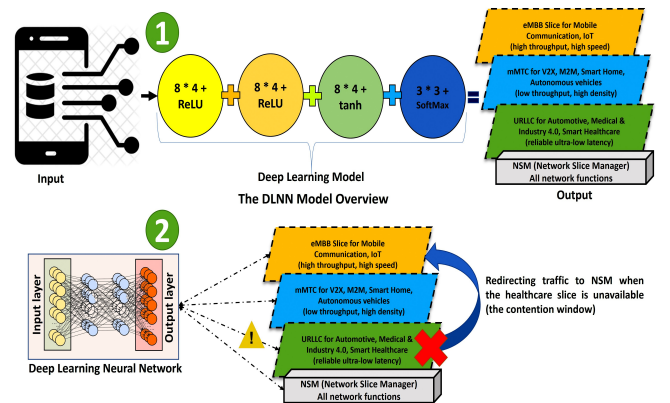


Fig. 3: ML and DLNN model overview and workflow

The dataset consists of the most important KPIs of 5G-guided devices. Detailed type-specific representations of the devices are shown in Table (1). The KPI data can be measured from the control packets traversing from the user entity (UE) to

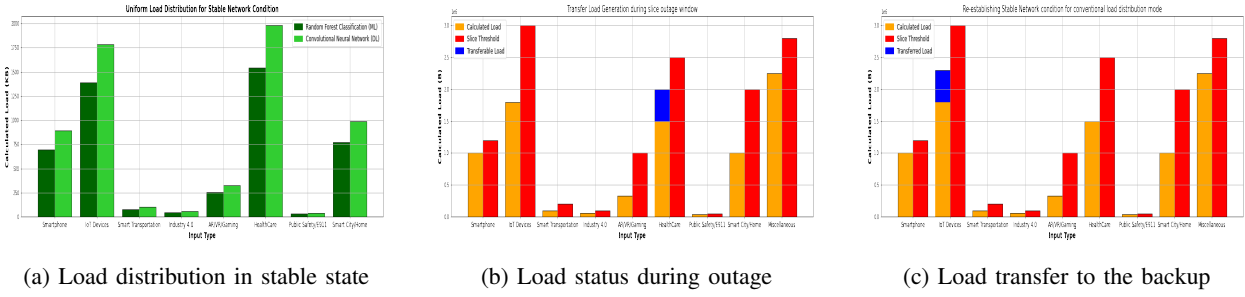


Fig. 4: Load balancing and slice back-up formation during slice resource outage

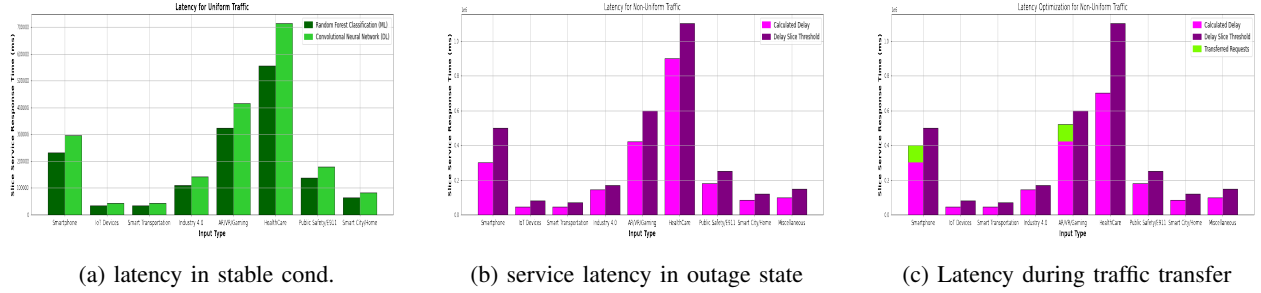


Fig. 5: Service latency distributions over the contention window

the network. The information is stored and accessed readily as the model run internally on the network. The incoming service requests have pre-allocated QCI values assigned based on characteristics. Likewise, 5G services contain a 5G QoS Identifier (5QI) (e.g. packet loss rate, packet delay budget) which we have considered in our model.

The average duration time of a signal is the time spent inside the system. This duration is recorded for each service type for further processing. Three major signal flow categories are used here namely mMTC, URLLC, and eMBB (Enhanced Mobile Broadband, massive Machine Type Communication and Ultra Reliability and Low Latency Communication.) The NSM (Network Slice Manager) does all the allocation, reallocation, and feasible routing of the flow requests. The incoming flows pass through the NSM. The centralization of NSM gives a global view to monitoring the status of each slice. A process scheduling is done if any reallocation request is generated due to slice unavailability. The assigned guest slice process the requests of the host slice for a certain time followed by an inter-slice load transfer.

