Addressing Data Sparsity with GANs for Multi-fault Diagnosing in Emerging Cellular Networks

A. Rizwan, A. Abu-Dayya, F. Filali
Qatar Mobility Innovations Centre
Qatar University, Doha, Qatar
[arizwan, adnan, filali]@qmic.com

A. Imran
University of Oklahoma
Oklahoma, USA
ali.imran@ou.edu

***Abstract*—Data-driven machine learning is considered a means to address the paramount challenge of timely fault diagnosis in modern and futuristic ultra-dense and highly complex mobile networks. Whereas diagnosing multiple faults in the network at the same time remains an open challenge. In this context, the data sparsity is hindering the potential of machine learning to address such issues. In this work, we have proposed a data augmentation scheme comprising Pix2Pix Generative Adversarial Network (GAN) and a customized loss function never used before, to address the data sparsity challenge in Minimization of Drive Tests (MDT) data. Our proposed unique augmentation scheme generates images of MDT coverage maps with Peak signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) values of 25 and 0.97 respectively, which are significantly higher than those achieved without our customized loss function. The performance of data augmentation scheme used is further evaluated with a Convolutional Neural Network (CNN) model for simultaneously detecting most commonly occurring network faults, such as antenna up-tilt, antenna down-tilt, transmission power degradation, and cell outage. The CNN applied on the data generated from the 1% of the MDT data with the proposed augmentation scheme has lead to a gain of 550% in the detection of all classes, including the four faults and cell with normal behavior, as compared to when it is applied on the data generated without our customized loss function.**

***Index Terms*—GAN, ZSM, Fault diagnosis, Automation, Machine Learning, Deep learning, Wireless cellular networks.**

I. Introduction

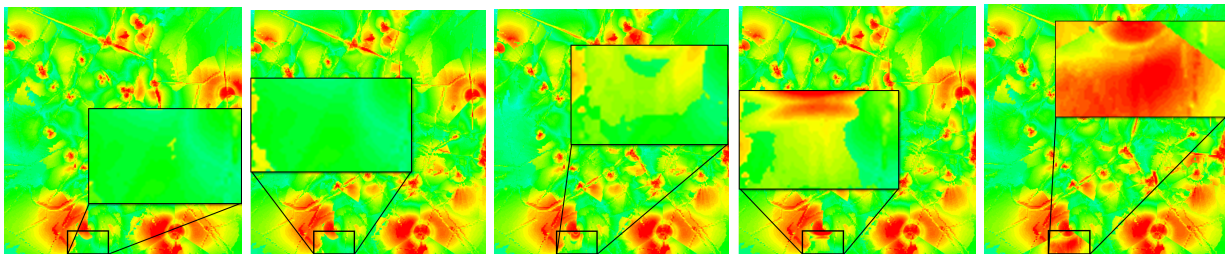
Network Performance Management(NPM) has always been a strenuous job highly dependent on skilled human resources. Network operators spend a significant share of their OPEX on NPM. But with the emergence of contemporary and futuristic technologies aka 5G, beyond 5G (B5G) and 6G, it has become essential to automate the functions of NPM. 3GPP introduced Self Organising Networks (SON) for automating network operations grouped in three main categories of self-healing, self-configuration, and self-optimization. The research on SON has lead to significant progress made so far in automating network operations in 5G, and it also provides the ground base for the projects like Hexa-X and ETSI ZSM aiming on Artificial Intelligence (AI) driven automation in B5G and 6G [1]. Research work on self-healing function in SON has set the premises for the automated NPM by introducing the concept and solutions for the automation of fault detection and diagnosis. AI equipped SON heavily relies on data-driven machine learning for the automation of network operations.

An important and equally challenging task of NPM in emerging and futuristic networks is automating the detection and diagnosis of cells facing some type of technical issues. Thanks to 3GPP release 10 for introducing Minimisation Drive Test (MDT) that enabled network operators to collect key data from users' equipment, rather than conducting drive test incurring too much operational cost and unnecessary delays. MDT data and machine learning can help in devising solutions for the automation task here but the sparsity of MDT data limits its potential. There exist different approaches like in-painting techniques and machine learning based models that can help in addressing data sparsity. They all have there pros and cons. But one machine learning tool recently developed, Generative Adversarial Network (GAN), has gathered the much attention of the researchers for its performance and efficacy in applications like addressing data sparsity.

