Deep Learning-based Framework for Multi-Fault Diagnosis in Self-Healing Cellular Networks

Muhammad Sajid Riaz*, Haneya Naeem Qureshi*, Usama Masood*,
Ali Rizwan†, Adnan Abu-Dayya‡, and Ali Imran*
*AI4Networks Research Center, School of Electrical and Computer Engineering, University of Oklahoma, Tulsa, USA
†Qatar Mobility Innovations Centre, Qatar University, Doha, Qatar
‡Department of Electrical Engineering, Qatar University, Doha, Qatar
Email: {riazsajid,haneya,usama.masood,ali.imran}@ou.edu, arizwan@qmic.com, adnan@qu.edu.qa

2022 IEEE Wireless Communications and Networking Conference (WCNC) | 978-1-6654-4266-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/WCNC51071.2022.9771947

***Abstract*—Fault diagnosis is turning out to be an intense challenge due to the increasing complexity of the emerging cellular networks. The root-cause analysis of coverage-related network anomalies is traditionally carried out by human experts. However, due to the vast complexity and the increasing cell density of the emerging cellular networks, it is neither practical nor financially viable. To address this, many studies are proposing artificial intelligence (AI)-based solutions using minimization of drive test (MDT) reports. Nowadays, the focus of existing studies is either on diagnosing faults in a single base station (BS) only or diagnosing a single fault in multiple BS scenarios. Moreover, they do not take into account training data sparsity (varying user equipment (UE) densities). Inspired by the emergence of convolutional neural networks (CNN), in this paper, we propose a framework combining CNN and image inpainting techniques for root-cause analysis of multiple faults in multiple base stations in the network that is robust to the sparse MDT reports, BS locations and types of faults. The results demonstrate that the proposed solution outperforms several other machine learning models on highly sparse UE density training data, which makes it a robust and scalable solution for self-healing in a real cellular network.**

***Index Terms*—Root cause analysis, multi-fault diagnosis, cellular data sparsity, minimization of drive tests, convolutional neural networks, radio environment map inpainting, network automation, self-healing**

I. Introduction

The growing complexity of network architecture, increase in network traffic, and increase in network parameters in the emerging cellular networks are making management tasks increasingly complex [1]. Out of several management challenges in the emerging cellular networks, one challenge is the detection and diagnosis of faults. Faults in cellular networks, which can lead to soft outages (partial service degradation) or hard outages (complete coverage degradation), can occur due to several reasons. One reason is poor network planning, which results in the improper configuration of parameters, such as optimal number, types, and location of base stations, the antenna height, the number of sectors, the sector orientation, tilt, power, frequency reuse pattern, or the number of carriers, among others. Other types of faults can occur due to hardware, software, or functionality failures (e.g., power supply or radio board and network connectivity failures) [2].

Traditionally, outages are detected by either using alarms, performance counters, or by complaints filed by network subscribers [2]. This can take hours and at times days to resolve outage issues. Therefore, for better Quality of Experience (QoE) and Quality of Service (QoS), the network providers must spend a lot of money, time, and energy to do coverage testing via drive tests. This challenge of fault detection and root cause analysis is especially aggravated in ultra-dense network vision of emerging networks, where the same advances in network design that bring advantages such as higher data rates and capacity as compared to the legacy networks, i.e., densification, also leads to growing complexity of the network, making it difficult to manually detect and diagnose faults. The additional burden of growing operational and capital expenditures is making matters worse.

To address these challenges, network automation solutions, i.e., self-healing solutions are needed that can automate the process of fault detection and diagnosis. Only when the outages and their root cause are detected in a timely manner, will the network be able to take actions to compensate for these outages autonomously.

A. Related Work

Root cause analysis of network coverage outages has two components; namely, anomaly detection, and diagnosis. Although outage detection has been studied extensively in the literature, such as in [3], [4], [5], among many others, relatively a small number of studies have focused on outage diagnosis [6], [7]. Minimization of drive test (MDT)-based data can be utilized for both detection [5] and diagnosis [8], [9] of outages. However, utilizing such an approach also has problems, e.g., sparse MDT data [10].

