Deep Learning-Based OSNR Estimation from Constellation Diagrams of DP m-PSK and DP m-QAM in Flexible Coherent Optical Receivers

Mateus F. de Araújo[1], Myke D. M. Valadão[2], Antônio M. C. Pereira[2], Éderson R. da Silva[1], Waldir S.S. Júnior[2], André L. A. da Costa[1]
[1]Federal University of Uberlândia (UFU), MG-Brazil
[2]Federal University of Amazonas (UFAM), AM-Brazil
Emails: mateusflausino@ufu.br, mykedouglas@ufam.edu.br, antoniopereira@ufam.edu.br, ersilva@ufu.br, waldirjr@ufam.edu.br, alacosta@ufu.br

***Abstract*— Accurate OSNR estimation is essential for maintaining signal quality in coherent optical communication systems. This work proposes a deep learning-based method for OSNR estimation using constellation diagrams as input. A dataset of over 19,000 images was generated through simulations with various modulation formats. We evaluated 15 CNN architectures, including MobileNetV3, ConvNeXt, DenseNet, and EfficientNet. ConvNeXtBase achieved the best results, with a MAE below 0.43 dB and R^2 above 0.98. The results demonstrate the effectiveness of computer vision models for accurate and non-intrusive OSNR prediction.**

***Keywords*— OSNR estimation, constellation diagrams, deep learning, coherent optical communications**

I. Introduction

Optical Signal-to-Noise Ratio (OSNR) is a fundamental parameter that directly influences the performance and reliability of modern optical communication systems [1]. As data traffic continues to grow exponentially due to the proliferation of cloud services, high-definition streaming, and 5G networks, the need for high-capacity and low-latency optical links has become critical. In such systems, OSNR serves as a primary indicator of signal quality, especially in dense wavelength-division multiplexing (DWDM) environments. Accurate OSNR estimation is essential not only for link maintenance and fault diagnosis but also for enabling adaptive modulation schemes that optimize transmission based on current channel conditions.

Recent advances in deep learning have opened new avenues for performance monitoring and signal analysis in optical communications [2]. Instead of relying solely on traditional signal processing techniques, deep neural networks have shown remarkable capabilities in modeling complex relationships within data, including tasks such as symbol detection, nonlinearity mitigation, and channel estimation [3] [4] [5]. These models can learn directly from raw or minimally processed data and generalize across various transmission conditions, offering high accuracy and robustness. As a result, deep learning is becoming a valuable tool for enhancing the intelligence and autonomy of optical networks.

One promising approach to OSNR estimation involves analyzing constellation diagrams, which visually represent the modulation state of a received signal [2]. These diagrams encapsulate key information about signal impairments, such as phase noise, amplitude distortions, and polarization effects [6]. By treating constellation diagrams as images, it becomes feasible to apply computer vision techniques to extract meaningful patterns and predict quality metrics like OSNR. Convolutional Neural Networks (CNNs), known for their success in image classification and regression, are particularly well-suited for this task, enabling accurate and efficient estimation even in the presence of channel impairments.

In this study, we propose a deep learning-based approach for OSNR estimation from constellation diagrams generated in a flexible coherent optical receiver setup. We evaluate a wide range of CNN architectures using a dataset of over 19,000 images across various modulation formats (DP m-PSK and DP m-QAM). The models are trained to perform direct OSNR regression from image inputs. Our best-performing models achieved mean absolute error (MAE) below 0.43 dB and R^2 scores above 0.98 in high-resolution settings, demonstrating the potential of deep learning and vision-based methods in advancing intelligent optical network monitoring.

A. Contributions

The main contributions of this study are as follows: (1) we propose a computer vision-based approach that treats constellation diagrams as images and employs state-of-the-art CNNs for direct OSNR estimation, eliminating the need for traditional signal processing pipelines; (2) a comprehensive dataset of over 19,000 labeled images was generated using a back-to-back simulation setup in VPIphotonics, covering various modulation formats and a wide range of OSNR values to enhance model robustness and generalization; (3) we conduct an extensive comparative evaluation of multiple deep learning architectures, assessing their performance under different conditions using standard regression metrics; and (4) our best-performing models, particularly ConvNeXtBase, achieved OSNR estimation with a mean absolute error below 0.43 dB and R^2 scores above 0.98, demonstrating state-of-the-art performance.

