Machine Learning-based Modulation Format Identification and Optical Performance Monitoring Techniques Implementation

by Puspa Devi

Submission date: 29-Sep-2021 11:00AM (UTC+0700)
Submission ID: 1660374567
File name: ng-based_Modulation_Format...Implementation_Final_Paper_2.0.pdf (526.75K)
Word count: 4447
Character count: 23595

Machine Learning-based Modulation Format Identification and Optical Performance Monitoring Techniques Implementation

Pukhrambam Puspa Devi
Department of Electronics and Communication Engineering National Institute of Technology Silchar
Silchar 788010, India
puspa.devi@ece.nits.ac.id

Vincent
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung 40132, Indonesia
23221049@std.stei.itb.ac.id

Joni Welman Simatupang
Study Program of Electrical Engineering
President University
Bekasi 17530, Indonesia
joniwsmtp@president.ac.id

***Abstract*—In this paper, machine learning-based techniques are used to solve and analyze the modulation format recognition problem. The combination of intelligent software and high performance hardware provides a large scope for innovation in optical networking. Machine learning algorithms can use a large amount of data available from the network monitors to learn and make the network more robust. This is a problem in optical communication that consists of defining the type of digital modulation process in which an electrical signal should be sent. A dataset to represent realistic transmission behaviors was generated using a simulator based on a Gaussian noise model. A multi-layer perceptron was used and tested with different architectures to show that a high level of accuracy is achievable with machine learning. An analysis of the input features was made by using the select *K* best features method. Finally, an attempt to visualize the data in 2-dimension was made using the Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) methods to reduce the dimensionality of the input features and see their relationships.**

Keywords—*Machine Learning, Modulation Format Recognition, Optical Network, PCA, t-SNE*

I. INTRODUCTION

Optical networks establish the essential actual framework of all large-provider networks around the world. On account of the high limit, minimal effort, and numerous different properties, there is no sign that a substitute innovation may show up withing a reasonable time frame. It constitutes the physical infrastructure of all major network service providers worldwide, and is able to provide high data transmission speed while having a low cost, as well as other properties. Judging from the perspective of internet traffic, it is expected that the number of services supported by optical networks would significantly increase in the upcoming years. Thus, the network should be adaptable to it by using the data from previous measurements [1]–[4].

Modern optical transmitters and receivers provide great flexibility in modulation format, carrier frequency, and utilized bandwidth. Given that an arbitrary modulation format can be adopted at the transmitter side, the information regarding the modulation format may not always be available at the receiver side. It may affect signal detection, signal demodulation, as well as signal processing [5]–[7].

Given the condition where the detection may not be straightforward, it can be approached by using machine learning techniques. Even so, the use of machine learning techniques with different neural network training in optical communication is still in its infancy. It is considered a paradigm shift for designing optical networks and the systems in the future. Confidently, it is possible to identify the most important features that affect the accuracy and efficiency of identifying the modulation format at the receiver side by having rigorous training and analyzing the data using machine learning [8]–[12].

In this paper, we propose machine learning-based techniques to identify and analyze the modulation format. Both tasks are done at the same time to get a better understanding of it. It will further help in optical signal monitoring and future elastic optical networks [13].

The rest of this paper is organized as follow. Chapter 2 surveys the literatures that conducted research on similar topic. Then Chapter 3 elaborates the methodology used for this work. Next, Chapter 4 discusses the obtained results. Finally, Chapter 5 concludes this paper.

II. LITERATURE SURVEY

In adaptive optical communication systems, users' demand and channel's condition determine the rate of data transmission channel. Since the modulation format may be changed at the transmitter side, the receiver requires the information in order to identify different modulation formats. Hence coherent receivers with Modulation Format Identification (MFI) can be used to address this problem [14]–[17].

Several works on MFI that had been conducted by other researchers are summarized in

Table 1.

III. METHODOLOGY

In this paper, we focus on MFI with neural network-based implementation, optimization, feature selection, and data visualization. The strategies include the optimization of the model to achieve more realistic predictions based on input features. The work plan for this paper is summarized in Fig. 1.