The work presented in [6] used Bayesian classifiers for fault diagnosis based on different network key performance indicators (KPIs). The solution proposed in [7] used unsupervised machine learning, self-organizing maps (SOM) for outage diagnosis. However, SOMs are not robust to varying distributions of data, which makes them nongeneralizable to use with MDT reports, due to the sparse nature of MDT data.

Most relevant to our research are the works presented in [8] and [9], where the authors proposed a fault diagnosis solution using neuromorphic artificial intelligence (AI) and classical machine learning methods, respectively. They used synthetic MDT reports to generate coverage maps. Their

analysis showed that random forest (RF) outperformed convolutional neural networks (CNN) on coverage maps when MDT data is available from the entire coverage area. However, real network data is expected to be both sparse (due to variable user equipments (UEs)) and noisy. Hence, the assumption in [8], [9] is not practical for MDT-based data collection from a real cellular network. Also, both these studies [8], [9] considered faults in a single base station (BS), i.e., they assume a single fault at a time which makes it not generalizable to the scenario where multiple BSs can have faults simultaneously.

B. Contributions and organization

In this paper we address the limitations of the available solutions for fault diagnosis highlighted in Section I-A with the following key contributions:

1) To the best of the authors' knowledge, this is the first solution that can reliably diagnose multiple faults in multiple BSs in the network, caused by both hard outages(network failures leading to no coverage) or soft outages (occurring due to the inefficient configuration of network parameters).
2) We propose a solution that is robust to not only different kinds of faults and BS locations, but also to variable user densities in the network. This makes the proposed solution more feasible to implement in a real cellular network, where user density and distribution never remain static. In addition, in a real network, faults can occur in different BSs, and they can be of different types.
3) To address user sparsity, we investigate different image enrichment methods to reconstruct radio environment maps (REMs) from raw/sparse data acquired from MDT reports and perform a comprehensive performance evaluation of image enhancement techniques and quantify their effect on various machine learning (ML) and deep learning (DL) models from literature as well as on the proposed model.

The rest of the paper is organized as follows: network topology and data acquisition are presented in Section II. The proposed framework based on CNN and data enrichment is described in Section III. Results and insightful performance analysis are provided in Section IV and Section V concludes this study.

II. Network Topology and Data Acquisition

A. Network Topology

The framework we present in this paper is designed for a real network but due to the unavailability of real data, a realistic commercial RF planning software Atoll[11] is used to collect MDT reports. The network topology considering an area from Brussels City, Belgium is shown in Fig. 1. We consider 15 different clutter types based on environmental conditions and terrain profiles. Aster propagation (advanced ray-tracing) is used as the propagation model because of its ability to better capture the idiosyncrasies in the environment as compared to empirical propagation models. We use the same locations and configuration parameters of BS used by a real network provider for its deployment in Belgium. Table I reports these settings. Therefore, the obtained coverage data can be assumed a very close representation of the ground truth of the area used in the simulation. The area of simulation is 13.292 km^2 with 24 macrocell BSs (72 cells) to generate data with multiple fault classes in the multiple BSs simultaneously. The rest of the network settings are reported in Table I.