Different GAN architectures proposed in recent research mostly aim at image-based applications like image completion, image super-resolution, image transformation, etc. But these GAN are proved to be very effective for such tasks as addressing data sparsity in the form of image completion. The key advantage of GANs is that, instead of just creating copies or averaging out values, they learn the data distribution patterns and create the new samples from those distributions. The newly generated samples are similar but not the same as the original ones. This little variance in the GAN generated data samples can reflect the real-world variations in the outcomes of the same network deployment schemes.

In this article, we propose a novel scheme for automating the detection of multiple commonly occurring network faults from the sparse MDT data. To the best of our knowledge, it is the first time that GANs are used to address sparsity challenges in mobile networks MDT data, presented as images, for detecting multiple faults in the network. The contributions we have made in this study are summarised as follows:

- We have introduced a Pix2Pix GAN-based unique data augmentation scheme to address data sparsity in MDT reports. The proposed scheme can generate a complete coverage map from the 1% data.
- We have introduced our own customized perceptual loss function, never used before, that has increased the performance of the GAN model manifolds.
- We have introduced a CNN model that can successfully



(a) Transmission power (b) Cell outage (c) Antenna up-tilt (d) Antenna down-tilt (e) Normal coverage

Fig. 1: Signatures of normal coverage and different network faults

detect multiple faults in the network even in sparse data and yields manifold gain for the detection of multiple faults on the data augmented with our proposed scheme.

The rest of the paper is organized as follows. The Section II offers a survey on state-of-the-art. Section III provides a brief description of data generated from the simulator. Section IV discuss the methodology adopted for the data enrichment and fault diagnosis. Section V presents discussion on the results and Section VI concludes the study.

II. Related Works

There are many studies that propose schemes and solution for cell outage detection [2], [3] and many of them take it as anomaly detection problem. But, there are very limited studies that focus on the detection of other network issues than the cell outage,that may cause sub-optimal performance in specific cells in the network. The studies on the diagnosis of such faults are even rare [4]. Whereas the detection of multiple faults, instead of just one single fault, is still an open research issue. One of the such studies is [5], where the authors have applied a semi-supervised learning scheme on real network Call Detail Record (CDR) data for grouping cells into multiple classes based on their performance. They have detected and diagnosed cells with sub-optimal performance and identified reasons of sub-optimal performance. But this study used CDR data form real network and rely on the input from the expert for the labelling of cells based on their performance, and presents only broad level network performance issues.

In a recent study [4] authors have proposed solution for the diagnosis of commonly occurring multiple faults such as site outage, transmission power, antenna up-tilt and antenna down-tilt. In addition to issues addressed in [4] authors in study [6] have introduced solution for diagnosing even more advanced network issues like Too Late Handover (TLHO), Inter-Cell Interference (ICI), and Cell Overload (CO). But main bottleneck of these two studies is, they are using complete coverage map generated from simulators. Whereas in the real network data points are very sparse. But the aspect of addressing data sparsity for the NPM management tasks is missing in literature.

Recently GANs have been very popular as a tool for data augmentation and have been used for diverse applications. Most popular applications of GANs are image based, like for image generation, image to image translation, image super-resolution, semantic segmentation etc., [7]. For images, two main application of GANs have been image quality enhancement and image completion. GANs have been used to successfully generate complete images from the incomplete images [8] and improve the resolution of the images [9]. The same concepts can be applied in our case for generating complete MDT network coverage map from the incomplete coverage map. This approach has potential to address the data sparsity issue in MDT reports.

III. Dataset description

In this study for the MDT data generation, we have used Atoll, an RF planning software capable to generate real scenario data. To make sure that the data generated is close to real-world network data, we have considered network topology and parameters from a real mobile network operator in Brussels. Besides that the network is also simulated over an area in Brussels, Belgium considering 15 types of clutters based on different terrain and environmental profiles. The simulated network comprises 24 sites (macrocells) and 72 transmission antennas (cells), with 3 antennas deployed on each base station, overall covering an area of 15 km^2. Detail about the network parameters used is listed in Table I.