TABLE 1. SUMMARY OF WORKS PRESENTED IN THIS FIELD

Authors	Objective	Method	Advantage	Disadvantage
Khan et al. (2016) [18]	MFI in coherent receivers using Deep Machine Learning	Uses amplitude histograms obtained with the help of constellation diagrams and employs ANN and DNN for MFI	High accuracy	It could not recognize the high order PSK formats due to the similitude of their amplitude histograms
Bilal et al. (2015) [19]	Blind MFI in coherent receivers using Deep Machine Learning	Peak-to-Average Power Ratio (PAPR) of the received data samples	High accuracy	It requires previous knowledge as well as huge values of the Optical Signal to Noise Ratio (OSNR)
Adles et al. (2014) [20]	Blind MFI from physical layer characteristics	Electric field distributions and histograms	Excellent classification rates	Computation process is quite complex
Liu et al. (2014) [21]	MFI based on power distributions of received signals for digital coherent receivers	Proposes a way to use the distrbution properties of received signals. The MFI used power distribution	Perform well identification	Adjustment of various thresholds are requires
Bo et al. (2016) [22]	Blind MFI software-defined optical network using techniques of image processing	Used Voronoi diagrams and obtained the binary images for MFI	High accuracy	Requires huge amount of samples for 16-QAM identification

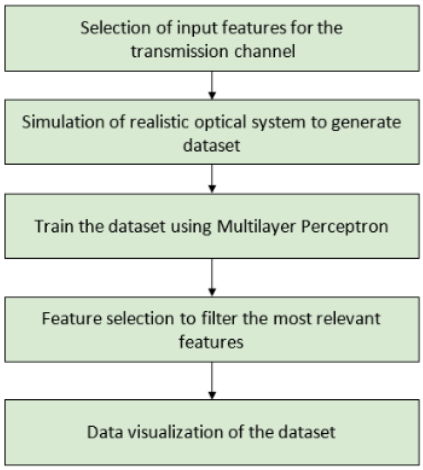


Fig. 1. Work plan for the paper

Since the analysis requires dataset, and only big network providers have the access to it, this work utilizes a simulation model an optical transmission channel – including its non-linearities– to generate the dataset. It represents a realistic transmission behaviour based on Gaussian Noise (GN) model.

With the completion of input feature selection and simulation of the non-linear optical transmission channel, a dataset of 100,000 samples is generated. This dataset is then be fed into the machine learning technique to develop the model. Since the model may have a lot of features put into high dimension, and high dimensionality creates overfitting model that leads to the decrement of capability to generalize apart from the entries of the training set due, thus the number of features is reduced. Besides, measurement of some of the features can be costly for an organization, while some others are difficult to be measured. Thus, having a smaller number of input features while having an excellent classification score will be highly desirable. In this work, Principal Component Analysis (PCA) is employed. It recognizes a rundown of the principal axes to portray the fundamental dataset prior to positioning them as per the measure of variance caught by each.

There are 11 input features used in this paper, and they are shown in Table 2. The first one is symbol rate, which defines the number of changes in the symbol occurring every second. One symbol may carry multiple bits. This parameter is denoted in baud. The second one is roll-off, which defines the bandwidth occupied beyond Nyquist bandwidth. This parameter varies between 0 to 1. Next is the launch power, which determines the amount of distortion. Besides, in the channel transmission, data is monitored throughout the network while load balancing is done to maintain the bearable load on each channel. After that, there is also a property by which a signal spreads in time, known as dispersion. Then, as light travels through a medium, nonlinear effects are present. Its index is calculated for linearly polarized light. During transmission, noise is normally non-neglectable. A noise figure is a measure of the degradation in signal to noise ratio, and can be used in association with receiver sensitivity. At the same time, signal power lost during transmission is known as loss, which is normally denoted in dB. Last, the channel grid and span are also used.

Most of the features are independent, although some of them are not. For instance, launch power and channel grid are dependent on symbol rate. The dataset is created using Python, and the output is in the form of Comma-separated Values (CSV) format.

TABLE 2. INPUT FEATURES USED IN THIS PAPER

Feature	Value
Symbol Rate	30 – 90 Gbaud/s
Roll-off	0.01 – 0.05
Launch Power	f (symbol rate)
Channel Load	1 – 120

Dispersion	4 – 21 ps/nm/km
Non-linear Index	0.8 – 1.6 l/W.km
Loss	0.15 – 0.2 dB/km
Span Count	1 – 50
Span Length	40 – 120 km
Noise Figure	4 – 6.5 dB
Channel Grid	f (symbol rate)

In generating the dataset, the transmission is simulated along with GN. It is desirable to have lowest possible loss to improve the OSNR, as well as a large area to avoid nonlinear distortion [23], [24]. However, it was practically impossible to have one due to huge accumulated dispersion. As the technology has advanced, it is now possible to have optical dispersion compensation in the transmission links. For instance, a 5500 km channel operates at 40 Gbaud/s having interference of approximately 800 baud. Because of the overlap, optical field in time-domain randomizes, resulting in a statistically independent zero-mean Gaussian. The amplitude of the optical field that becomes random along with the intra-channel and inter-channel interactions that are nonlinear. Thus, the non-linear interference can be interpreted as random noise by the receiver. Statistical distribution of the non-linear interference noise (NLI) is statistically independent and zero-mean Gaussian distributed. Thus, for low to moderate NLI, the effect can be assumed to be GN.