II. Related Works

Pereira et al. [1] conducted a comprehensive study on automatic modulation classification for flexible coherent optical

receivers, aiming to enhance transmission adaptability in next-generation optical networks. Utilizing a back-to-back simulation setup between a transmitter and a receiver, the authors generated a substantial dataset comprising 93,600 signals modulated using dual-polarization DP m-PSK and DP m-QAM schemes across thirteen distinct OSNR levels. The extracted signal features—specifically, histograms from in-phase and quadrature components—were used to train classical machine learning classifiers, including Random Forest, AdaBoost, and Decision Trees. Among the evaluated models, Random Forest achieved accuracy rates exceeding 97%, even when relying solely on a single polarization component, underscoring the viability and efficiency of lightweight classification approaches in optical communication systems.

In [7], Valadão et al. proposed a deep learning-based method for estimating noise power density using spectrograms derived from wireless signals. The authors introduced a robust dataset composed of signals subjected to varying levels of additive white Gaussian noise (AWGN) under diverse transmission conditions, including user mobility and channel impairments. To extract meaningful spectral features, the Hilbert transform was applied before spectrogram generation. These spectrograms served as inputs to several state-of-the-art CNN architectures for both classification and regression tasks. The results demonstrated high predictive accuracy, with deep architectures like ResNet152 achieving over 97%, highlighting the method's potential for reliable and efficient noise estimation in dynamic wireless environments.

In their study, Cho et al. [2] proposed a multitask CNN capable of simultaneously estimating linear ASE noise, nonlinear interference, and generalized optical signal-to-noise ratio (GOSNR) in wavelength-division multiplexed (WDM) fiber communication systems. The method extracts features from I-Q constellation density matrices obtained from dual-polarized 32-GBaud 16-QAM signals and demonstrates accurate noise component separation with mean absolute errors below 0.5 dB for $OSNR_{ASE}$, $OSNR_{NL}$, and GOSNR. Importantly, their approach does not require high-speed sampling, specialized symbols, or additional hardware, making it suitable for real-time monitoring in deployed metro networks. The study also validates the model's robustness across fiber types and explores the potential for universal training by incorporating link-specific normalization.

In [8], Wang et al. introduced a deep learning-based intelligent constellation diagram analyzer that simultaneously performs modulation format recognition (MFR) and OSNR estimation. Leveraging a CNN, their approach operates directly on raw constellation diagram images, eliminating the need for manual feature extraction or statistical preprocessing. The system was trained and tested on six widely-used modulation formats—QPSK, 8PSK, 8QAM, 16QAM, 32QAM, and 64QAM—across a range of OSNR values. The CNN achieved 100% classification accuracy for MFR and maintained OSNR estimation errors below 0.7 dB in both simulation and experimental validations. The study also benchmarked the CNN against traditional machine learning models, such as decision trees, SVMs, and KNNs, demonstrating superior accuracy and practical inference speed, highlighting its suitability for real-time optical performance monitoring.

III. Methods

This section describes the methodological pipeline adopted for OSNR estimation, including the system model, data generation process, image preprocessing steps, the deep learning architectures evaluated, and the performance metrics used for model assessment.

A. System model

The proposed system model follows a structured pipeline for OSNR estimation using constellation diagrams as inputs to deep learning models. The methodology begins with the generation of dual-polarization signals modulated with DP m-PSK and DP m-QAM formats in a back-to-back coherent optical transmission setup, as described in Subsection III-B. These signals are processed up to the stage of phase and quadrature compensation in the digital signal processor (DSP), after which constellation diagrams are extracted. The diagrams are saved as RGB images with controlled resolution and symbol opacity (α), simulating different visualization conditions. These images undergo preprocessing steps including resizing, normalization, and data augmentation, detailed in Subsection III-C. The resulting dataset is then used to train and evaluate a diverse set of CNN architectures, discussed in Subsection III-D, using regression metrics as presented in Subsection III-E. This framework enables a fully image-based, non-intrusive OSNR estimation approach that avoids conventional signal analysis algorithms and leverages recent advances in computer vision for performance monitoring in optical networks.

B. Data generation

The simulation setup illustrated in Figure 1 is based on a back-to-back configuration using VPIphotonics. It consists of a parallel-to-serial converter and a modulation mapper at the transmitter. The mapped signals are converted from digital to analog and used to modulate a laser carrier at $f_0 = 192.1$ THz, polarized in the $\overrightarrow{E_x}$ and $\overrightarrow{E_y}$ polarization states [1].