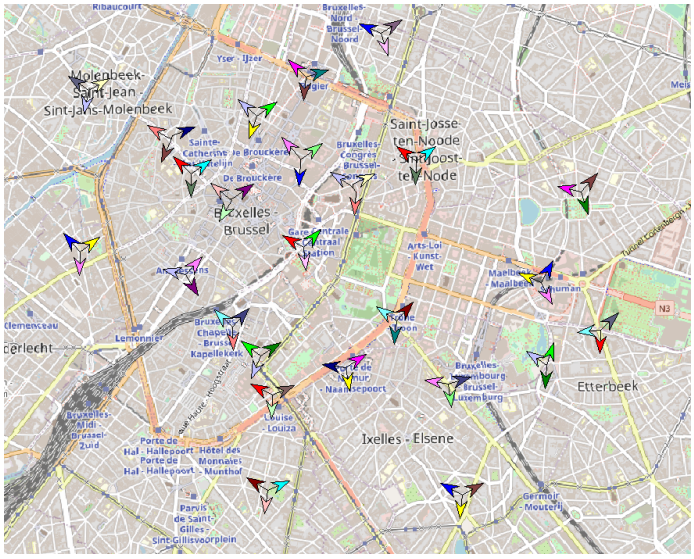


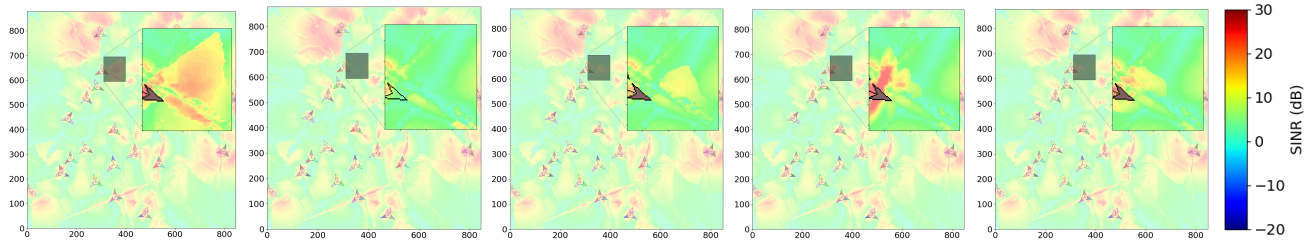
Fig. 1. Network topology for generation of synthetic data

TABLE I
NETWORK SCENARIO SETTINGS

Network Parameters	**Values**
Network layout	24 Macrocell BSs (eNodeBs)
Sectors per BS	3 sectors/cells per BS
Carrier frequency	2100 MHz
Simulation area	13.292 km^2
Bin size	30m $\times$ 30m
Antenna height	Actual site heights
Propagation Model	Aster Propagation Model (Ray-tracing)
Clutter types	15 classes
Maximum transmission power	43 dBm
Cell individual offset (CIO)	0 dB
Antenna tilt	0^o
Antenna gain	18.3 dBi
Geographical information	Digital Terrain Model (Ground heights) + Digital Land Use Map (clutter classes)

B. Data Acquisition

We acquire MDT reports with four highly used fault classes in literature for root cause analysis and self-healing frameworks: cell outage, low transmission power, excessive antenna uptilt, and excessive antenna downtilt [7], [8], [9]. Fig. 2 presents a visualization using signal-to-interference-plus-noise



(a) Normal Coverage Map (b) Cell Outage (c) Low Transmit Power (d) Excessive UpTilt (e) Excessive DownTilt

Fig. 2. Coverage maps of different network conditions. (a) Normal (b) Cell/Site outage (c) Low transmission power, this image is showing when transmission power drops to 25dBm (d) Excessive antenna uptilt, this is $+20^o$ tilt (e) Excessive antenna downtilt, -20^o tilt.

ratio (SINR) maps of different fault classes when induced on a selected cell in the designed network in the simulator. Fig. 2(a) represents a normal coverage scenario and the impact of other fault classes on cell coverage is illustrated in Fig. 2 (b-e). The parameter configuration of the four fault classes are described as follows:

1) **Cell Outage (CO):** To simulate cell outage, we deactivate the transmitter on a selected site in the simulator. This will simulate a no-coverage fault scenario around that cell. Fig. 2(b) is presenting CO scenario for highlighted cell.
2) **Low Transmission Power (LTP):** The maximum transmission power is 43 dBm for a normal BS in our designed network based on recommended value by [12]. We simulate LTP fault scenario by reducing the maximum transmit power of a cell to 25 dBm (we select this value based on the industry experience of co-authors). Fig. 2(c) shows an LTP scenario.
3) **Excessive Antenna Downtilt (EAD):** To induce excessive antenna downtilt we change the tilt value from 0^o to 20^o. Both antenna uptilt and downtilt values are selected based on co-authors' industry experience. Fig. 2(d) present an EAD scenario.
4) **Excessive Antenna Uptilt (EAU):** Normal antenna tilt is 0^o. We change the tilt value from 0^o to -20^o, we select this value based on our industry experience. The impact of EAU can be seen in 2(e) for a selected cell.

