Neuromorphic AI Empowered Root Cause Analysis of Faults in Emerging Networks

Shruti Bothe, Usama Masood, Hasan Farooq and Ali Imran
AI4Networks Research Center, Dept. of Electrical & Computer Engineering, University of Oklahoma, USA
Email: {shruti, usama.masood, hasan.farooq, ali.imran}@ou.edu

***Abstract*—Mobile cellular network operators spend nearly a quarter of their revenue on network maintenance and management. A significant portion of that budget is spent on resolving faults diagnosed in the system that disrupt or degrade cellular services. Historically, the operations to detect, diagnose and resolve issues were carried out by human experts. However, with diversifying cell types, increased complexity and growing cell density, this methodology is becoming less viable, both technically and financially. To cope with this problem, in recent years, research on self-healing solutions has gained significant momentum. One of the most desirable features of the self-healing paradigm is automated fault diagnosis. While several fault detection and diagnosis machine learning models have been proposed recently, these schemes have one common tenancy of relying on human expert contribution for fault diagnosis and prediction in one way or another. In this paper, we propose an AI-based fault diagnosis solution that offers a key step towards a completely automated self-healing system without requiring human expert input. The proposed solution leverages Random Forests classifier, Convolutional Neural Network and neuromorphic based deep learning model which uses RSRP map images of faults generated. We compare the performance of the proposed solution against state-of-the-art solution in literature that mostly use Naive Bayes models, while considering seven different fault types. Results show that neuromorphic computing model achieves high classification accuracy as compared to the other models even with relatively small training data.**

***Index Terms*—Self Organizing Networks, data mining, fault diagnosis, future mobile networks, neuromorphic computing, self-healing.**

I. Introduction

Unprecedented growth in the cellular network usage in the last decade has made system operations more complex. Billions of dollars are spent in the United States alone by cellular operators to detect and diagnose faults [1]. With the genesis of 5G ultra-dense network, fault management will be a primary challenge. It is comprehensible to foresee that with such an amalgam of technologies and operational complexity required to configure, operate, optimize and maintain the networks the biggest challenges in the emerging cellular networks [2] would be to automate the fault diagnosis process.

Even single fault in the network system always produces a large amount of alarm information. Operators then have to depend on domain experts to diagnose the exact cause and devise a solution. The drawback of this is that it is very time consuming and is prone to human errors given the unfathomable complexity of the system. From the perspective of a network operator, achieving timely and accurate diagnosis of the cause of the faults is critical for both improving subscriber perceived experience and maintaining network reliability.

Network applications necessitate that network fault issues be resolved in a very short time. The complexity of fault diagnosis has led to emergence of current solutions that use a combination of expert input and an automated system. Although research is on-going, most of these fault diagnosis systems were designed on make-shift and incoherent basis. This involved simply dissemination of human expert knowledge into a system that has the capability of solving problems in a specific domain. But, by taking into consideration future emerging cellular networks, their complexity and shrinking profit margins; minimizing human interaction for fault diagnosis for Self Organizing Networks (SON) would unquestionably be desirable [1]–[3]. This means that automation for reducing costs, handling complexity and maximizing resources efficiency will not only become a necessity, but future cellular networks will depend on it [1]. The main task of an automated fault diagnosis tool is to identify the root cause of problem. This paper focuses on the fault diagnosis process since the accuracy and timeliness of the fault management process is highly dependent on it.

A. Relevant Work

Several recent studies have explored fault diagnosis in cellular networks and primarily focus on certain predefined faults. Authors in [4] have proposed the use of Bayesian networks to diagnose faults. In [5], authors described the functionalities that targeted automated troubleshooting in Radio Access Network (RAN) by using 'if-then' rules. Similarly, in [6] an approach for automated diagnosis in Universal Mobile Telecommunications System (UMTS) networks using a Bayesian network (BN) is proposed. Alarms are modeled as discrete random variables with two states: OFF or ON. KPIs are modeled as discrete random variables with two, three, or more states. The authors in [7] propose a system for automated diagnosis using Naive Bayesian (NB) classifier, which is used to identify the cause of fault in 3G, GSM/GPRS or multi-systems networks. In [8] authors presented a self-healing framework for detecting and compensating cell outages and evaluated it in realistic environment. Authors in [9] have compared thresholds obtained with 3 different algorithms such as knowledge-based method (TEXP), Entropy Minimization Discretization (EMD) and Selective Entropy Minimization Discretization. In [10],