In the simulation, the transmitter side considered ordinary highlights such as span count, span length, and symbol rate on an optical connection. It also employs several bit-modulation techniques, and the OSNR penalties have been taken as 0.5 dB for Dual-Polarized BPSK (DP-BPSK), 1 dB for DP-QPSK, 2 dB for DP-16QAM, and 3 dB for DP-64QAM. The optical transmission channel is modelled with GN approach.

After having the dataset, it is partitioned into three parts. The first one is validation data, which compromises of 20,000 entries. Its presence is able to prevent overfitting, thus improving the model's accuracy. Next, the training data has 60,000 entries. Finally, the test dataset is compromised with 20,000 entries. The dataset is fed into Multilayer Perceptron (MLP) as a feedforward model with supervised learning. The training is done via backpropagation algorithm.

Then, the feature is analysed using a feature selection method. The aim is to check if is it possible to decrease the number of features used in the modulation classification without a loss in terms of accuracy. In this work, select *K* best features method is utilized. It is a forward sequential selection method, which takes *K* features that achieve the highest scores according to individual performance measurement.

Finally, the result is visualized into a two-dimensional space. There are two algorithms utilized in this work. The first one is Principal Component Analysis (PCA). However, it only captures the one that has linear relationship [25]. The second one is t-Distributed Stochastic Neighbour Embedding (t-SNE). It is able to improve the data-visualization as it captures components in non-linear relationship.

IV. Result and Discussion

This chapter begins with brief overview on the generated dataset. The discussion then continues to the MLP section, which covers the classification accuracy, training time, as well as the confusion matrices. After that, the performance of each feature is evaluated based on its accuracy. Finally, the data visualization with PCA and t-SNE is elaborated by the end of this chapter.

The generated dataset has the eleven features included to form 100,000 samples in CSV file. The data is utilized to serve as dataset for MLP models. The snapshot of the dataset is given in Fig. 2.

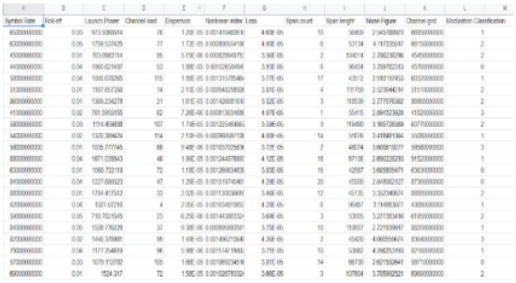


Fig. 2. Snapshot of the dataset

Next, the MLP is designed to have four architectures. The first one has one hidden layer with five neurons. Unless stated otherwise, this architecture is written as [1:(5)] for the rest of the paper. The second one has 1 hidden layer with ten neurons, and is written as [1:(10)]. Then the third one has one hidden layer with a hundred neurons, and is written as [1:(100)]. Finally, the last one has two hidden layers with a hundred neurons and ten neurons each [2:(100x10)]. The layers use Rectified Linear Unit (ReLU) activation function, and the output layer uses soft-max activation function. One batch has a size of a hundred samples, and has learning rate of 0.01. the Random Uniform function, which generates weights with a uniform distribution, is used to initialize the weights between -0.05 to 0.05.

The models are then run with the dataset that consists of training, validation, and testing data. Fig. 3 shows the classification accuracy on validation data on the base of the number of epochs run. In this work, the maximum epoch is set to be 500. It is observed that all architectures follow a similar saturation behavior after reaching a certain number of epochs. The model with architecture of [2:(100x10)] reaches convergence earlier than the other three models. It is followed with the one with architecture of [1:(100)].