To ensure the robustness of our solution, while generating simulated MDT reports, we randomly select four cells out of 72 cells in total and induce a random fault in them through a different independent random process. In this way, not only we can have a different site (based on location in the network) in every instance but also different types of fault (CO, LTP, EAD, or EAU). We have 19933 different instances of the network, each having 4 anomalous and 68 normal cells and each anomalous cell with a different fault.

We then convert raw MDT reports into SINR radio environment maps(REMs). These REMs can be sparse due to the sparse nature of the data captured by the MDT reports from a real cellular network. We perform image inpainting to enrich the sparse REMs.

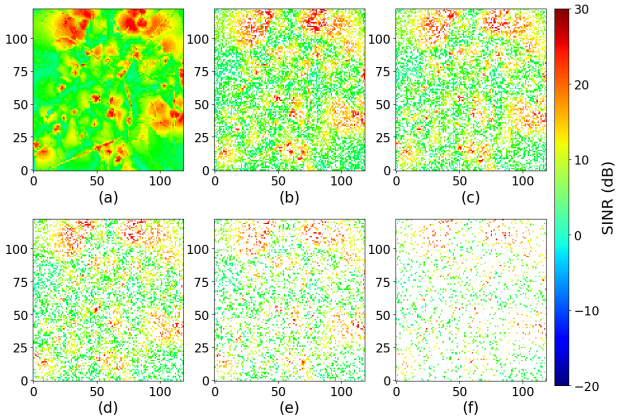


Fig. 3. Network coverage maps with various user densities (a) Full coverage map (203 UEs/cell (1101 UEs per km^2)). (b) 100 UEs/cell (550 UEs per km^2). (c) 80 UEs/cell (440 UEs per km^2). (d) 60 UEs/cell (330 UEs per km^2). (e) 40 UEs/cell (220 UEs per km^2). (f) 20 UEs/cell (110 UEs per km^2)

III. Proposed Framework

This section presents the overall proposed framework consisting of two major components; the pre-processing techniques used for data enrichment, and the root cause analysis, described as follows.

A. Data Pre-Processing

To ensure the robustness of our fault diagnosis solution we consider various user densities for generating sparse coverage maps, as shown in Fig. 3. Data pre-processing is an important part of any framework which involves machine learning. However, in our framework, its importance is even more because there are hardly any self-healing solutions involving image data for root cause analysis. A detailed view of data pre-processing is present in the first block of Fig. 4. We are converting raw MDT reports to SINR image maps, which are sparse. To address sparsity, we use image inpainting methods to enrich data before passing it on for root cause analysis. We surveyed and selected the available methods, based on accuracy and efficiency, in other domains, and applied those methods to recover missing SINR values in the REMs. We applied TELEA, biharmonic-equation based method, and frequency selective reconstruction (FSR) methods to enrich the