V. PROPOSED SOLUTION

The proposed model calculates and stores the loads on each slice based on the previous data of incoming connections. It keeps track of the load, used capacity and available capacity of individual network slices. As there is no exact count of incoming traffic, a DLNN is used to predict the resource availability over the slices. The DLNN also supports dynamic slicing as its adaptive nature readjusts in every network outage.

Deep Learning Neural Network

The dataset under consideration contains unstructured data. We train several neurons of our DLNN with the same dataset,

and it predicts the proper network slice depending on any input from the UE information. We use our DLNN to identify the optimal slice for unknown device kinds since it can forecast quite well. If load balancing is necessary for the network slices, as well as in the event of a network slice failure, it assists in redirecting traffic to the NSM. We forecast the network load of each network slice based on the incoming connection in our proposed model and keep track of which output slice is being used the most. The distribution of the incoming traffic is done between URLLC, eMBB, and mMTC slices analysing the output given by the NSM.

VI. RESULT ANALYSIS AND DISCUSSIONS

In this section, the system performance evaluation is done over a use case. We have examined our model to verify how well it can perform load balancing, slice prediction, and service availability. The validation of our proposed approach is done by showing the detection accuracy of resource unavailability and shifting the excess traffic to the best available slice. The load balancing improves the slice performance by restricting any irregular handovers within the slices. The following two scenarios explain the desirable outcomes.

A. Network Slice Future Scenario

As mentioned earlier, we have assumed a complete failure of a URLLC slice as shown in figure (2). The proposed framework directs all the incoming flow to the NMS to avoid any packet drops. Though the sudden outage can not guarantee the integrity of the ongoing processes, the majority of the tasks are kept protected. The centralized NMS has the access to sufficient network resources to host the new traffic flow while searching for a suitable slice to carry forward the tasks. In case of unavailability of any slices, NMS self address the

functions for a good amount of time compromising some of the QoS parameters. Figures (4d), (4e), and (4f) show the time boundaries of the requests following the DLNN techniques in a generalized way.

B. Load Balancing

Figure (4) shows the load balancing scenario during the service outage. We have assumed a contention window of 5 sec. during which the healthcare slice becomes unavailable. Now, figure (4a) shows the network condition during the stable flow condition (we have considered random inflow of traffic). The NSM found the IoT Device slice to be the most suitable one to serve as a temporary backup through the DLNN technique. Figures (4b) and (4c) show how the additional traffic gets shifted from the healthcare to the IoT slice. This way the service flow remains intact and the overall load stays within the threshold.

We have simulated the proposed model using the Keras library in python 3.0 which supports deep learning-related functions. Further improvement in complexity can be done by imposing suitable algorithms which is one of our future targets as well.

VII. CONCLUSION AND FUTURE WORKS

The use of the PDNS architecture is growing faster and its advanced versions will be applied to the next-generation communication networks. Though the applications are ever-expanding, we have focused on one of the critical areas in this work. We have developed a dynamic slice backup algorithm using a deep learning technique that can withstand the service outage and keep the flow intact over a designated time period. Suitable load balancing is done between the slices considering slice capacity and other relevant parameters. A network can only adopt this technique if the supportive measures provide a good boundary for other ongoing service demands. Massive devices make a gigantic load over any particular slice, often it is required to borrow the unused instances from the adjacent one. The same concept is applied here as well. The result shows that following this technique can save the sudden break-up of events and provide a reliable and secure experience. We have assumed secured slicing here which may not be practical in the real world. So, in the future, one of our primary goals would be securing the resource exchange between slices using federated learning techniques.

VIII. ACKNOWLEDGEMENT

This work has been funded by the IHUB-DATA Mobility Group, IIIT-Hyderabad, India and supported by the National Science Foundation, under award number 2219741. We extend our sincere gratitude to the collaborators.