In the signal generation process, at the receiver side, the modulated signals pass through an optical band-pass filter (OBPF) centered at $f_c = 192.1$ THz, followed by a variable optical attenuator (VOA) to maintain a constant power of -2 dBm at the photodetector [1]. After attenuation, the polarization states are demultiplexed and fed into 2×4 ninety-degree hybrid couplers, along with the local oscillator signal centered at $f_{\text{LO}} = 192.1$ THz. The resulting in-phase and quadrature components are detected by photodetectors, amplified by transimpedance amplifiers, and filtered by low-pass filters (LPFs), before being processed by the DSP [1].

Within the DSP, the signal is converted from analog to digital, and phase and quadrature delays are compensated in the IQ compensation block. Chromatic dispersion and timing offsets are corrected by the dispersion compensation and time recovery blocks. Polarization-mode dispersion, frequency offset, and phase noise are subsequently addressed by the

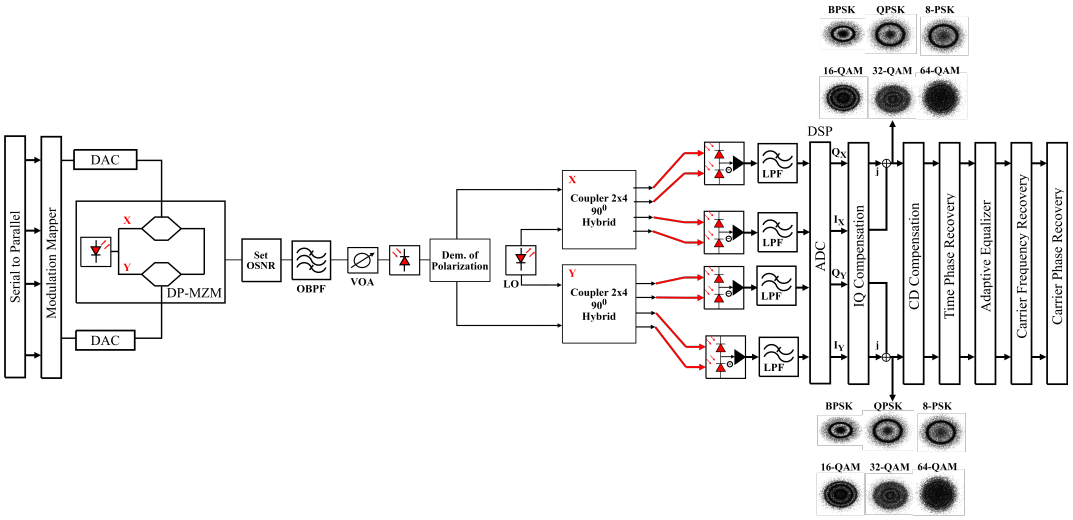

Fig. 1: Block diagram of the coherent optical transmitter and receiver used to generate the signals for training and testing [1]. Serial-to-parallel converter; DAC digital-to-analog converter; Laser diode; DP-MZM: dual-control Mach-Zehnder modulator; *set* OSNR: defines the OSNR value in dB/0.1 nm; LO: Local Oscillator; PD: Photodiode; Coherent optical receiver; ADC: Analog-to-digital converter.

dynamic equalizer, frequency recovery, and phase recovery blocks, respectively [1].

In this study, full digital signal processing is not required to obtain the constellation diagrams. After compensating for phase and quadrature delays, the dual-polarization BPSK, QPSK, 8PSK, 16QAM, 32QAM, and 64QAM signals are extracted and used to generate training and testing samples for the proposed neural network models. The resulting constellation diagrams are saved as images, whose visual clarity and distinguishability of symbol clusters are influenced by two key parameters: dots-per-inch (DPI), which defines the resolution of the image, and the α, which controls the overlap visibility between symbols. By varying these parameters, it is possible to simulate different levels of visual noise and representation granularity, which directly affect the learning behavior of the models during training and evaluation.

