# Convolutional and Variational Autoencoders with Supervised Classifiers for RF Spectrum Anomaly Detection

Szilárd László Takács*, and Konrád Kajdy

***Abstract*—Effective and secure monitoring of the radio frequency spectrum is essential for the reliable operation of modern wireless communication systems. Automated detection of spectral anomalies can support regulatory and operational monitoring tasks. This study evaluates three autoencoder architectures — Vanilla Autoencoder (AE), Convolutional Autoencoder (CAE), and Variational Autoencoder (VAE) — for anomaly detection using waterfall image representations derived from FM-band (86.5–108 MHz) measurement data. The models were assessed both as standalone reconstruction-based detectors and as feature extractors combined with supervised classifiers. Standalone autoencoders achieved F1-scores between 0.794 and 0.813. Higher performance was obtained when encoder-derived latent representations were used with supervised models, where the best result (F1-score: 0.915) was achieved by the CAE + ExtraTrees combination. The results indicate that hybrid encoder–classifier approaches can provide an effective practical solution for anomaly detection in spectrum monitoring environments. While promising, the findings are limited to the investigated FM-band dataset and require further validation across broader spectrum environments.**

***Index Terms*—spectrum monitoring; autoencoder; anomaly detection; RF spectrum anomalies**

## I. Introduction

Ensuring uninterrupted wireless communication has become increasingly challenging due to the growing number of radio systems and devices sharing the spectrum. Interference, unauthorized emissions, and technical malfunctions may affect multiple frequency bands, including the frequency modulated (FM) radio band (86.5–108 MHz). Consequently, efficient monitoring tools are required to support timely anomaly detection and response.

In Hungary, wireless communication is regulated under Act C of 2003 on Electronic Communications [1]. Spectrum supervision is supported by a nationwide monitoring infrastructure consisting of 67 measurement stations. In current operational practice, substantial parts of anomaly assessment still require expert interpretation. As spectrum usage becomes increasingly dynamic, automated decision-support methods may improve efficiency and consistency.

Szilárd László Takács Department of Electrical Engineering and Infocommunications, Széchenyi István University, Győr, Hungary; (e-mail: takacs.szilard.laszlo@sze.hu).

Konrád Kajdy Széchenyi István University, Győr, Hungary; (e-mail: kajdykonrad@gmail.com).

Recent advances in machine learning have created new opportunities for supporting automated RF spectrum analysis. In particular, image-based representations such as waterfall diagrams enable the application of computer vision and representation-learning methods to spectrum monitoring tasks. However, comparative evidence on how different autoencoder architectures perform on real FM-band monitoring data remains limited.

The contribution of this study is primarily empirical. First, three commonly used autoencoder architectures are compared on real FM-band waterfall measurements. Second, the usefulness of combining encoder-derived latent representations with conventional supervised classifiers is examined. The goal is not to introduce a novel architecture, but to evaluate a practical hybrid pipeline for operational RF anomaly detection.

## II. Related Work

Anomaly detection in RF spectrum monitoring is inherently challenging because abnormal events may vary substantially in frequency, duration, intensity, and temporal structure. Frequently recurring and well-defined events, such as transmission start/stop patterns, persistent level changes, or known interference types, can often be addressed using supervised machine learning methods when labeled examples are available. However, rare, previously unseen, or weakly characterized anomalies remain more difficult to model in a purely supervised setting. For this reason, unsupervised representation-learning approaches, including autoencoders, continue to receive attention.

The application of machine learning to radio spectrum analysis is an emerging research area. Prior studies have investigated unsupervised spectrum anomaly detection, spectrum sensing, and interference recognition. The SAIFE framework, based on an adversarial autoencoder, demonstrated anomaly detection and localization using power spectral density data [2]. Stacked autoencoder models have also been applied to OFDM spectrum sensing, showing robustness against noise and synchronization uncertainties [3]. These studies indicate that latent-feature learning can be useful when explicit anomaly labels are limited. Beyond RF applications, autoencoder-based methods have also been studied in adjacent signal and image domains, including hyperspectral image classification and industrial fault detection

DOI: 10.36244/ICJ.2026.2.7

[4]–[7]. Although these domains differ from RF monitoring, they provide evidence that learned low-dimensional representations may improve downstream classification tasks.