REFERENCES

- [1] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, wide-area, and non-terrestrial iot: A survey on 5g advances and the road towards 6g," *IEEE Communications Surveys & Tutorials*, 2022.
- [2] G. F. Huseien and K. W. Shah, "A review on 5g technology for smart energy management and smart buildings in singapore," *Energy and AI*, vol. 7, p. 100116, 2022.
- [3] D. Basu, R. Datta, and U. Ghosh, "Softwarized network function virtualization for 5g: Challenges and opportunities," *Internet of Things and Secure Smart Environments: Successes and Pitfalls (2020)*, vol. 147, 2020.
- [4] M. Attaran, "The impact of 5g on the evolution of intelligent automation and industry digitization," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–17, 2021.
- [5] D. Basu, A. Jain, U. Ghosh, and R. Datta, "A reverse path-flow mechanism for latency aware controller placement in vsdn enabled 5g network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 6885–6893, 2020.
- [6] D. Basu, V. Krishnakumar, S. Parui, W. Mansoor, and U. Ghosh, "Optimum e-health care resource provisioning using network slicing approach for future 5g networks," in *2021 4th International Conference on Signal Processing and Information Security (ICSPIS)*. IEEE, 2021, pp. 88–91.
- [7] S. Wijethilaka and M. Liyanage, "Survey on network slicing for internet of things realization in 5g networks," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 957–994, 2021.
- [8] D. Basu, A. Sankara Rao, U. Ghosh, and R. Datta, "Realization of a techno-economic controller deployment architecture for vsdn enabled 5g networks," in *Adjunct Proceedings of the 2021 International Conference on Distributed Computing and Networking*, 2021, pp. 49–55.
- [9] M. H. Abidi, H. Alkhalefeh, K. Moiduddin, M. Alazab, M. K. Mohammed, W. Ameen, and T. R. Gadekallu, "Optimal 5g network slicing using machine learning and deep learning concepts," *Computer Standards & Interfaces*, vol. 76, p. 103518, 2021.
- [10] S. Khan and A. Akhunzada, "A hybrid dl-driven intelligent sdn-enabled malware detection framework for internet of medical things (iomt)," *Computer Communications*, vol. 170, pp. 209–216, 2021.
- [11] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, "From network sharing to multi-tenancy: The 5g network slice broker," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 32–39, 2016.
- [12] P. Du and A. Nakao, "Deep learning-based application specific ran slicing for mobile networks," in *2018 IEEE 7th international conference on cloud networking (CloudNet)*. IEEE, 2018, pp. 1–3.
- [13] L. Ma, X. Wen, L. Wang, Z. Lu, and R. Knopp, "An sdn/nfv based framework for management and deployment of service based 5g core network," *China Communications*, vol. 15, no. 10, pp. 86–98, 2018.
- [14] S. Khan, A. Hussain, S. Nazir, F. Khan, A. Oad, and M. D. Alshehri, "Efficient and reliable hybrid deep learning-enabled model for congestion control in 5g/6g networks," *Computer Communications*, vol. 182, pp. 31–40, 2022.
- [15] T. Yoo, "Network slicing architecture for 5g network," in *2016 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2016, pp. 1010–1014.
- [16] R. A. Paropkari, A. A. Gebremichael, and C. Beard, "Fractional packet duplication and fade duration outage probability analysis for handover enhancement in 5g cellular networks," in *2019 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2019, pp. 298–302.
- [17] F. Kurtz, C. Bektas, N. Dorsch, and C. Wietfeld, "Network slicing for critical communications in shared 5g infrastructures-an empirical evaluation," in *2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft)*. IEEE, 2018, pp. 393–399.
- [18] R. A. Paropkari, C. Beard, and A. Van De Liefvoort, "Handover performance prioritization for public safety and emergency networks," in *2017 IEEE 38th Sarnoff Symposium*. IEEE, 2017, pp. 1–6.
- [19] D. Basu, V. Krishnakumar, U. Ghosh, and R. Datta, "Softhealth: Softwarized 5g-driven network slicing for real-time e-healthcare applications using ml," in *2022 14th International Conference on Communication Systems & Networks (COMSNETS)*. IEEE, 2022, pp. 222–226.
- [20] R. Abhishek, D. Tipper, and D. Medhi, "Network virtualization and survivability of 5g networks: Framework, optimization model, and performance," in *2018 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2018, pp. 1–6.
- [21] Q. Wang, J. Alcaraz-Calero, R. Ricart-Sanchez, M. B. Weiss, A. Gavras, N. Nikaein, X. Vasilakos, B. Giacomo, G. Pietro, M. Roddy *et al.*, "Enable advanced qos-aware network slicing in 5g networks for slice-based media use cases," *IEEE transactions on broadcasting*, vol. 65, no. 2, pp. 444–453, 2019.