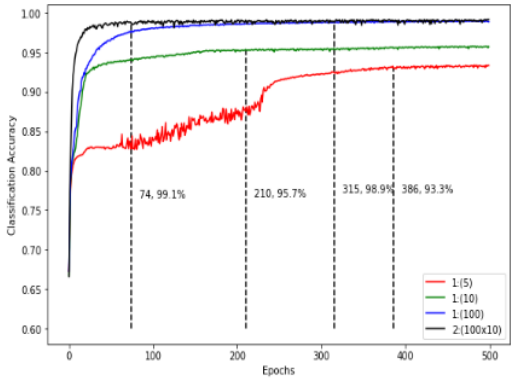


Fig. 3. Accuracy of classification per epochs for the four models

Next, the performance is evaluated by taking the training time into account. Fig. 4 shows the training time required for each model to reach convergence. All of the models do not have significant difference. The longest time of 1200 s is required for the model of [2:(100x10)], while the shortest one of 1000 s is required for the model of [1:(10)].

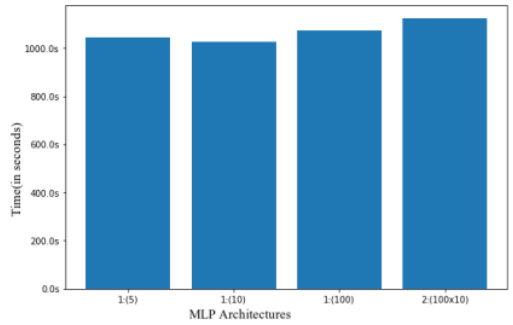


Fig. 4. Model training time to reach convergence

After that, the evaluation takes the confusion matrices of each architecture into account. It depicts the number of correct classification as well as the wrong one, including where it goes wrong. Overall, the accuracy ranges between 93.59% to 98.55%.

Fig. 5 shows the confusion matrix for the first architecture. The classification is confused mostly in the DP-BPSK and DP-QPSK scenario, with accuracy of 83.63% and 90.34%, respectively. Even though DP-16QAM and DP-64AM also has error, it is significantly less than PSK scenario with the accuracy of 99.89% and 99.78%, respectively. Overall, this architecture achieves accuracy of 93.59% throughout all scenarios.

Fig. 6 shows the confusion matrix of the second architecture. The result is similar to the previous architecture, where PSK-type modulation has lower accuracy compared to QAM-type. In this case, the model has accuracy of 90.60% and 95.23% for DP-BPSK and DP-QPSK, respectively. On the other hand, DP-16QAM and DP-64QAM are able to be classified perfectly with accuracy of 100%. Overall, this architecture achieves accuracy of 96.64% throughout all scenarios.

Fig. 7 shows the confusion matrix of the third architecture. It is able to achieve higher accuracy compared to the previous one, where the classification of DP-BPSK and DP-QPSK has accuracy of 94.64% and 98.58%, respectively. Similar to the second architecture, DP-16QAM and DP-64QAM are able to be classified perfectly with the accuracy of 100%. Overall, this architecture achieves accuracy of 98.55% throughout all scenarios.

Fig. 8 shows the confusion matrix of the fourth architecture. It is able to achieve higher accuracy compared to the previous one, but only for DP-BPSK. The accuracy for this bit-modulation is 97.97%. for DP-QPSK, the accuracy is less than the third architecture with the number of 96.62%. As for the DP-16QAM and DP-64QAM, the accuracy reaches 100%, similar to the second and the third architecture. Overall, this architecture achieves accuracy of 98.42% throughout all scenarios, which is 0.13% lower than the third architecture.

Based on the evaluations, the fourth architecture is the best option to choose. Even though it requires longest training time, and has slightly lower classification accuracy compared to the third architecture, it converges fastest among the other architectures.

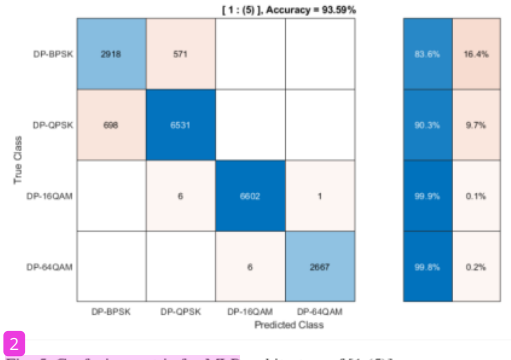


Fig. 5. Confusion matrix for MLP architecture of [1:(5)]

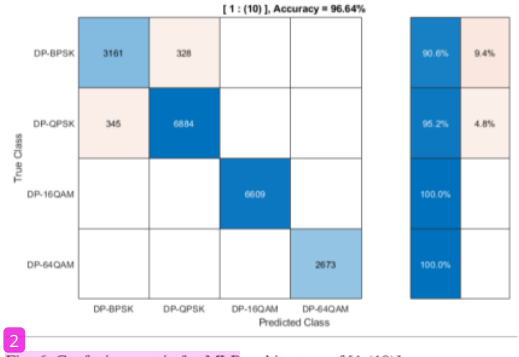


Fig. 6. Confusion matrix for MLP architecture of [1:(10)]