Using that simulation environment, from the MDT reports we have generated SINR based network coverage map of 72 cells with 68 cells performing normally and four randomly selected cells are induced with any of the four faults listed below. Samples of the each possible status of the cells are shown in Figure 1. The color range from green to red present

TABLE I: Network parameters used for MDT data generation

Network Parameters	**Values**
Propagation Model	Aster Propagation Model (Ray-tracing)
Maximum transmission power	43 dBm
Cell individual offset (CIO)	0 dB
Antenna tilt	0
Antenna gain	18.3 dBi
Carrier frequency	2100 MHz
Network layout	24 Macrocell BS
Simulation area	15 km2
Transmitters (sectors) per BS	3
Base station height	Actual site heights
Clutter types	15 classes
Geographical information	Digital Terrain Model (Ground heights)
	Digital Land Use Map (clutter classes)

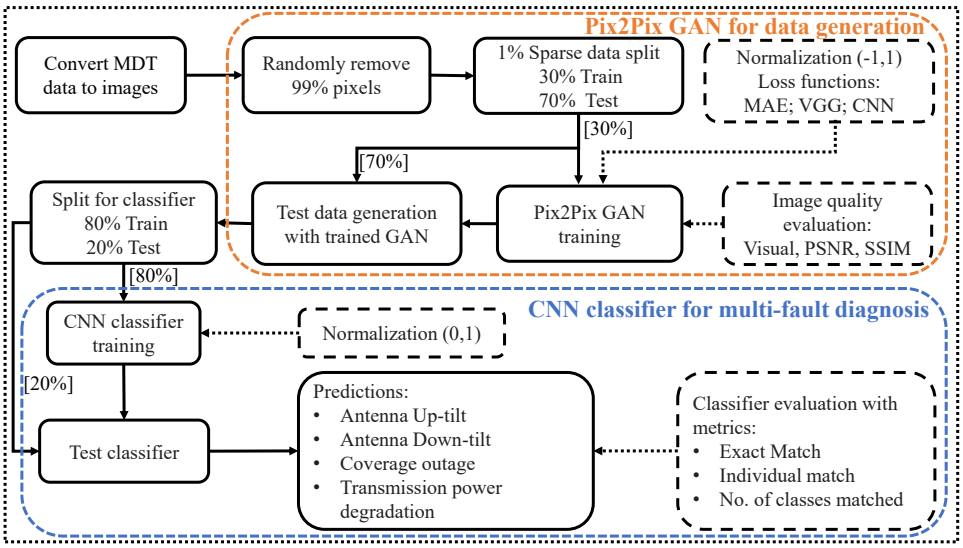


Fig. 2: Process flow diagram for MDT data augmentation and classification of multiple faults

poor to good SINR values in the network. In total 22,864 coverage map images are generated.

- Low Transmission Power (LTP): Maximum transmission power of a normally functioning cell is 43 dBm. But, based on common experience in the industry, here we have reduced it to 25 dBm.
- Cell Outage (CO): In this type of issue cell is not functional at all, it can be caused by transmitter deactivation.
- The antenna is tilted by an angle of -20^o from the standard normal antenna angle of 0^o.
- This fault is induced by changing the tilt value of the simulation from 0^o to 20^o.

IV. Proposed Scheme for Data Augmentation and Fault Diagnosis

The important steps of the methodology adopted in this study to accomplish two main tasks, addressing data sparsity and diagnosing multiple network faults along with a preliminary step of generating relevant sparse data images, are listed in Figure 2 and discussed in this section.

A. Generating sparse data

One of the main objectives of this study is to address the challenge of data sparsity in mobile networks. The sparse data images are, therefore, generated from the complete images of simulated coverage maps already discussed in Section III. Since each pixel in the images is a representative of an SINR value from a UE, therefore, to create an extreme data sparsity scenario, for this study, we have removed 99% of the pixels from the images that left only 1% of as much information as present in original complete images. The original images are resized to 256 by 256 dimensions before pixel removal for convenience in image processing and consistency in implementation of deep learning. Since each pixel represents a user, therefore, a complete 256 by 256 image means, the $15km^2$ coverage area in the image has 65,536 users and the number of users per cell is around 910. So when we remove 99% of the pixels then the new incomplete images have around 655 users in the $15km^2$ coverage area i.e around 9 users per cell or base station BS and around 43 users per km^2.