At the same time, several limitations of reconstruction-based anomaly detection have been reported. Prior work has shown that anomalous samples may sometimes be reconstructed with low error, which can reduce detection sensitivity [8]. Dataset topology may also affect reconstruction behavior [9], while contaminated training data can degrade robustness [10]. In addition, adversarial manipulation has been identified as a potential deployment risk [11]. These findings suggest that reconstruction error alone may be insufficient in some practical scenarios.

Motivated by these limitations, hybrid approaches that combine unsupervised feature extraction with supervised classification may offer a more reliable alternative. However, comparative studies evaluating multiple autoencoder architectures on real FM-band waterfall measurements remain limited. This study addresses that gap by comparing vanilla, convolutional, and variational autoencoders both as standalone detectors and as feature extractors for downstream classifiers.

More recent studies have also explored convolutional neural networks and transformer-inspired architectures for modulation recognition, interference classification, and wideband spectrum sensing, indicating a broader shift toward end-to-end learned RF representations [12], [13].

Unlike prior studies focusing on single-model designs, this work emphasizes comparative benchmarking under identical measurement conditions

## III. Materials and Methods

This section describes the measurement dataset, preprocessing workflow, model architectures, training configuration, and evaluation protocol used in the study. Particular attention was given to preventing data leakage and obtaining robust performance estimates despite the limited dataset size.

### *A. Data Collection and Dataset Construction*

Data collection was carried out at two stations of the Hungarian national spectrum monitoring network, located in Dobogókő and Kisvárda, within the FM broadcast band (86.5–108 MHz). Waterfall diagrams were generated from the measurements, where the x-axis represents frequency, the y-axis represents time (400-second intervals), and the color scale represents received signal strength.

To ensure consistency across stations, measured values were linearly rescaled to the interval −20 to 120 dBμV. The measurement frequency resolution was 5 kHz.

The resulting dataset contained 1,787 training images and 280 evaluation images. The hold-out evaluation set contained 135 anomalous and 145 normal samples, while the training partition consisted only of samples not appearing in the evaluation subset. Because the available dataset size is moderate for deep learning applications, regularization, augmentation, and cross-validation were incorporated into the evaluation pipeline.

### *B. Data Preprocessing and Annotation*

Anomaly labels were assigned through expert inspection by multiple spectrum monitoring specialists. Independent assessments were reconciled through consensus review to reduce subjectivity.

From the original waterfall diagrams, image patches of 640 × 640 pixels were extracted and resized to 128 × 128 pixels for model input. Data augmentation included random horizontal and vertical flips, brightness/contrast variation, and normalization.

No synthetic anomalies were introduced. No explicit denoising was applied, as preserving naturally occurring RF noise patterns was considered important for operational realism.

### *C. Autoencoder Architectures*

Three encoder–decoder architectures were evaluated: a fully connected autoencoder (AE), a convolutional autoencoder (CAE), and a variational autoencoder (VAE). These models were selected to compare deterministic dense representations, spatially aware convolutional representations, and probabilistic latent representations within a common experimental framework.

- Vanilla AE: A symmetric, fully connected neural network that reduces the 49,152-dimensional input space to a 512-dimensional latent representation through successive layers of 4096, 2048, and 1024 neurons using ReLU activation. The decoder mirrors the encoder in reverse, with a sigmoid output activation ensuring reconstructions remain in the [0,1] range.
- CAE: A convolutional architecture comprising five sequential Conv2D layers (32–256 filters, 4×4 kernels, stride = 2, padding = 1) with ReLU activations. Compression is finalized by an Adaptive Average Pooling layer, followed by a linear projection into a 128-dimensional latent space. Reconstruction is achieved through transposed convolutional layers, gradually reducing the filter size from 256 to 32, and finalized with a sigmoid activation.
- VAE: Architecturally similar to the CAE encoder, but the latent space is modeled as a probabilistic distribution. Two fully connected layers output the mean and log-variance of the latent space, from which a 128-dimensional latent vector is sampled. The decoder is identical to that of the CAE.