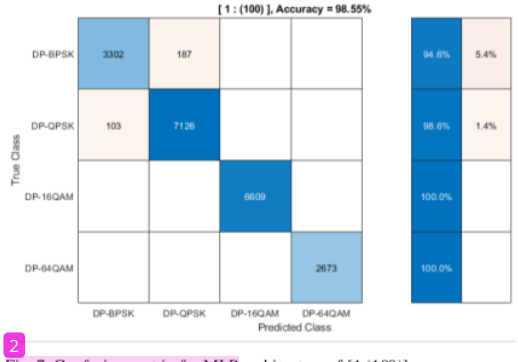


Fig. 7. Confusion matrix for MLP architecture of [1:(100)]

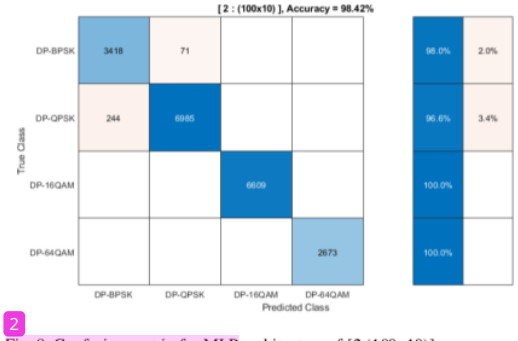


Fig. 8. Confusion matrix for MLP architecture of [2:(100x10)]

Then, the features are evaluated based on its accuracy. The measurement is performed by using χ^2 test. Each feature is fitted into the fourth architecture. The accuracy level and where the convergence takes place are observed for each feature.

Table 3 shows the evaluation for each feature. It is sorted in descending order based on the χ^2. The first three features score below 70% on the accuracy, and only the last four features score higher than 90%. According to the data, it is possible to achieve high accuracy while saving resources at the same time by using just four features.

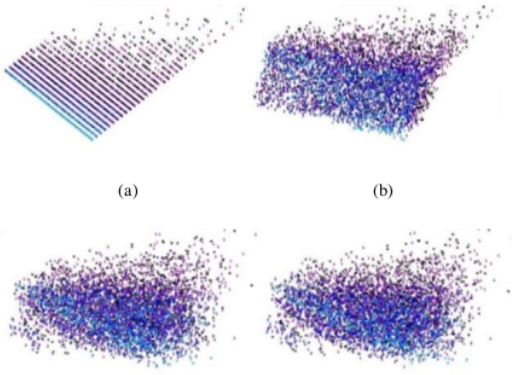
TABLE 3. MLP ACCURACY USING K BEST FEATURES

K	Feature	χ^2	MLP Accuracy	Epoch
1	Span Count	4625.19	61.76%	49
2	Symbol Rate	26.10	65.30%	54
3	Channel Grid	24.56	65.57%	59
4	Span Length	17.09	78.32%	81
5	Launch Power	13.79	81.55%	114
6	Noise Figure	5.78	86.86%	134
7	Dispersion	4.86	87.56%	127
8	Non-linear Index	2.74	92.20%	100
9	Loss	1.58	94.82%	106
10	Roll-off	0.55	95.72%	62
11	Channel Load	0.23	98.68%	98

Finally, the dimensionality is reduced and presented within two-dimensional space. The first one is reduced by using PCA algorithm, and the result is depicted in Fig. 9. The data has four parts, where each part has different number of features. The first one has 2 features, the second one has 4 features, the third one has 8 features, while the last one has all 11 features in. The DP-BPSK is presented in black, DP-QPSK is presented in purple, DP-16QAM is presented in blue, and DP-64QAM is presented in cyan. The pattern behind the data is hard to be interpreted. Since PCA captures data with linear relationship only, then it implies that the features do not have linear relationship.

Then Fig. 10 visualizes the data with t-SNE algorithm for 2, 4, 8, and 11 features. It provides a more interpretable data compared to PCA algorithm. The data shows that DP-16QAM and DP-64QAM can be differentiated relative easily. On the other hand, DP-BPSK and DP-QPSK are indistinct in most of the areas, making it hard to classify. Even though there is no definite cluster, there exist different areas being formed based on three divisions. The first one is for DP-64QAM, the second one is DP-16QAM, and the last one is the mixture between DP-BPSK and DP-QPSK.