B. Generating enriched data using GAN

Next important task is developing a scheme for generating complete network coverage map from the incomplete coverage map images of 1% data points. For that purpose we have used a conditional GAN architecture, Pix2Pix-GAN proposed in [10] with cutomized perceptual loss function in [11] also used in [9] with VGG-19 model. Pix2Pix GAN is a conditional GAN that takes an input image x, like an incomplete network coverage map shown in Figure 3a, along with a noise vector z, to learn a mapping function from x to y the target image which is the complete coverage map as shown in Figure 3b and to generate the outcome image $\hat{y}$ like shown in Figure 3c. We provide noise only in the form of dropout, applied on several layers of our generator at both training and test time. The dropout noise is observed to lead to minor stochasticity in the output of our model making generated images more close to the real network scenarios.

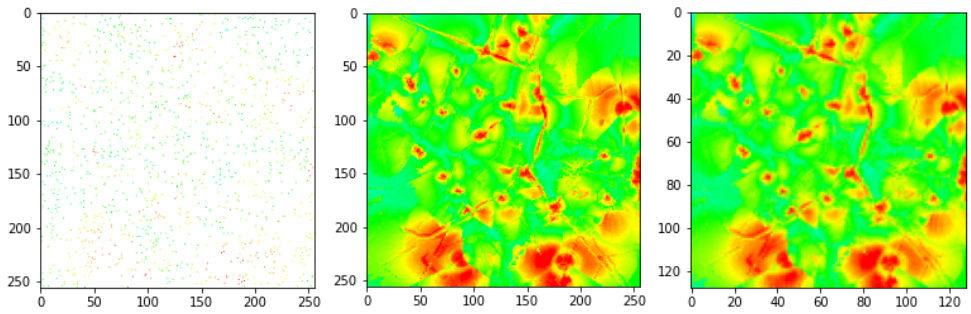
Like the standard architecture of GANs, Pix2Pix also comprises a generator and a discriminator. Where the main objective of the Pix2Pix proposed in [10] is:

$$G^* = arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L_1}(G) \quad (1)$$

Where $\mathcal{L}_{cGAN}(G, D)$ presents the objective function of typical conditional GAN computed as follows:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[1 - \log D(x, G(x, z))] \quad (2)$$

Here the generator G tries to minimize this objective against an adversarial discriminator D that tries to maximize it with following approach: The authors in [10] included the other part, one of the commonly used loss function, the L_{L_1}, to



(a) Incomplete 1% sparse data (b) Original complete map (c) Generated from the GAN

Fig. 3: Sample of images from input(left),target (middle),and output (right) in Pix2Pix GAN

compute the distance between the pixel values of generated and target image as follows:

$$\mathcal{L}_{cGAN}(G, D) = \mathop{\mathbb{E}}_{x,y.z}[\|y_i - \hat{y}_i\|_1] \tag{3}$$

However, the above MAE optimization can lead to satisfactorily high PSNR, but often it lacks high frequency content which results in perceptually unsatisfying solutions. We have, therefore, introduced a perceptual loss also known as feature reconstruction loss which is a type of content loss introduced in [11]. Rather than encouraging the pixel to pixel match of the output image $\hat{y}$ and target image y, we encourage GAN to learn similar feature representations as computed by the loss network ϕ. while processing the image x if the ϕ_j (x) is the activation of the jth layer of the network ϕ and j is a convolutional layer then $\phi_j(x)$ results in a feature map of shape $C_j \times H_j \times W_j$. The feature reconstruction loss is the normalized, squared Euclidean distance between feature representations computed as follows:

$$\ell^{\phi,j}_{feat}(\hat{y}, y) = \frac{1}{C_j W_j H_j}\|\phi(\hat{y}_i) - \phi(y_i)\|_2^2 \tag{4}$$

So in this study we have used two network functions ϕ to compute the above perceptual loss. One ϕ is VGG-19 inspired from [9] and the other ϕ is our own CNN network. We have computed the perceptual loss as an additional loss to the loss computed in equation IV-B. As a result the objective function in our case becomes:

$$\mathrm{G}^* = arg\min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda\mathcal{L}_{L_1}(G) + \alpha\ell^{\phi,j}_{feat}(\hat{y}, y) \tag{5}$$

where $\lambda = 100$ and $\alpha = 10^{-3}$ as suggested in [10] and [9] respectively. So using the using the objective function in equation IV-B in Pix2Pix GAN architecture proposed in [10] we have generated the network coverage maps from the sparse data coverage map. We have generated images using the perceptual loss in IV-B with two network VGG and our own network CNN separately and results for the both are compared in the results section.