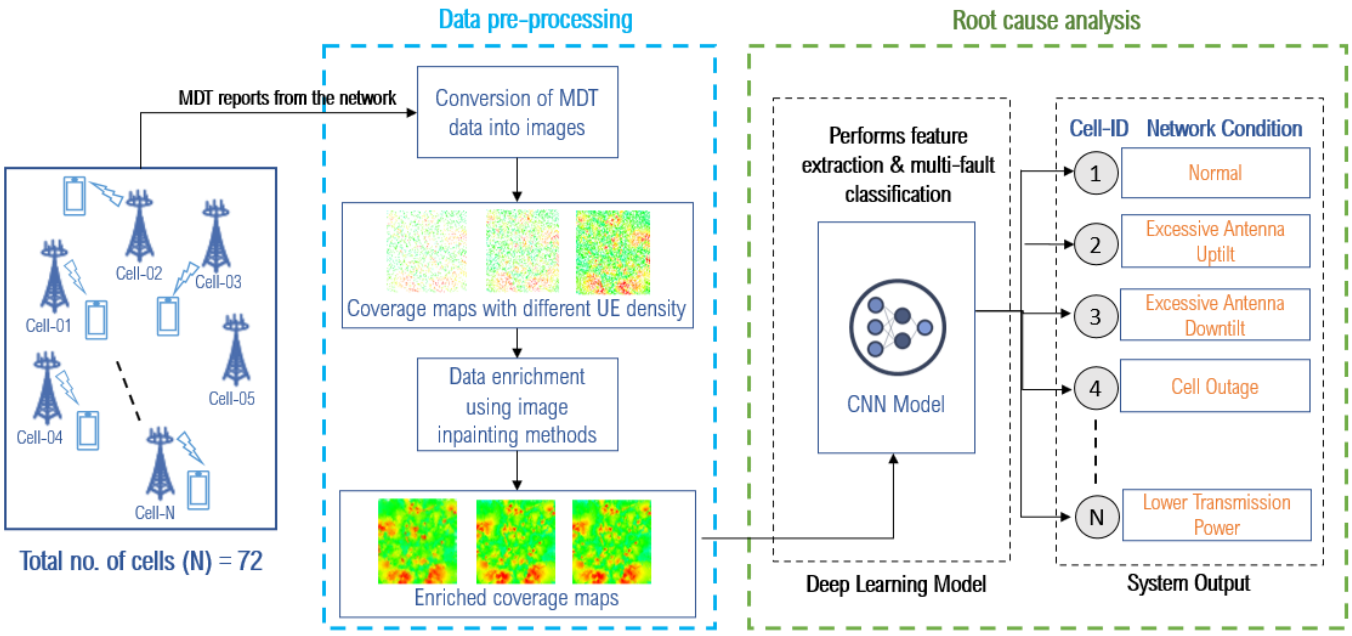


Fig. 4. Proposed Deep Learning-based Framework for diagnosing multiple faults in a multi BS network scenario Root Cause Analysis of multiple faults.

sparse coverage maps, but due to page limitation constraints, we will only describe the best performing image enrichment method below.

Frequency Selective Reconstruction (FSR): FSR reconstructs missing SINR values using Fourier basis functions from available neighboring SINR values in the REM. This is a computationally expensive method but is highly parallelizable and with the use of GPUs can achieve significantly accurate results in considerably less time [13].

To evaluate the performance of image enhancement methods we use root mean square error (RMSE), structural similarity measure index (SSIM), and peak signal to noise ratio (PSNR) as performance metrics; details and rationale of these metrics can be found in [14]. A summarized performance evaluation of these methods is given in Table II.

TABLE II
PERFORMANCE EVALUATION OF IMAGE INPAINTING METHODS

Method	RMSE (dB)	SSIM (%)	PSNR (dB)
FSR	8.6	90	22.81
Biharmonic Equation	9.3	87	21.23
TELEA	9.7	85	20.84

B. Root Cause Analysis

This section elaborates the **root cause analysis** block of Fig. 4. It explains the intuition behind using a CNN-based model and details the implementation and hyper-parameter configuration. It also elaborates the performance metrics to compare the proposed model against state-of-the-art used for root cause analysis.

1) **Why a CNN-based deep learning model?** The rationale behind using a CNN model for this study revolves around two significant reasons, the tailor-made nature of CNN for image data, and the robustness it offers towards noisy data [15], [16]. The results of this research show dominance of CNN against classical ML models on sparse and noisy data. The results are presented with a detailed explanation in Section IV. Fig. 5 provides the architectural design (no. of hidden layers, no. of neurons in each layer, kernel size, pooling, and dropout details) of the proposed CNN model. We trained models with a wide range of hyper-parameter values and tracked the loss and accuracy graphs using TensorBoard, to find an optimal CNN architecture. The hyperparameters presented in Fig. 5 are the configuration of our top-performing model, based on TensorBoard visualization. A detailed explanation of each feature extraction function is given as follows:

   a) Convolution: This operation extracts important hidden features e.g. boundary edges from the input coverage map. In our study boundaries are very important to distinguish between coverage regions of different sites. To keep the dimensions of the feature matrix unchanged we are using zero padding. Dimensions of the output matrix of convolution operation are defined as Equation 1 and 2.