C. Pre processing

The image preprocessing consisted of systematic steps aimed at standardizing and enriching the dataset. Initially, the images were loaded from the paths specified in a csv file and resized to a fixed resolution, ensuring uniform input dimensions for the neural network. After resizing, the images were converted into numerical matrices of type uint8 and normalized to the [0, 1] range by dividing their pixel intensity values by 255, thus preparing them for compatibility with deep learning models [9].

To increase the variability of the training set and enhance the model's generalization capabilities, a data augmentation strategy was employed. The applied transformations included horizontal and vertical flipping, random rotations, random brightness and contrast adjustments, and a combined application of minor translations, scalings, and rotations [9]. These operations were applied exclusively to the training images, while the test images underwent only resizing and normalization. This ensured that their original characteristics were preserved, enabling an unbiased evaluation of the model's performance.

D. Proposed deep learning models

We evaluated a diverse set of CNN architectures, Table I, ranging from lightweight models optimized for resource-constrained environments to deeper networks with enhanced representational capacity. The selected models—MobileNetV3, ConvNeXt, DenseNet, ResNet, EfficientNet, Xception, and InceptionV3—represent a broad spectrum of design strategies [10][11]. This diversity enabled a thorough performance comparison in OSNR estimation from constellation diagrams, considering both accuracy and computational efficiency.

TABLE I: Summary of CNN architectures and their variants.

Model	**Variants**
MobileNetV3 [10][11]	Small, Large
ConvNeXt [10][11]	Tiny, Small, Base
DenseNet [10][11]	121, 169, 201
ResNet [10][11]	50, 101, 152
EfficientNet [10][11]	B0, B3
Xception [10][11]	–
InceptionV3 [10][11]	–

E. Evaluation Metrics

To assess the regression models, we adopted four standard metrics: MAE, mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination (R^2), as in [5][4]. These metrics offer complementary insights into prediction quality:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4}$$

In the above, y_i and $\hat{y}_i$ denote the true and predicted values, $\bar{y}$ is the mean of the true values, and n is the number of samples.

IV. Experiments and Results

This section presents the experimental setup, training configuration, dataset characteristics, and quantitative results obtained from evaluating multiple CNN architectures on the OSNR estimation task using constellation diagram images.

A. Parameters

Table II summarizes the training configuration and model parameters used in this study. All models were trained using RGB constellation images resized to 224×224 pixels. A batch size of 32 and 100 training epochs were adopted, with 15% of the training data reserved for validation. The dataset was split into 80% for training and 20% for testing, and shuffling was applied to ensure randomness. The models were initialized with ImageNet-pretrained weights and fine-tuned using the Adam optimizer. MSE was used as the loss function, while the MAE served as the main evaluation metric. Fine-tuning was enabled by allowing the base model weights to be trainable. A global average pooling layer and a linear activation function were used in the final layers to support the regression objective of OSNR prediction.

TABLE II: Training configuration and model parameters.

Parameter	Value
Input Image Size	224×224 pixels
Number of Channels	3 (RGB)
Batch Size	32
Number of Epochs	100
Validation Split	15% (from training data)
Train/Test Split	80% train / 20% test
Shuffle	True
Random State (Seed)	Fixed (for reproducibility)
Initial Learning Rate	1×10^{-4}
Fine-tuning Learning Rate	1×10^{-5}
Optimizer	Adam
Loss Function	MSE
Evaluation Metrics	MAE
Pretrained Weights	ImageNet
Base Model Trainable	Yes (fine-tuning enabled)
Global Pooling Layer	GlobalAveragePooling2D
Final Activation	Linear
Output Dimension	1 (regression task)

B. Dataset

The dataset used was generated according to the specifications presented in Table III, totaling 19,050 images. The OSNR values were incremented with a step of 0.5. Additionally, the DPI and α levels were varied to increase dataset diversity, being set to 100 and 1.0, and 150 and 0.2. Figure 2 presents the constellation diagrams of a QPSK signal under two distinct OSNR conditions: 12 dB and 20 dB. These diagrams provide a visual representation of how noise affects signal quality in optical communication systems. At an OSNR of 12 dB (Figure 2a), the constellation points exhibit significant dispersion and overlap, which indicates a higher likelihood of symbol detection errors due to noise interference. In contrast, the constellation at 20 dB OSNR (Figure 2b) displays well-separated and tightly clustered points, reflecting a cleaner signal with improved detection reliability. Both images were generated at a resolution of 150 DPI to ensure high visual fidelity for detailed analysis.