Once the images are generated from the GAN a crucial task is to evaluate the quality of the generated images. Common evaluation metrics measure the pixel to pixel euclidean distance which does not take the contextual or structural information into consideration. Here we have therefore used two popular and relevant metrics, Peak Signal to Noise Ratio (PSNR) to evaluate the pixel to pix match and structural similarity index measure (SSIM) to evaluate structural similarity. PSNR and SSIM are good indicative of the image quality but better PSNR and SSIM values not necessarily mean that the images generated are true representation of original images. Apart form these evaluation tools, we have majorly relied on the performance of our classifier for the selection of images generated by the Pix2Pix GAN.

C. Classifying multiple faults

In total around 16000 images are generated from Pix2Pix GAN, which makes 70% of the total images generated from the simulator. The generated images have all pixel values and visually look similar to the original images. After the enriched images generation, the next step is to identify the four faults present in the network using those enriched images. For that goal, we fine-tuned many of the popular pre-trained models and also applied a custom CNN architecture. For the development of the classifier, we split the GAN-generated data into train and test data, such that 80% of the data is used for training classifiers and 20% of the data is used for evaluation.

The classifiers are developed such that they not only predict the four BS having any issue and type of the issue present but also predict which BS performs normally. Hence the classifier has 360 predicates in total, 72 cells having five possible outcomes, performing normal, or having issues due to up-tilt, down tilt, transmission power degradation, or complete outage. Overall accuracy can not be the representative metrics to evaluate the performance of such a classification scheme. Since we are trying to predict multiple faults so it is important to know that in how many cases(images) status of all 72 cells are detected correctly from five possible outcomes. Here, therefore we have used the exact match as a metric to find in how many cases all 360 predicates are predicted correctly. Besides that, we have also calculated the class-specific accuracy to reflect the accuracy rate of identification of any individual fault or normal behavior. The third factor we evaluate is the performance of the classifier for detecting the number of faults.

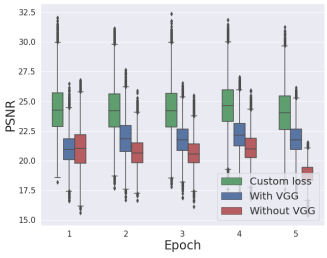
V. Results and Discussion

In this section, first, we present the results produced from the Pix2Pix GAN data augmentation scheme applied for addressing the data sparsity. Later, the results of the multi-fault classification scheme are presented.

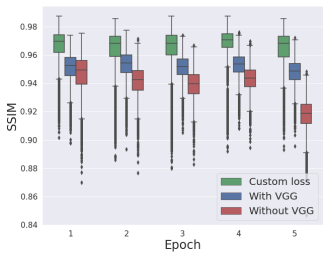
A. Results of data enrichment with Pix2Pix GAN scheme

A representative sample of the images generated as the result of the Pix2Pix model-based augmentation scheme is shown in Figure 3c. In Figure 3, the image 3a on the left presents the sample from the sparse data input images for the Pix2Pix GAN scheme used for data augmentation. Images like 3a are produced by removing 99% of the pixels from the complete coverage maps. The image 3c on the right presents a sample from the generated images with our proposed data augmentation scheme. Even visually it can be seen that 3c looks almost the same as the original complete coverage map 3b in the middle.

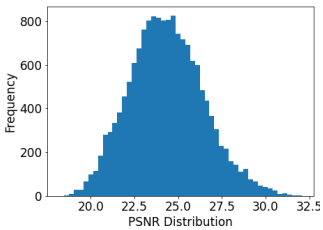
Figures 4a and 4b present the PSNR and SSIM value for the five epochs where the data generated showed the best five performances on classifier as reflected from the Figures 5 and 6. So when we applied the VGG and our custom loss functions in our Pix2Pix GAN models, it not only improved the PSNR and SSIM values for data generated at each epoch as shown in the Figures 4a and 4b but the performance of the classifier also improved significantly as it can be seen from the Figure 6. The histograms in Figures 4c and 4d present PSNR and SSIM values of all images generated from the GAN model with our custom loss function trained for epoch 40. The data generated from this epoch yields the best performance in the classification for the multi-fault diagnosis. It can be seen that, here, for the majority of the images generated PSNR value is between 22 and 27 similarly the SSIM for the majority of images generated is between 0.96 and 0.98. It is an indicator that the GAN has learned for features for the majority of the images and there are fewer images for which GAN could not grab the desired information.