$$O_r = \frac{(I_r - F + 2P)}{S} + 1 \tag{1}$$

$$O_c = \frac{(I_c - F + 2P)}{S} + 1 \tag{2}$$

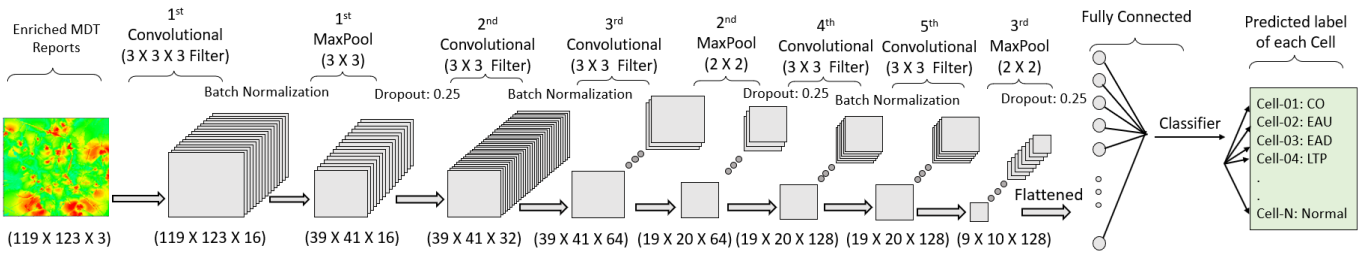


Fig. 5. Proposed CNN-based deep learning model, extracts hidden features from coverage maps and performs multi-fault classification operation.

where O_r , O_c , I_r, and I_c represent the number of rows and columns of the output and input matrix respectively, while F, P, and S represent the size of kernel, padding, and length of the stride. For example, the length of the input volume is 119×123 (as in 1st convolutional layer in Fig. 5), kernel size is 3×3, length of stride is 1 and we are using zero-padding of 1 (using Keras it can be set as 'same'). Then the length of the output feature matrix will be $119\times123\times$no.of neurons.

b) Batch Normalization: To accelerate the learning of a deep network and to address internal covariate shift, batch normalization is used. This transformation normalizes the input to a layer by maintaining its mean and standard deviation close to 0 and 1 respectively.

c) Pooling: The output feature map from a convolutional layer is location-sensitive to the input image. However, down-sampling reduces its sensitivity and makes the feature extraction process robust to changes. One of the techniques for down-sampling is pooling which essentially down-samples the features in patches of the feature map. In our case, we use max-pooling to ensure the presence of the most activated features.

2) **Performance Metrics** Since we propose a solution for the diagnosis of multiple faults in multiple BSs simultaneously (multi-label multi-class problem), it can not be fairly evaluated with accuracy (total correct predictions / total number of instances). Unlike a simple classification problem, anomaly detection problems require special performance measuring metrics, due to the biased nature of data towards normal class. Hence, we used the exact match ratio (subset accuracy) due to the nature of the problem we are solving. Every incorrect diagnosis is as vital as a correct diagnosis because it can lead to poor coverage around a BS due to a false alarm. For example, if our root cause analysis system diagnoses EAU fault for a normal site (false alarm), as a response it adds 20^o to the tilt value of that site, which will induce an EAD fault to an otherwise normal site, that will lead to poor coverage for users connected to the site.

Exact Match Ratio/Subset Accuracy (EMR): According to EMR, a diagnosis made by the model will be correct only if the network condition of all the sites in the network are diagnosed correctly. In this study, we have a network designed with 72 cells, even if the network condition of 1 out of 72 cells at a given instance is predicted incorrectly, that instance will be considered as an incorrect prediction. This is considered a very strict performance metric, but to present a critical performance analysis we include it in our results. EMR is defined by Equation 3