978-1-7281-7127-2/20/$31.00 ©2020 IEEE

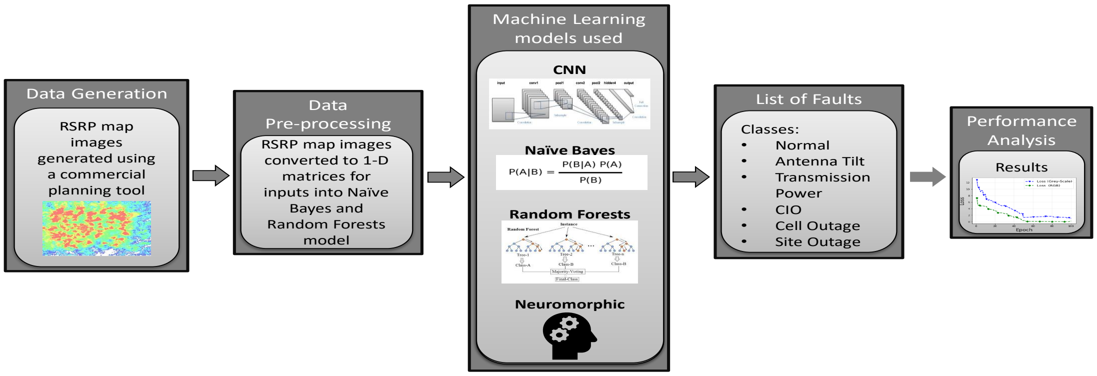

Data Generation
RSRP map images generated using a commercial planning tool
Data Pre-processing
RSRP map images converted to 1-D matrices for inputs into Naïve Bayes and Random Forests model
Machine Learning models used
CNN
Naïve Bayes
P(A|B) = P(B|A) P(A) / P(B)
Random Forests
Neuromorphic
List of Faults
Classes:
Normal
Antenna Tilt
Transmission Power
CIO
Cell Outage
Site Outage
Performance Analysis
Results

Fig. 1: Proposed framework for fault diagnosis in cellular networks

authors used unsupervised clustering for detection and diagnosis of the root cause of faults. Five classification algorithms, namely Chi-squared automatic interaction detection (CHAID), quick unbiased efficient statistical tree (QUEST), Bayesian network, support vector machine (SVM) and classification and regression tree (CRT) were used. Authors in [11] and [12] have designed a framework for detection and diagnosis to find the root cause which is based on expert knowledge. In [13], a few improvements to this framework are discussed. The work has included more sophisticated profiling and detection capabilities.

B. Contributions

Although all the above mentioned works contribute on similar lines in terms of automating the fault diagnosis process, the novel contributions of this paper are outlined below:

1) To the best of the authors' knowledge, none of the existing studies have used Convolutional Neural Networks (CNN), Random Forests (RF) and neuromorphic computing models for fault diagnosis.
2) Existing studies use small pre-existing data sets limiting the number and types of faults that can be diagnosed. To overcome this limitation, in this study for the first time we have generated a synthetic data set. The entire process of generating the dataset has been automated in a ray-tracing based commercial planning tool.
3) Most existing studies still rely on input/domain knowledge from human expert in one form or other. In the proposed solution, no input from human experts is needed. The models are purely trained on Reference Signal Received Power (RSRP) map images, which are labeled as faults. This paper presents a pure deep learning model where no additional external information provided by experts is needed.
4) The final contribution of the paper is a comparison between NB solution that is the most widely investigated approach in literature for fault diagnosis, and our proposed CNN, RF and neuromorphic based solutions. Results show that all the proposed models outperform the prevailing NB fault diagnosis solutions and RF even outperforms CNN.

The rest of the paper is structured as follows: Data generation and pre-processing methodology is presented in Section II. Section III presents the CNN, NB, RF and neuromorphic model frameworks used in this paper. Numerical results are presented in Section IV. Section V concludes the paper.

II. Data Generation and Pre-Processing

1) Data Generation: The dataset used in this paper consists of RSRP map images generated using a ray tracing based commercial planning tool used for wireless network design and optimization. The area under consideration is the city of Brussels spanning over 800 km^2 . As seen in Figure 2, 291 transmitters have been positioned around the city over 15 clutter classes, which have been included in this framework and each base-station (BS) has 3 transmitters connected to it. For our analysis, we have considered 120 transmitters i.e. 40 BS. Also, to minimize the boundary effect, all transmitters are kept functional.