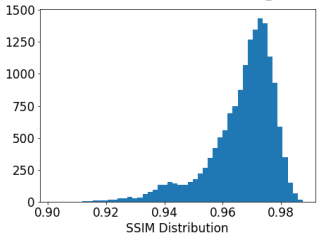
(a) PSNR over best five epochs



(b) SSIM over best five epochs



(c) PSNR Histogram of 40^{th} epoch



(d) SSIM Histogram of 40^{th} epoch

Fig. 4: SSIM and PSNR for data genrated from Pix2Pix GAN

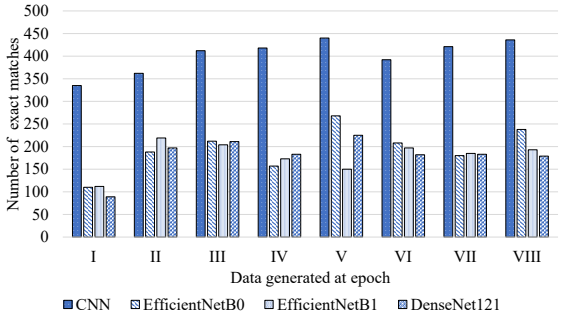


Fig. 5: Performance of classification models on VGG-GAN data

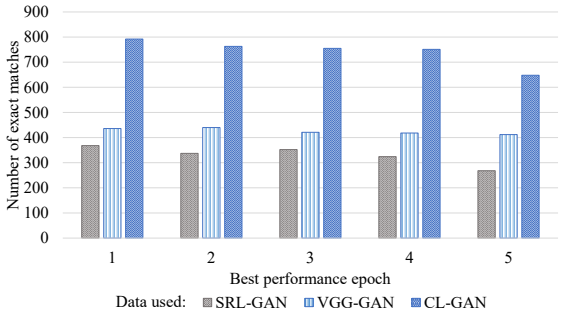


Fig. 6: Performance of CNN with different loss functions

B. Results of classification schemes

The performance of the classifiers is evaluated on the original complete map data, sparse data (with 1% data points), data generated from sparse data with Pix2Pix GAN using simple reconstruction loss function (SRL-GAN data), and data generated from Pix2Pix GAN with perceptual loss functions, one exploiting VGG model, labelled as VGG-GAN data, and the other with our Customized Loss Function (CL-GAN data) based on our own CNN classifier. The results of the some best performance predefined popular image-based classification models using transfer learning along with the results of our own CNN model, all on the data generated from VGG-GAN up to eight epochs, are presented in Figure 5. It can be seen that the CNN proposed outperforms predefined models. The performance of the predefined model is observed to be even more poor on the raw sparse data images, whereas CNN could identify all the classes in around 119 samples out of 3201 test samples when trained and evaluated on sparse data.

Since CNN has shown significantly better performance as compared to other models, therefore, further detailed results are presented for CNN only. Figure 6 shows the performance of CNN for correctly detecting all five classes when applied on the detests generated with Pix2Pix SRL-GAN, VGG-GAN, and CL-GAN. The results are presented for those five epochs where CNN has the best five performances, 1 for the first best and 2 for the second-best performance. Figure 6 clearly shows that the use of perceptual loss improves the quality of images and the performance of the classifier subsequently. By

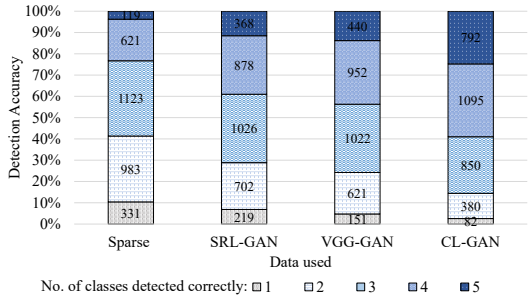


Fig. 7: Number of classes detected correctly by CNN on data generated from GAN with different loss functions

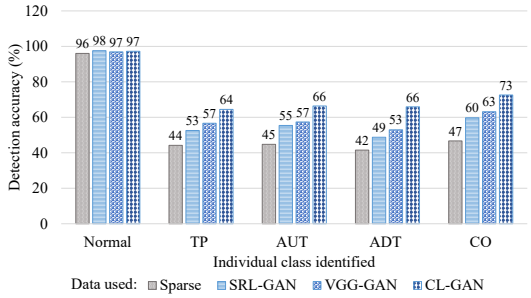


Fig. 8: Performance of CNN with different loss functions for the detection of individual class

detecting all five classes for 72 cells in 440 coverage maps, the use of the VGG-loss increased the exact match rate when compared to the data generated with simple SRL-GAN where for 368 coverage maps all classes were correctly predicted. Our customized loss function has lead to further significant improvement in the exact match rate by detecting all classes in 792 images out of 3201.