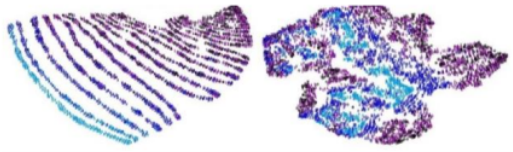
Also, using only 2 features or 4 make the data to have a direction that modifies the classification. In this case, blurriness occurs near the boundaries of the decision regions. By adding more features, e.g., by using 8 or 11 features, there are different directions to improve the classification. Then the area of DP-BPSK and DP-QPSK ends, it is normally leads to DP-16QAM, then followed by DP-64QAM. Using those features, the aforementioned more-specific cases can be separated with higher accuracy.



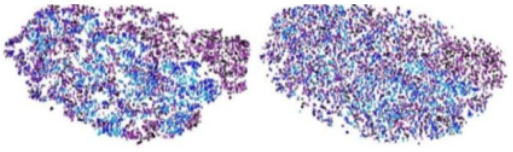
(a) (b)

(c) (d)

Fig. 9. Data visualization using PCA for (a) 2 features, (b) 4 features, (c) 8 features, and (d) 11 features.



(a) (b)



(c) (d)

Fig. 10. Data visualization using t-SNE for (a) 2 features, (b) 4 features, (c) 8 features, and (d) 11 features

V. CONCLUSION

Through this work, a machine learning-based techniques as an approach to solve modulation format prediction. Due to the lack of industrial data of optical transmission channels, a simulator is designed to mimic a non-linear optical transmission channel using 11 input features with industrial input ranges. Using the dataset generated using this simulator, a machine learning-based model is developed. The performance of various layouts of neural networks have been measured, as well as the attempt on achieving the mathematical performance of optimal decisions. The work also covers feature selection to reduce the number of input features. Confidently, this would help organizations in cutting the cost of measurement instruments. Finally, the data is visualized in different forms to help establish a relationship among the features.

The MLP model allows the modulation identification from a dataset generated using a simulator with high accuracy. However, there have been certain areas in this work which will require attention in future and can emerge great scope of development from the perspective of modulation format identification and optical performance monitoring. Other emerging technologies in the field of ML like Convolutional Neural Network (CNN) can be used to keep or improve the accuracy performance using less features. Besides, the industrial measurements from real optical transmission channels would provide better quality dataset to train the model. The data, –if and when available– can be used for the development. Finally, another possible work would be to convert this problem as regression. Instead of directly classifying the best method to send the electrical signal, the aim could be to predict the OSNR value.

REFERENCES

[1] R. Casellas, R. Martínez, R. Vilalta, and R. Muñoz, "Control, Management, and Orchestration of Optical Networks: Evolution, Trends, and Challenges," *J. Light. Technol.*, vol. 36, no. 7, pp. 274–282, 2018, doi: 10.1109/JLT.2018.2793464.

[2] J. Yu and J. Zhang, "Recent Progress on High-speed Optical Transmission," *Digital Communications and Networks*, vol. 2, no. 2. pp. 65–76, 2016, doi: 10.1016/j.dcan.2016.03.002.

[3] M. Channegowda, R. Nejabati, and D. Simeonidou, "Software-defined Optical Networks Technology and Infrastructure: Enabling Software-defined Optical Network Operations [Invited]," *J. Opt. Commun. Netw.*, vol. 5, no. 10, pp. 274–282, 2013, doi: 10.1364/JOCN.5.00A274.

[4] P. Samadi, D. Amar, C. Lepers, M. Lourdiane, and K. Bergman, "Quality of Transmission Prediction with Machine Learning for Dynamic Operation of Optical WDM Networks," 2017, doi: https://doi.org/10.1109/ECOC.2017.8346216.

[5] V. O. C. Dias, E. da V. Pereira, H. R. O. Rocha, M. E. V. Segatto, and J. A. L. Silva, "Performance Evaluation of CO-OFDM Systems based on Electrical Constant-envelope Signals," *Opt. Fiber Technol.*, vol. 37, pp. 30–34, 2017, doi: 10.1016/j.yofte.2017.06.010.

[6] H. Ren, J. Yu, Z. Wang, J. Chen, and C. Yu, "Modulation Format Recognition in Visible Light Communications based on Higher Order Statistics," in *2017 Conference on Lasers and Electro-Optics Pacific Rim, CLEO-PR 2017*, 2017, pp. 1–2, doi: 10.1109/CLEOPR.2017.8119030.