### *D. Model Training Procedure*

All autoencoder models were trained for 50 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 64. Mean squared error (MSE) was used as the reconstruction loss for all autoencoder architectures. Model parameters were updated through backpropagation. Training was performed on GPU when CUDA was available.

After training, only the encoder components were retained and used to generate latent feature representations. These extracted features were then used as inputs for supervised classifiers (Random Forest, Support Vector Machine, ExtraTrees, Multi-Layer Perceptron, and others), each optimized using its standard learning objective.

To reduce overfitting risk, data augmentation and held-out evaluation were applied. Future work may investigate larger datasets and early stopping strategies.

### E. *Feature Extraction*

After training, the encoder components were separated and used as feature extractors. This produced latent representations of dimension 512 for AE and 128 for both CAE and VAE.
These compact representations were then used as input features for conventional supervised classifiers.

### F. *Supervised Classifiers*

The following supervised models were evaluated: Random Forest, Gradient Boosting, XGBoost, Logistic Regression, Support Vector Machine, k-Nearest Neighbours, Gaussian Naive Bayes, ExtraTrees, Multi-Layer Perceptron, and LightGBM.
This diverse set was selected to compare linear, kernel-based, ensemble, probabilistic, and neural classification paradigms.

### G. *Validation Strategy*

The hold-out test set was fixed prior to model development and remained untouched until final evaluation.
Autoencoder models were trained on a fully separate unsupervised training subset. Supervised classifiers were then trained on encoder-derived latent representations using a distinct labeled training subset. No samples, overlapping waterfall segments, or duplicated measurements were shared across the autoencoder training, classifier training, and hold-out test partitions.
Five-fold cross-validation was applied only within the supervised classifier training subset for model selection and hyperparameter tuning. Final performance metrics were computed once on the independent hold-out test set.

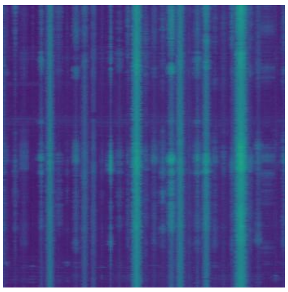

Fig. 1. Example of a waterfall diagram in the FM frequency band (86.5 – 108 MHz). The x-axis represents frequency, the y-axis represents time, and the color scale encodes signal strength.

TABLE I
THE BEST AND AVERAGE PERFORMANCE OF SUPERVISED CLASSIFIERS ON THE TEST SET FOR EACH AUTOENCODER ARCHITECTURE.

| Autoencoder | AE | CAE | VAE |
|---|---|---|---|
| Classifier | SVM | ExtraTrees | MLP |
| Accuracy | 0.893 | 0.911 | 0.893 |
| Precision | 0.839 | 0.844 | 0.862 |
| *Recall* | 0.963 | 1 | 0.926 |
| *F1-score* | 0.897 | 0.915 | 0.893 |
| *ROC AUC* | 0.861 | 0.966 | 0.950 |

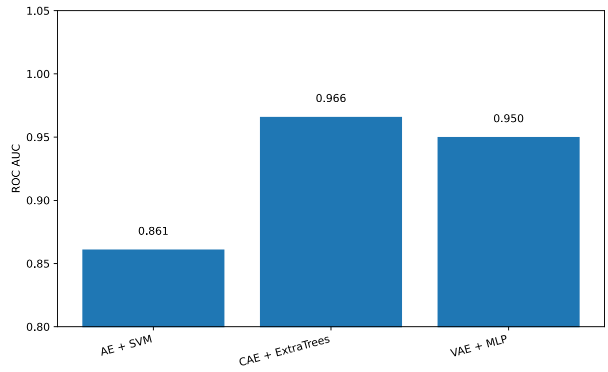


Fig. 2. ROC AUC of the best classifier–encoder combinations: AE + SVM, CAE +ExtraTrees, VAE + MLP.