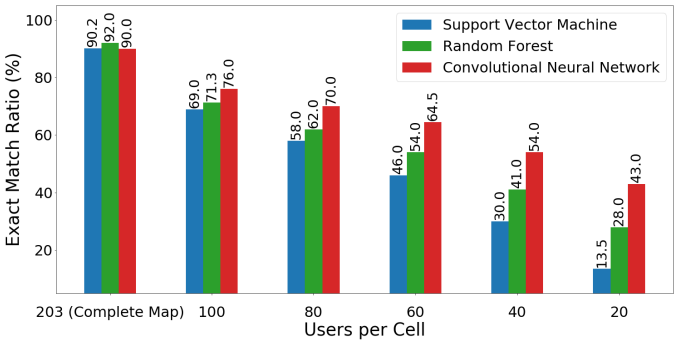
$$EMR = \frac{1}{N}\sum_{i=1}^{N} I(P_i = T_i) \tag{3}$$

where I is a proposition function, which returns 1 if the network condition of all N sites is correctly predicted, else returns 0.

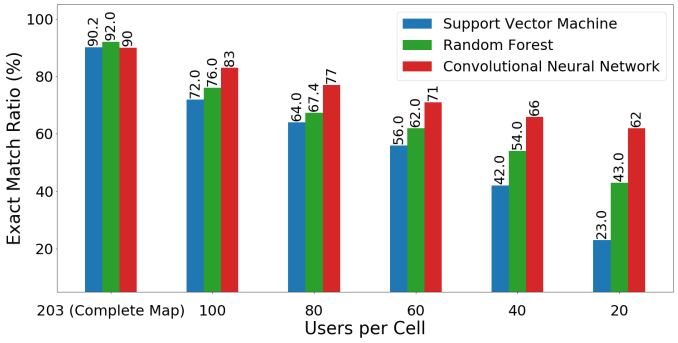
IV. Results and Comparative Analysis

We compare the performance of the proposed CNN-based solution against several methods used for detection and diagnosis of outages from literature, including SVM and Random forest (RF) [8], [9]. We present an investigation of our approach at different UE densities to analyze its efficacy in realistic settings (i.e. robustness to the sparsity of MDT reports in a cell/area). Fig. 6 presents the performance evaluation of mentioned ML/DL models on sparse data (considering various UE densities) and enriched data (enhanced using FSR image inpainting mentioned in Section III-A). In Fig. 6(a)-(b) we provide an EMR based comparative analysis on sparse and enriched data.

Fig. 6(a), shows that SVM and RF both perform slightly better than CNN on full coverage maps (complete map, coming from the simulator without a need of any data enrichment). This justifies the popularity of RF [8], [9] for self-healing in literature. But, a drastic drop in diagnosis accuracy can be seen for SVM and RF on sparse data, EMR drops from 90.2% to 69% and from 92% to 71.3% respectively, as the density of users drops from 203 to 100 users/cell. The downward trend in performance continues as the number of users decreases per cell. EMR of SVM drops to 13.5% when users per cell decrease to 20. On the other hand, the results show that the proposed CNN-based solution is showing robustness to the



(a) Exact match ratio performance on sparse data



(b) Exact match ratio performance on enriched data

Fig. 6. Performance comparison of proposed CNN-based solution against classical ML models on sparse and enriched coverage maps data (using FSR image inpainting). Note: (Data having 203 users per cell are full coverage maps, so does not require enrichment, that is why performance remains same for complete maps on sparse and enriched data) (a) Exact match ratio comparison on sparse data, (b) Exact match ratio comparison on enriched data

user sparsity and can diagnose faults with an EMR of 43%, even when users are as sparse as just 20 users/cell.

An encouraging insight is shown in Fig. 6(b) is the performance of CNN on enriched data, as compared to sparse data (shown in Fig. 6(a)). From Fig. 6(b), it is observed that it can diagnose faults with an EMR of 62% on data enriched from as sparse as 20 users/cell. This is a promising 19% improvement in EMR, as compared to RF, and a whopping 39% improvement over the SVM algorithm, in exact match ratio on equally sparse data. These results justify our intuition for using a CNN-based root cause analysis solution because it is a universal learner and inherently robust to noise, as highlighted in Section III-B.