TABLE III: Statistical summary by modulation type

Modulation	Image Count	OSNR Min	OSNR Max
64QAM	4452	22.0	33.0
16QAM	3799	17.0	26.0
QPSK	3400	12.0	20.0
BPSK	2999	8.0	15.0
32QAM	2600	15.0	21.0
8PSK	1800	17.0	21.0

(a) QPSK constellation at 12 dB OSNR

(b) QPSK constellation at 20 dB OSNR

Fig. 2: Constellation diagrams for QPSK modulation at different OSNR levels.

To improve the generalization ability of the model and reduce overfitting, a data augmentation pipeline was applied using the `Albumentations` library. The following transformations were composed:

- Horizontal Flip with a probability of 50%;
- Vertical Flip with a probability of 50%;
- Random Rotation within a range of ± 30 degrees, applied with 50% probability;
- Random Brightness and Contrast Adjustment with 50% probability;
- Shift, Scale, and Rotate: includes random shifting (up to 5% of the image), scaling (up to 10%), and rotation (within ± 15 degrees), each applied with 50% probability.

This augmentation strategy increases the variability of the training set and simulates different real-world conditions, thus helping the model to learn more robust and invariant features.

C. OSNR estimation

Tables IV and V present the final test results of all evaluated models under two different visualization settings for constellation diagrams: DPI = 100 with $\alpha = 1.0$, and DPI = 150 with $\alpha = 0.2$. The models were assessed using four regression metrics: RMSE, MAE, MSE, and R^2. In both scenarios, the ConvNeXtBase model achieved the best overall performance, with a minimum MAE of 0.4241 and an R^2 of 0.9900 when using higher resolution and lower opacity (DPI = 150, $\alpha = 0.2$). Other high-performing models include ConvNeXtSmall and ResNet152, which also presented MAE values well below 0.6 in the enhanced setting.

Compared to related studies in the literature, which reported MAE values of approximately 0.5 [2] and 0.7 [8] for OSNR estimation using deep learning techniques, our results demonstrate a clear improvement in this metric. These gains can be attributed to the use of high-resolution constellation images, extensive data augmentation, and the evaluation of recent convolutional architectures specifically designed for visual pattern recognition. The low MAE values and high R^2 scores achieved confirm the effectiveness of the proposed image-based approach for OSNR estimation in coherent optical systems.

TABLE IV: Final test results for each model using RMSE, MAE, MSE, and R^2 metrics (DPI = 100 and α = 1).

Model	RMSE	MAE	MSE	R^2
MobileNetV3Small	2.6867	1.9759	7.2183	0.7893
MobileNetV3Large	1.8833	1.3889	3.5467	0.8465
ConvNeXtTiny	1.3413	1.0056	1.7991	0.9475
ConvNeXtSmall	1.4429	1.0824	2.0821	0.9392
ConvNeXtBase	1.0073	0.7260	1.0146	0.9698
DenseNet121	1.1965	0.8514	1.4317	0.9573
DenseNet169	1.3416	0.9723	1.7998	0.9464
DenseNet201	1.1847	0.8448	1.4036	0.9582
ResNet50	1.5255	1.0931	2.3270	0.9307
ResNet101	1.4209	1.0141	2.0189	0.9412
ResNet152	1.2849	0.9035	1.6510	0.9519
EfficientNetB0	1.9216	1.3805	3.6924	0.8924
EfficientNetB3	1.6092	1.2015	2.5894	0.9245
InceptionV3	1.8921	1.3888	3.5799	0.8957
Xception	2.1604	1.5914	4.6673	0.8640

TABLE V: Final test results for each model using RMSE, MAE, MSE, and R^2 metrics (DPI = 150 and α = 0.2).