Table I summarizes the network default settings before any fault is induced. In this data generation technique, 7 faults have been generated one at a time. These are generated in the BS in the area of interest. RSRP map images were created for each instance when an error is triggered and are later fed into the aforementioned models. Figure 3 represents the RSRP map images for 'transmitter off' and 'site outage' faults in one of the BS as encircled.

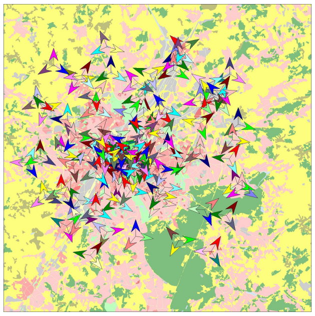


Fig. 2: Spatial representation of the Base-Station positions

Seven most commonly occurring faults in real networks were simulated to generate the synthetic data as described below:

1) **Cell Outage** (C.O): This error was created by turning each transmitter off one at a time.
2) **Site Outage** (S.O): A BS consisting of 3 transmitters was turned off at one time.
3) **Transmission Power** (TxP): An error in the transmission power of the BS was induced. The default value of 43 dBm was varied between 25 dBm to 35 dBm with a step size of one to trigger the error.
4) **CIO Positive** (CIO+): To account for handover parameter error, the Cell INdividual Offset (CIO) was increased to 10 dB from its default value of 0 dB.
5) **CIO Negative** (CIO-): A second error in CIO was created by reducing the CIO value from its default value of 0 dB to negative 10 dB.
6) **Antenna Uptilt** (A$_U$): The tilt angle of the antenna was increased from 0 degrees to 25 degrees.
7) **Antenna Downtilt** (A$_D$): Antenna tilt was decreased from its original value of 0 degrees to negative 25 degrees.

2) Data Pre-Processing: The dataset consists of 1961 colored images of size 553x578 split into 8 classes namely, Normal Scenario, C.O., S.O., TxP, CIO+, CIO-, A$_U$ and A$_D$. The train test division is 70:30. The advantage of using RSRP map images over raw data for fault diagnosis, as most prior studies do, is that there is no need for identification of smart input features or labeling required from experts. No preprocessing of the data is done before feeding it to the CNN. Since images are simply RSRP maps, in real network such maps can be generated using Minimization of Drive Test (MDT) data and then fed to the proposed diagnosis framework.

TABLE I: Network Scenario Default Settings

System Parameters	Values
Cellular Layout	120 Macrocell sites
Sectors	3 sectors per BS
Simulation Area	800 km^2
Path Loss Model	Ray-Tracing
Land Cover (Clutter) Types	15 different classes
BS Transmit Power	43 dBm
Cell Individual Offset	0 dBm
Antenna Tilt	0 deg
Antenna Gain	18.3 dBi
Carrier Frequency	2100 MHz
Geographic Information	Ground Heights + Building Heights + Land Use Map

However, since NB and RF do not take images as inputs, 1-D matrices corresponding to the RSRP map images are generated. The associated label for each class are induced in the matrix. A digital grayscale image is represented by a pixel matrix. In such an image each pixel is represented by one integer from {0,...,255}. Zero represents black pixels and 255 as white. RGB images are represented with three grayscale images matrices (one for each red, green and blue color). Figure 1 summarizes the framework for fault diagnosis model. All the RSRP map images are also converted to grayscale to be fed in these models to compare the networks' performance between grayscale and RGB input data. A representation of this is shown in Fig 4.

III. Proposed Model Framework

A. Convolutional Neural Network Model

CNNs are Artificial Neural Networks (ANN) that have been most commonly used for analyzing images. This pattern detection is what makes CNNs so useful for images analysis. The hidden convolutional and non-convolutional layers makes them different from the standard multi-layer perceptron(MLP). Since its introduction, CNNs have been applied in various research such as sentence classification [14], face recognition [15], traffic signal recognition [16] and many more. 3 operations mainly:convolution, transformation and pooling are performed on the input image. In this model, we have provided RGB and grayscale RSRP map images.