[7] A. Nag, M. Tornatore, and B. Mukherjee, "Optical Network Design with Mixed Line Rates and Multiple Modulation Formats," *J. Light. Technol.*, vol. 28, no. 4, pp. 466–475, 2010, doi: 10.1109/JLT.2009.2034396.

[8] T. Gui, C. Lu, A. P. T. Lau, and P. K. A. Wai, "Opportunities for Machine Learning in Optical Communication Systems and Networks," 2017, doi: http://dx.doi.org/10.1364/ACPC.2017.Su4B.1.

[9] W. S. Saif, M. A. Esmail, A. M. Ragheb, T. A. Alshawi, and S. A. Alshebeili, "Machine Learning Techniques for Optical Performance Monitoring and Modulation Format Identification: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 22, no. 4, pp. 2839–2882, 2020, doi: 10.1109/COMST.2020.3018494.

[10] M. A. Esmail, W. S. Saif, A. M. Ragheb, and S. A. Alshebeili, "Free Space Optic Channel Monitoring Using Machine Learning," *Opt. Express*, vol. 29, no. 7, p. 10967, 2021, doi: 10.1364/oe.416777.

[11] D. Rafique and L. Velasco, "Machine Learning for Network Automation: Overview, Architecture, and Applications [Invited Tutorial]," *J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. 126–143, 2018, doi: 10.1364/JOCN.10.00D126.

[12] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," *J. Light. Technol.*, vol. 37, no. 2, pp. 493–516, 2019, doi: 10.1109/JLT.2019.2897313.

[13] S. Shahkarami, F. Musumeci, F. Cugini, and M. Tornatore, "Machine-learning-based Soft-failure Detection and Identification in Optical Networks," in *Optics InfoBase Conference Papers*, 2018, vol. Part F84-O, doi: 10.1364/OFC.2018.M3A.5.

[14] L. Guesmi and M. Menif, "Modulation Formats Recognition Technique using Artificial Neural Networks for Radio Over Fiber Systems," in *International Conference on Transparent Optical Networks*, 2015, vol. 2015-Augus, pp. 1–4, doi: 10.1109/ICTON.2015.7193508.

[15] P. Poggiolini and Y. Jiang, "Recent Advances in the Modeling of the Impact of Nonlinear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems," *J. Light. Technol.*, vol. 35, no. 3, pp. 458–480, 2017, doi: 10.1109/JLT.2016.2613893.

[16] J.-X. Cai, G. Mohs, and N. S. Bergano, "Ultralong-distance Undersea Transmission Systems," in *Optical Fiber Telecommunications VII*, 1st ed., A. Willner, Ed. 2019.

[17] D. Rafique, "Machine Learning Based Optimal Modulation Format Prediction for Physical Layer Network Planning," in *20th International Conference on Transparent Optical Networks*, 2018, pp. 1–4, doi: https://doi.org/10.1109/ICTON.2018.8473593.

[18] F. N. Khan, K. Zhong, W. H. Al-Arashi, C. Yu, C. Lu, and A. P. T. Lau, "Modulation format Identification in Coherent Receivers Using Deep Machine Learning," *IEEE Photonic Technol. Lett.*, vol. 28, no. 17, pp. 1886–1889, 2016, doi: https://dx.doi.org/10.1109/LPT.2016.2574800.

[19] S. M. Bilal, G. Bosco, Z. Dong, A. P. T. Lau, and C. Lu, "Blind Modulation Format Identification for Digital Coherent Receivers," *Opt. Express*, vol. 23, no. 20, pp. 26769–26778, 2015, doi: https://dx.doi.org/10.1109/JLT.2014.2307555.

[20] E. J. Adles *et al.*, "Blind Optical Modulation Format Identification from Physical Layer Characteristics," *J. Light. Technol.*, vol. 32, no. 8, pp. 1501–1509, 2014, doi: https://dx.doi.org/10.1109/JLT.2014.2307555.

[21] J. Liu, Z. Dong, K. Zhong, A. P. T. Lau, C. Lu, and Y. Lu, "Modulation Format Identification based on Received Signal Power Distributions for Digital Coherent Receivers," 2014, doi: https://dx.doi.org/10.1109/OFC.2014.6886833.

[22] T. Bo, J. Tang, and C. C. K. Chan, "Blind Modulation Format Recognition for Software-defined Optical Networks Using Image Processing Techniques," in *Optical Fiber Communications Conference and Exhibition, OFC*, 2016, pp. 1–3, doi: 10.1364/ofc.2016.th2a.31.