V. Conclusion

This paper presents a framework for root cause analysis of coverage-related anomalies in emerging cellular networks. Our focus during this research remained on the practicality of the solution. Firstly, we considered a complex scenario of multiple faults in multiple BSs simultaneously. Secondly, we presented our analysis on varying levels of MDT data sparsity (user densities) in the network. To address the data sparsity issue, we used image inpainting methods to enrich sparse coverage maps. The sparsity itself and then the inpainting methods add noise to the data, which makes diagnosis harder. However, our proposed CNN-based framework maneuvers well with the noise. Results show that the CNN-based model can diagnose faults with an exact match ratio of 62% on highly sparse MDT reports data in a multi-BS multi-fault scenario.

VI. Acknowledgment

This research is based upon work supported by the National Science Foundation (NSF) under Grant numbers 1619346, 1730650 and Qatar National Research Fund (QNRF) under Grant no. NPRP12-S 0311-190302 . For more details about the projects, please visit: https://www.ai4networks.com

References

[1] A. Asghar, H. Farooq, and A. Imran, "Self-healing in emerging cellular networks: review, challenges, and research directions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1682–1709, 2018.

[2] M. Amirijoo, L. Jorguseski, T. Kurner, R. Litjens, M. Neuland, L. Schmelz, and U. Turke, "Cell outage management in LTE networks," in *2009 6th International Symposium on Wireless Communication Systems*, pp. 600–604, IEEE, 2009.

[3] P. Szilágyi and S. Nováczki, "An automatic detection and diagnosis framework for mobile communication systems," *IEEE transactions on Network and Service Management*, vol. 9, no. 2, pp. 184–197, 2012.

[4] W. Xue, M. Peng, Y. Ma, and H. Zhang, "Classification-based approach for cell outage detection in self-healing heterogeneous networks," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2822–2826, IEEE, 2014.

[5] U. Masood, A. Asghar, A. Imran, and A. N. Mian, "Deep learning based detection of sleeping cells in next generation cellular networks," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 206–212, IEEE, 2018.

[6] R. Barco, V. Wille, and L. Díez, "System for automated diagnosis in cellular networks based on performance indicators," *European Transactions on Telecommunications*, vol. 16, no. 5, pp. 399–409, 2005.

[7] A. Gómez-Andrades, P. Muñoz, I. Serrano, and R. Barco, "Automatic root cause analysis for LTE networks based on unsupervised techniques," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2369–2386, 2016.

[8] S. Bothe, U. Masood, H. Farooq, and A. Imran, "Neuromorphic AI empowered root cause analysis of faults in emerging networks," in *2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 1–6, IEEE, 2020.

[9] J. B. Porch, C. H. Foh, H. Farooq, and A. Imran, "Machine learning approach for automatic fault detection and diagnosis in cellular networks," in *2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 1–5, IEEE, 2020.

[10] H. N. Qureshi, A. Imran, and A. Abu-Dayya, "Enhanced MDT-based performance estimation for AI driven optimization in future cellular networks," *IEEE Access*, vol. 8, pp. 161406–161426, 2020.

[11] Forsk, "Atoll overview, [online], available: http://www.forsk.com/atoll-overview. [accessed: 12-oct-2018]."

[12] 3rd Generation Partnership Project, "Universal terrestrial radio access (UTRA) and evolved universal terrestrial radio access (E-UTRA); radio measurement collection for minimization of drive tests (MDT); overall description; stage 2 (release 10), 3GPP standard ts 37.320, version 10.2.0, tech. rep., june 2011."

[13] A. Regensky, S. Grosche, J. Seiler, and A. Kaup, "Real-time frequency selective reconstruction through register-based argmax calculation," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1–6, 2020.

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[15] D. Rolnick, A. Veit, S. Belongie, and N. Shavit, "Deep learning is robust to massive label noise," *arXiv preprint arXiv:1705.10694*, 2017.

[16] T. Xiao, T. Xia, Y. Yang, C. Huang, and X. Wang, "Learning from massive noisy labeled data for image classification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2691–2699, 2015.