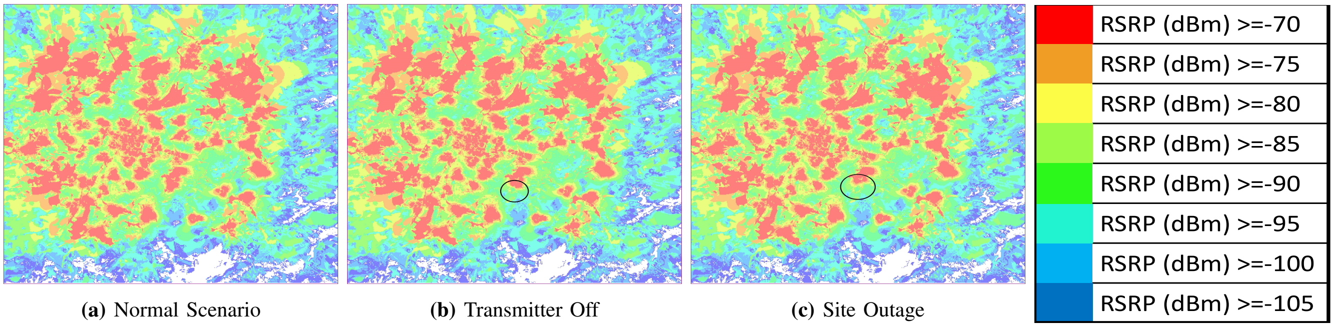
The pre-defined number of filters in the convolutional layers learn the features of the input image. During a convolution, the filters effectively slide over the input feature map grid horizontally and vertically, one pixel at a time, extracting each corresponding tile. For each filter-tile pair, CNN performs element-wise multiplication of the filter matrix and the tile matrix, and then sums all the elements of the resulting matrix to get a single value. Each of these resulting values are the outputs in the convoluted feature matrix. During training, the CNN 'learns' the optimal values for the filter matrices that enable it to extract meaningful features from the input feature map. As the number of filters applied to the input increases, so does the number of features the CNN can extract. However, the trade-off is that filters compose the majority of resources expended by the CNN, so training time also increases as more filters are added.

Following each convolution operation, the CNN applies a Rectified Linear Unit (ReLU) transformation to the convoluted feature, in order to introduce nonlinearity into the model. The equation for ReLU can be written as follows:

$$\mathbf{f(x) = max(0, x)} \tag{1}$$

where x is the input to the neuron. ReLU can allow the model to account for non-linearities and interactions.

Neurons are only locally connected by filters, followed by a pooling layer of fixed size 2x2. For this model we have used max-pooling. This max-pooling layer down-samples the convoluted feature, reducing the number of dimensions of the feature map, while still preserving the most critical feature information and to control over fitting. In a fully connected layer all the neurons consider every activation in the previous layer. Each layer learns its weights and biases using gradient descent in small mini-batches of training samples. In this model, the batch size to train is 32 and the number of output classes is 8. Learning rate is set to 0.001 and the maximum number of epochs is 100.



(a) Normal Scenario

(b) Transmitter Off

(c) Site Outage

Fig. 3: Comparision of different faults generated in the network

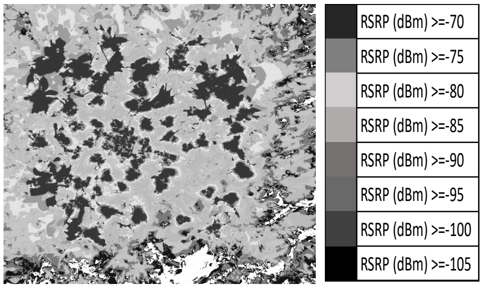


Fig. 4: Grayscale RSRP map used as input to CNN model

Since the neurons in a fully connected layer have connections to all activations in the previous layer, their activations can be computed with a matrix multiplication followed by a bias offset. In the convolutional layer the neurons are connected only to a local region in the input, and only those many the neurons in a convolutional volume share parameters. However, the neurons in both layers still compute dot products, so their functional form is identical. The last fully-connected layer holds the output.