Figure 7 further reflects the potential of our augmentation and classification scheme. It shows that for the best quality data set generated with CL-GAN, the classifier could identify all classes including four faults in around 25% of the test samples, whereas it could identify four and three classes in almost 35% and 26% of the test samples. So overall it can be seen that in around 86% of cases at least three classes could be identified correctly. For the exact match detection rate for all classes only, it can be clearly seen that our proposed data-augmentation scheme comprising Pix2Pix CL-GAN, yields a gain of around 115% as compared to data generated by SRL-GAN and a gain of around 565% as compared to all fault detection rate on raw data.

The detection accuracy of each individual class is shown in Figure 8, where it can be seen that the normal cell behavior could be detected easily even when CNN is applied on the raw data leading to around 96% detection accuracy. The other comparatively easily detected fault is coverage outage which is detected with an accuracy of around 73% for the CL-GAN data. The rest of the faults have almost the same detection rate of around 65% for the CL-GAN data.

VI. Conclusion

In this study, we have presented a comprehensive scheme comprising a unique data augmentation method for addressing data sparsity and a classification model to diagnose multiple faults in the network. As part of the data augmentation scheme, we have introduced a customized perceptual loss that helps a Pix2Pix-GAN model generate images of high quality with PSNR and SSIM values of around 25 and 0.97 respectively. We have evaluated the performance of our augmentation scheme using a CNN model that yields a gain of 550% in the detection of all five classes, including four faults as compared to when it is applied on the sparse data sample with 1% of the information available. Our proposed scheme can not only help in addressing the data sparsity challenge in MDT but also provides a solution for the multiple-fault diagnosis an important task in network performance management.

ACKNOWLEDGMENT

This work is supported by the Qatar National Research Fund (QNRF) (a member of The Qatar Foundation) under Grant No. NPRP12-S 0311-190302. The statements made herein are solely the responsibility of the authors.

References

[1] K. Koufos, K. Haloui, M. Dianati, M. Higgins, J. Elmirghani, M. Imran, and R. Tafazolli, "Trends in intelligent communication systems: Review of standards, major research projects, and identification of research gaps," *Journal of Sensor and Actuator Networks*, vol. 10, no. 4, p. 60, 2021.

[2] A. Asghar, H. Farooq, H. N. Qureshi, A. Abu-Dayya, and A. Imran, "Entropy field decomposition based outage detection for ultra-dense networks," *IEEE Access*, 2021.

[3] T. Zhang, K. Zhu, and D. Niyato, "Detection of sleeping cells in self-organizing cellular networks: An adversarial auto-encoder method," *IEEE Transactions on Cognitive Communications and Networking*, 2021.

[4] 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)*. IEEE, 2020, pp. 1–5.

[5] A. Rizwan, J. P. B. Nadas, M. A. Imran, and M. Jaber, "Performance based cells classification in cellular network using cdr data," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–7.

[6] W. Zhang, R. Ford, J. Cho, C. J. Zhang, Y. Zhang, and D. Raychaudhuri, "Self-organizing cellular radio access network with deep learning," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 429–434.

[7] L. Wang, W. Chen, W. Yang, F. Bi, and F. R. Yu, "A state-of-the-art review on image synthesis with generative adversarial networks," *IEEE Access*, vol. 8, pp. 63 514–63 537, 2020.

[8] R. Wang, Z. Fang, J. Gu, Y. Guo, S. Zhou, Y. Wang, C. Chang, and J. Yu, "High-resolution image reconstruction for portable ultrasound imaging devices," *EURASIP Journal on Advances in Signal Processing*, vol. 2019, no. 1, pp. 1–12, 2019.

[9] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[10] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[11] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European conference on computer vision*. Springer, 2016, pp. 694–711.