[23] D. Wang *et al.*, "Modulation Format Recognition and OSNR Estimation Using CNN-Based Deep Learning," *IEEE Photonics Technol. Lett.*, vol. 29, no. 19, pp. 1667–1670, 2017, doi: 10.1109/LPT.2017.2742553.

[24] X. Lin, O. A. Dobre, T. M. N. Ngatched, Y. A. Eldemerdash, and C. Li, "Joint Modulation Classification and OSNR Estimation Enabled by Support Vector Machine," *IEEE Photonics Technol. Lett.*, vol. 30, no. 24, pp. 2127–2130, 2018, doi: 10.1109/LPT.2018.2878530.

[25] C. Panjaitan, A. Silaban, M. Napitupulu, and J. W. Simatupang, "Comparison K-Nearest (K-NN) and Artificial Neural Network (ANN) in Real Time Entrants Recognition," *2018 Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2018*, pp. 1–4, 2018, doi: 10.1109/ISRITI.2018.8864366.

Machine Learning-based Modulation Format Identification and Optical Performance Monitoring Techniques Implementation

ORIGINALITY REPORT

9%	3%	7%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Rania A. Eltaieb, Ahmed E. A. Farghal, HossamEl-din H. Ahmed, Waddah S. Saif et al. "Efficient Classification of Optical Modulation Formats Based on Singular Value Decomposition and Radon Transformation", Journal of Lightwave Technology, 2020 Publication	1%
2	Ángel Manuel Guerrero-Higueras, Noemí DeCastro-García, Vicente Matellán. "Detection of Cyber-attacks to indoor real time localization systems for autonomous robots", Robotics and Autonomous Systems, 2018 Publication	1%
3	www.osapublishing.org Internet Source	1%
4	dokumen.pub Internet Source	1%
5	Francesco Musumeci, Cristina Rottondi, Avishek Nag, Irene Macaluso, Darko Zibar, Marco Ruffini, Massimo Tornatore. "An	1%

	Overview on Application of Machine Learning Techniques in Optical Networks", IEEE Communications Surveys & Tutorials, 2019 Publication	
6	Submitted to University of Huddersfield Student Paper	1 %
7	Submitted to University of Energy and Natural Resources Student Paper	<1 %
8	Zhao Zhao, Aiying Yang, Peng Guo, Wenyu Tang. "A Modulation Format Identification Method Based on Amplitude Deviation Analysis of Received Optical Communication Signal", IEEE Photonics Journal, 2019 Publication	<1 %
9	Rania A. Eltaieb, Heba A. E. Abouelela, Waddah S. Saif, Amr Ragheb et al. "Modulation format identification of optical signals: an approach based on singular value decomposition of Stokes space projections", Applied Optics, 2020 Publication	<1 %
10	www.tandfonline.com Internet Source	<1 %
11	arxiv.org Internet Source	<1 %

12	Rohit Yadav, Sagar Pande, Aditya Khamparia. "Chapter 27 Breast Cancer Classification Using Convolution Neural Network (CNN)", Springer Science and Business Media LLC, 2021 Publication	<1 %
13	iopscience.iop.org Internet Source	<1 %
14	"Computer Vision – ECCV 2018", Springer Science and Business Media LLC, 2018 Publication	<1 %
15	Danish Rafique. "Machine Learning Based Optimal Modulation Format Prediction for Physical Layer Network Planning", 2018 20th International Conference on Transparent Optical Networks (ICTON), 2018 Publication	<1 %
16	F. Meng, S. Yan, K. Nikolovgenis, Y. Ou, R. Wang, Y. Bi, E. Hugues-Salas, R. Nejabati, D. Simeonidou. "Field Trial of Gaussian Process Learning of Function-Agnostic Channel Performance Under Uncertainty", Optical Fiber Communication Conference, 2018 Publication	<1 %
17	Waddah S. Saif, Maged A. Esmail, Amr M. Ragheb, Tariq A. Alshawi, Saleh A. Alshebeili. "Machine Learning Techniques for Optical	<1 %

Performance Monitoring and Modulation Format Identification: A Survey", IEEE Communications Surveys & Tutorials, 2020
Publication

18	"Advanced Informatics for Computing Research", Springer Science and Business Media LLC, 2021 Publication	<1 %
19	trilogi.ac.id Internet Source	<1 %
20	Latifa Guesmi, Amr Mohamed Ragheb, Habib Fathallah, Mourad Menif. "Experimental Demonstration of Simultaneous Modulation Format/Symbol Rate Identification and Optical Performance Monitoring for Coherent Optical Systems", Journal of Lightwave Technology, 2018 Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		