B. Naive Bayes Classifier Model

NB classifiers are a collection of classification algorithms with strong (naive) independence assumptions between the features in which every pair of feature being classified is independent of each other. NB is particularly useful for large data sets. To input the data-set consisting of images, a matrix in which each row is a 1-D array associated to an image is created. Next, an array which contains the associated label for each class, i.e all the fault labels used in this paper as described above was added to the matrix. The NB classifier combines this model with a function which maps an observation to an appropriate action. This classifier is the function that allocates the class label. The classifier used in this paper is Gaussian NB classifier and can be expressed as:

$$P(x_i \mid y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right) \tag{2}$$

where x_i denotes the feature vector and y is the class variable. σ_y and μ_y are estimated using maximum likelihood. The class with the highest probability is considered as the most likely class. This is also known as Maximum A Posteriori (MAP). The MAP for a hypothesis is:

$$\begin{aligned} MAP(H) &= max(P(H|E)) && (3)\\ &= max((P(E|H) * P(H))/P(E)) && (4)\\ &= max(P(E|H) * P(H)) && (5) \end{aligned}$$

P(E) is evidence probability, and it normalizes the result.

The advantage of using NB is that it performs well in multiclass prediction and is fast to predict class of test data set. In cases where assumption of independence is valid, a NB classifier performs better compared to logistic regression and less training data is needed. But the disadvantage of using NB is that if categorical variables have a category in test data set, which were not observed in training data set, then the model will assign a 0 (zero) probability and will be unable to make a prediction. This is often known as "Zero Frequency".

C. Random Forests Classification

A third model in the proposed framework is the RF classifier. A Decision Tree (DT) builds a set of rules that predict a class for classifying complex situations. DT when used as ensemble can even outperform neural networks in some cases [17]. The most popular ensemble technique is RF. A RF is a collection of DT whose results are aggregated into one final result. RFs limit over fitting without substantially increasing error due to bias by training on different samples of the data. Boostrap resampling method aids in extracting multiple samples from the original samples and to construct sub-datasets. The sub dataset are then used to form the base decision tree and train on it. This method is a statistical technique for estimating quantities about a population by averaging estimates from multiple small data samples. A second method is by using a random subset of features. This means, for a set of 8 features, RF will only train on a certain number of features in each model. Thus, in each tree a minimum number of random features can be utilized. In a forest, eventually most or all of the features would have been included making them much more robust than a single decision tree. RF also offers a good feature selection indicator. RF use Gini importance or mean decrease in impurity (MDI) to calculate the importance of each feature. The Gini coefficient (GC) measures the inequality among values of a frequency distribution where zero expresses perfect equality and 1 or 100% expresses maximal inequality among values. The RF classifier's parameters that have to be set, consist of number of trees, number of features, impurity function and stop criteria. Generally, the default values have been used for this. The maximum number of iteration chosen is 100. The depth of tree is limited to 5 levels. The overall accuracy and Kappa Coefficient obtained by RGB images are 77.82% and 0.63, respectively. For the grayscale images, these values are 87.16%; and 0.79, respectively.

D. Neuromorphic Computing model

The goal of this approach is to assess the similarity of the decision making mechanisms of convolutional neural networks and corresponding mechanisms in the human cortex. This tool greatly helps in understanding the dynamic method of learning and development in the brain. This can be an inspiration which can be applied on generic cognitive computing. In late 1980, this concept was developed by Carver Mead [18], describing the use of very-large-scale integration (VLSI) systems containing electronic analog circuits to mimic neurobiological architectures present in the nervous system. We may expect neuromorphic technologies to deliver a range of applications more efficiently than conventional computers. For instance, to deliver speech and image recognition capabilities in smart phones. Currently, such capabilities are available only using powerful cloud resources to implement the recognition algorithms.

1) Nengo: Nengo [19] is a neural simulator based on a framework called the Neural Engineering Framework (NEF). The Neural Engineering Framework (NEF) is a set of theoretical methods that are used in Nengo for constructing neural models. The NEF is based on Eliasmith and Anderson's (2003) book from MIT Press. Nengo is highly extensible and flexible. One can define their own neuron types and learning rules, get input directly from hardware, build and run deep neural networks. It is a large-scale modeling approach that can leverage single neuron models to build neural networks with demonstrable cognitive abilities.