## IV. RESULTS

This section summarizes the empirical performance of standalone autoencoders and hybrid encoder–classifier pipelines on the evaluated FM-band anomaly detection task.

### A. *Autoencoder-Based Reconstruction Performance*

The three autoencoder variants were first evaluated as reconstruction-based anomaly detectors. All models achieved comparable performance, with F1-scores ranging from 0.794 to 0.813.
Among the standalone approaches, the VAE produced the strongest result, followed closely by the vanilla AE, while the CAE achieved slightly lower performance. These findings suggest that latent-space regularization may provide modest benefits, although standalone reconstruction-based detection remained limited overall.

### B. *Feature Extraction and Supervised Classification*

Substantial improvements were observed when encoder outputs were used as feature representations for supervised learning. Across all three encoder families, multiple classifiers outperformed the corresponding standalone autoencoders.
The highest measured performance within the evaluated configurations was obtained by the CAE + ExtraTrees combination, which achieved the highest F1-score and recall among the evaluated models. The AE + SVM pipeline also performed strongly, particularly in sensitivity-oriented detection scenarios, while VAE + MLP yielded the most balanced trade-off across evaluation metrics.
Relative to standalone autoencoders, the hybrid pipelines improved F1-score by approximately 10.3% (AE), 15.2% (CAE), and 10.0% (VAE).
Detailed numerical results are summarized in Tables I–III, while ROC-based comparisons are presented in Fig. 2 and aggregate metric comparisons in Figs. 3–5.

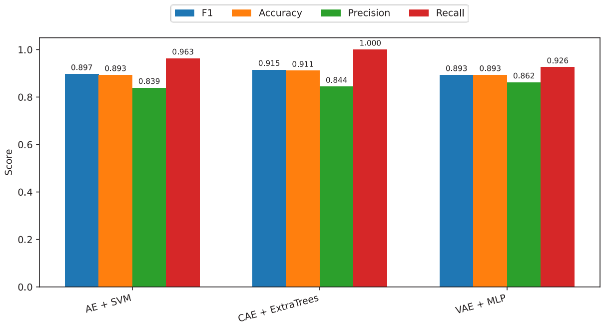


Fig. 3. Comparison of F1, Accuracy, Precision, and Recall for the best combinations.

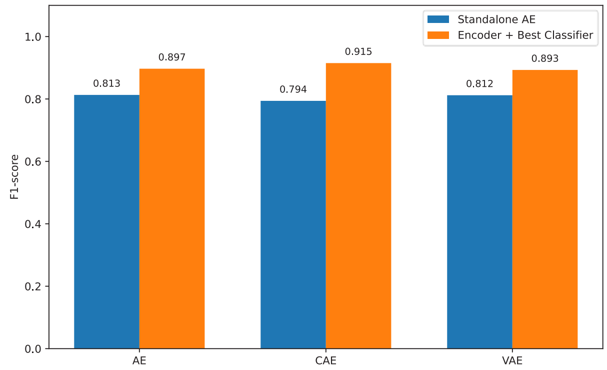


Fig. 4. F1-score comparison between standalone autoencoders and the best encoder–classifier pairs (per encoder).

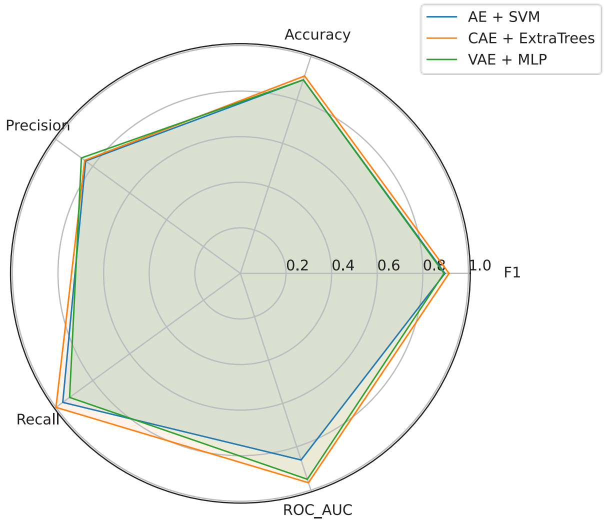


Fig. 5. Radar chart summarizing F1-score, Accuracy, Precision, Recall, and ROC AUC for the best-performing encoder–classifier combinations.