Model	RMSE	MAE	MSE	R^2
MobileNetV3Small	2.4648	1.8346	6.0754	0.8292
MobileNetV3Large	6.176	4.8008	38.1512	-0.0549
ConvNeXtTiny	0.8381	0.5970	0.7024	0.9805
ConvNeXtSmall	0.7869	0.5488	0.6192	0.9828
ConvNeXtBase	0.6007	0.4241	0.3608	0.9900
DenseNet121	0.9412	0.6715	0.8859	0.9755
DenseNet169	0.8425	0.6088	0.7099	0.9803
DenseNet201	0.8925	0.6374	0.7966	0.9779
ResNet50	0.9637	0.6962	0.9287	0.9743
ResNet101	0.7314	0.5059	0.5349	0.9850
ResNet152	0.6487	0.4450	0.4208	0.9882
EfficientNetB0	1.5831	1.2731	2.5062	0.9295
EfficientNetB3	1.0524	0.7738	1.1077	0.9689
InceptionV3	0.9616	0.6854	0.9246	0.9740
Xception	1.3035	0.9400	1.6991	0.9522

V. Conclusions

This study presented a deep learning-based approach for estimating OSNR directly from constellation diagrams generated in flexible coherent optical receivers. By leveraging a diverse dataset encompassing various DP m-PSK and DP m-QAM modulation formats and evaluating multiple state-of-the-art convolutional neural networks, we demonstrated that models such as ConvNeXtBase can accurately predict OSNR with a MAE below 0.43 dB and R^2 scores exceeding 0.98. The results confirm the potential of computer vision techniques for performance monitoring in optical networks, offering a simplified and efficient alternative to traditional OSNR estimation methods. As future work, we intend to further enhance model robustness by expanding the dataset with greater variability in parameters such as linewidth, symbol rate, DPI, α, and polarization effects. Additionally, we plan to explore transformer-based architectures, including Vision Transformers (ViT), which may offer improved generalization capabilities in complex scenarios due to their global attention mechanisms.

References

[1] Antônio MC Pereira, Rafael S Furtado, Diego A Amoedo, Myke DM Valadão, Celso B Carvalho, André LA da Costa, and Waldir SS Júnior. Classificacao automatica de modulacoes dp m-psk e dp m-qam em receptores opticos coerentes flexiveis. In *XXXIX Simposio Brasileiro de Telecomunicacoes e Processamento de Sinais (SBrT). SBrT*, pages 1–5, 2021.

[2] Hyung Joon Cho, Daniel Lippiatt, Varghese A Thomas, Siddharth Varughese, Steven Searcy, Thomas Richter, Sorin Tibuleac, and Stephen E Ralph. Constellation-based identification of linear and nonlinear osnr using machine learning: a study of link-agnostic performance. *Optics Express*, 30(2):2693–2710, 2022.

[3] Shilian Zheng, Shurun Chen, Tao Chen, Zhuang Yang, Zhijin Zhao, and Xiaoniu Yang. Deep learning-based snr estimation. *IEEE Open Journal of the Communications Society*, 5:4778–4796, 2024.

[4] Myke Valadão, Diego Amoedo, André Costa, Celso Carvalho, and Waldir Sabino. Predicting noise and user distances from spectrum sensing signals using transformer and regression models. *Applied Sciences*, 15(8):4296, 2025.

[5] Myke Douglas de Medeiros Valadão. *Sensing, Estimation, and Security of the Frequency Spectrum Using Shallow and Deep Learning Techniques*. Tese de doutorado, Universidade Federal do Amazonas, Outubro 2024. https://tede.ufam.edu.br/handle/tede/10605.

[6] Van-Sang Doan, Thien Huynh-The, Cam-Hao Hua, Quoc-Viet Pham, and Dong-Seong Kim. Learning constellation map with deep cnn for accurate modulation recognition. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pages 1–6. IEEE, 2020.

[7] Myke DM Valadão, André LA da Costa, Éderson R da Silva, Alexandre C Mateus, and Waldir SS Júnior. Noise power density estimation based on deep learning using spectrograms extracted from wireless signals. In *XLII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT). SBrT*, pages 1–5, 2024.

[8] Danshi Wang, Min Zhang, Jin Li, Ze Li, Jianqiang Li, Chuang Song, and Xue Chen. Intelligent constellation diagram analyzer using convolutional neural network-based deep learning. *Optics express*, 25(15):17150–17166, 2017.

[9] Kiran Maharana, Surajit Mondal, and Bhushankumar Nemade. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1):91–99, 2022.

[10] S Deepa, J Loveline Zeema, and S Gokila. Exploratory architectures analysis of various pre-trained image classification models for deep learning. *Journal of Advances in Information Technology*, 15(1):66–78, 2024.

[11] Hafeez Anwar. Selecting the best backbone model: A comprehensive evaluation of deep learning models for satellite image-based land cover classification. *Earth Science Informatics*, 18(2):334, 2025.