2) Nengo Neural Network: Synapses are unidirectional connections that neurons use to communicate. When a neuron spikes, the neurotransmitter released across the synapse, causes some quantity of current to be transmitted in the postsynaptic (downstream) neuron. Factors affecting the amplitude of the current can be summarized in a scalar connection weight. To compute any function, the weights of the connection are set as the product of the decoding weights for that function in the first population, the encoding weights for the downstream population, and any linear transform. Non-neural information (sensory inputs and motor outputs) are represented by the Node. A complex experimental setup that provides input and responds to the output can be modeled. The Connection describes how ensembles and nodes are connected. The Probe gathers data during a simulation for later analysis. The Network encapsulates a functionally related group of interconnected nodes and ensembles. The Model encapsulates a Nengo model. Several signals are created for each high-level Nengo object; i.e. the simulator creates signals that represent the high-level input signal that will be encoded to input currents, and the encoding weights. The ensemble also contains a neural population, for which the simulator creates signals that represent input currents, bias currents, membrane voltages, and refractory times for each cell.

IV. Performance analysis

All the models explained earlier perform better for grayscale images than RGB images. As seen in Fig 5, a visible difference in accuracy can be noted. This difference in accuracy can be attributed to the increase in the number of kernels. The kernel size for grayscale image is k x k x 1 where as it is k x k x 3 for an RGB image. Depending on the number of kernels, the number of parameters increases proportionally. Furthermore, large dimensional vectors are almost equally spaced and hence are difficult to separate by a classifier. This can be accredited to the curse of dimensionality phenomena. Although no substantial information is added in an RGB image, it in turn adds to more compute and memory intensive training time due to large number of input parameters. Hence, feeding these model with a grayscale image rather than an RGB image results in a better classification accuracy.

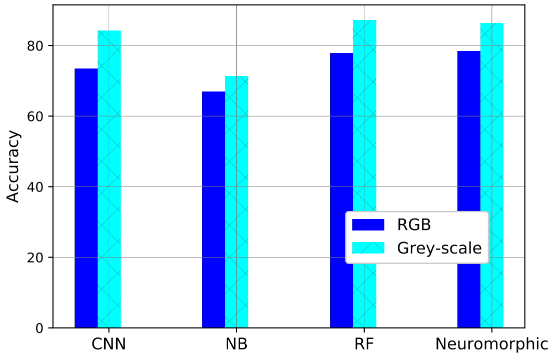


Fig. 5: Graph of Accuracy for grayscale and RGB images

Advantages of neuromorphic computing as compared to the other approaches are execution speed, adaptability, energy efficiency, the ability to learn, robustness against local failures and diverse cell types at individual nodes. There are primarily two ways in which the scene perception or image recognition is performed in the human brain:

1) an object-centered approach, in which components of a scene are segmented and serve as scene descriptors;
2) space-centered approach, in which spatial layout and global properties of the whole image or place act as the scene descriptors.

The RF model has the highest classification accuracy followed by the neuromorphic model, CNN and NB. Some classifier combination techniques like ensembling, bagging and boosting may improve NB classification accuracy but these methods would not help since their purpose is to reduce variance. NB has no variance to minimize. Accuracy of CNN saturates after a particular number of epoch as the model has limited number of data samples to train from. Neural Networks show best results when the number of data points or images are very large in number. Loss values of the CNN model on the test dataset are represented in Fig 6. Classification accuracy of individual fault parameters obtained from RF classifier are represented in Fig 7. Classification accuracy for S.O. is the highest for RGB images while C.O. are better classified in grayscale images. Also, a very small degree of antenna tilt results in a noticeable change in the RSRP map images therefore enabling RF to better classify them as compared to TxP faults.

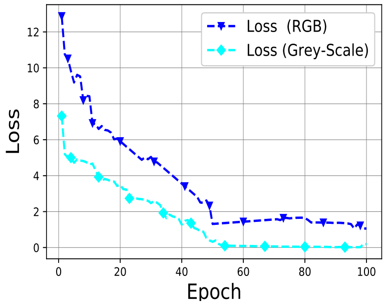


Fig. 6: CNN model loss values for grayscale and RGB images

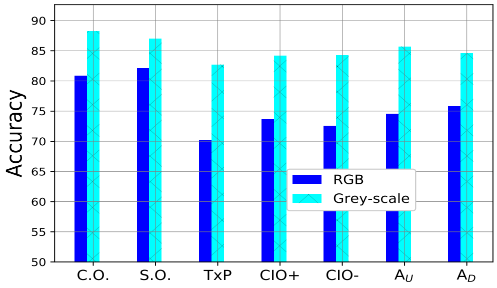


Fig. 7: RF prediction accuracy for individual faults

V. Conclusion and Future Work

Fault diagnosis is a highly desirable feature in emerging complex and dense cellular networks to not only reduce operational costs but also to improve quality of experience. In this paper, we present a first of its kind neuromorphic based AI model fault diagnosis solution that outperforms the prevailing, CNN, RF and NB fault diagnosis schemes in existing literature. The key advantage of proposed solution is that, it does not require human experts for manual feature extraction unlike existing solutions in literature. This advantage is achieved by feeding RSRP data in form of images to train the models instead of conventional approach of using raw fault data. In real networks, these images can be created using MDT reports or RSRP measurements gathered from other sources. The results indicate that the proposed model can flag all seven faults in the fed images. This framework can be used as a part of the self-healing module in emerging networks. For future works, we plan to investigate the performance of the proposed solution in the presence of multiple faults concurrently as well as fault detection.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant Numbers 1559483, 1619346 and 1730650. For more details, please visit 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, thirdquarter 2018.

[2] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: how to empower SON with big data for enabling 5G," *IEEE Network*, vol. 28, no. 6, pp. 27–33, Nov 2014.

[3] 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)*. IEEE, 2018, pp. 206–212.

[4] R. Barco, L. Nielsen, R. Guerrero, G. Hylander, and S. Patel, "Automated troubleshooting of a mobile communication network using Bayesian networks," in *4th International Workshop on Mobile and Wireless Communications Network*, Sep. 2002, pp. 606–610.

[5] E. J. Khatib, R. Barco, A. Gómez-Andrades, P. Muñoz, and I. Serrano, "Data mining for fuzzy diagnosis systems in LTE networks," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7549–7559, Nov. 2015. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2015.05.031

[6] R. M. Khanafer, B. Solana, J. Triola, R. Barco, L. Moltsen, Z. Altman, and P. Lazaro, "Automated diagnosis for UMTS networks using Bayesian network approach," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2451–2461, July 2008.

[7] 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. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.1060

[8] M. Z. Asghar, S. Hämäläinen, and T. Ristaniemi, "Self-healing framework for LTE networks," in *2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sep. 2012, pp. 159–161.

[9] R. Barco, V. Wille, L. Díez, and M. Toril, "Learning of model parameters for fault diagnosis in wireless networks," *Wirel. Netw.*, vol. 16, no. 1, pp. 255–271, Jan. 2010. [Online]. Available: http://dx.doi.org/10.1007/s11276-008-0128-z

[10] S. Rezaei, H. Radmanesh, P. Alavizadeh, H. Nikoofar, and F. Lahouti, "Automatic fault detection and diagnosis in cellular networks using operations support systems data," *NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium*, pp. 468–473, 2016.

[11] P. Szilagyi and S. Novaczki, "An automatic detection and diagnosis framework for mobile communication systems," *IEEE Transactions on Network and Service Management*, vol. 9, no. 2, pp. 184–197, June 2012.

[12] P. Szilagyi and S. Novaczki, "An automatic detection and diagnosis framework for mobile communication systems," *IEEE Transactions on Network and Service Management*, vol. 9, no. 2, pp. 184–197, June 2012.

[13] S. Nováczki, "An improved anomaly detection and diagnosis framework for mobile network operators," in *2013 9th International Conference on the Design of Reliable Communication Networks (DRCN)*, March 2013, pp. 234–241.

[14] A. Hassan and A. Mahmood, "Convolutional Recurrent Deep Learning model for sentence classification," *IEEE Access*, vol. 6, pp. 13 949–13 957, 2018.

[15] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep face recognition: A survey," in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, Oct 2018, pp. 471–478.

[16] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German traffic sign recognition benchmark: A multi-class classification competition," in *The 2011 International Joint Conference on Neural Networks*, July 2011, pp. 1453–1460.

[17] B. Abraham, A. Mandya, R. Bapat, F. Alali, D. E. Brown, and M. Veeraraghavan, "A comparison of machine learning approaches to detect Botnet traffic," in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–8.

[18] "Carver Mead. Available at: http://www.carvermead.caltech.edu."

[19] "Nengo. Available at: https://www.nengo.ai/."