TABLE II
MEAN F1-SCORE OF SUPERVISED CLASSIFIERS ON THE VALIDATION SET FOR EACH AUTOENCODER ARCHITECTURE.

| Classifier | AE | CAE | VAE |
|---|---|---|---|
| RandomForest | $0.840 \pm 0.015$ | $0.798 \pm 0.040$ | $0.848 \pm 0.025$ |
| GradientBoosting | $0.858 \pm 0.025$ | $0.818 \pm 0.018$ | $0.864 \pm 0.039$ |
| *XGBoost* | $0.828 \pm 0.041$ | $0.825 \pm 0.042$ | $0.864 \pm 0.039$ |
| *LogisticRegression* | $0.818 \pm 0.039$ | $0.777 \pm 0.068$ | $0.858 \pm 0.026$ |
| *SVM* | $0.856 \pm 0.023$ | $0.781 \pm 0.045$ | $0.875 \pm 0.031$ |
| *kNN* | $0.852 \pm 0.021$ | $0.813 \pm 0.033$ | $0.847 \pm 0.039$ |
| *GaussianNB* | $0.803 \pm 0.051$ | $0.805 \pm 0.051$ | $0.816 \pm 0.045$ |
| *ExtraTrees* | $0.831 \pm 0.025$ | $0.817 \pm 0.045$ | $0.863 \pm 0.053$ |
| *MLP* | $0.841 \pm 0.063$ | $0.807 \pm 0.068$ | $0.874 \pm 0.042$ |
| *LightGBM* | $0.805 \pm 0.030$ | $0.823 \pm 0.025$ | $0.856 \pm 0.027$ |

TABLE III
CROSS-VALIDATION F1-SCORES OF SUPERVISED CLASSIFIERS FOR EACH ENCODER FEATURE SPACE.

| Classifier | AE | CAE | VAE |
|---|---|---|---|
| RandomForest | 0.864 | 0.826 | 0.884 |
| GradientBoosting | 0.894 | 0.844 | 0.930 |
| *XGBoost* | 0.878 | 0.889 | 0.930 |
| *LogisticRegression* | 0.857 | 0.864 | 0.905 |
| *SVM* | 0.889 | 0.844 | 0.930 |
| *kNN* | 0.870 | 0.851 | 0.894 |
| *GaussianNB* | 0.884 | 0.895 | 0.895 |
| *ExtraTrees* | 0.870 | 0.889 | 0.930 |
| *MLP* | 0.909 | 0.870 | 0.930 |
| *LightGBM* | 0.844 | 0.851 | 0.900 |

## V. DISCUSSION

The results indicate that combining encoder-derived latent representations with supervised classifiers appears to be an effective strategy for RF spectrum anomaly detection on the evaluated FM-band dataset. In all three investigated encoder families, hybrid pipelines outperformed the corresponding standalone reconstruction-based autoencoders, suggesting that latent features contained discriminative information not fully exploited by reconstruction error alone.

Among the evaluated configurations, the CAE + ExtraTrees model achieved the best performance among the evaluated configurations. This may indicate that convolutional encoders are well suited to capturing local spatial structures present in waterfall images, while ensemble classifiers can effectively separate normal and anomalous feature patterns. The VAE + MLP configuration produced a competitive and balanced alternative, while AE + SVM also yielded strong sensitivity-oriented performance.

These findings are consistent with prior studies showing that hybrid representation-learning pipelines can be useful when labeled anomaly data are limited [2]–[6]. Direct CNN baselines were outside the scope of this initial benchmarking study and are reserved for future work. In practical monitoring environments, such approaches may offer operational advantages because feature extraction and classification can be updated independently.

At the same time, the results should be interpreted within the scope of the present study. The experiments were conducted on a relatively limited dataset size for deep learning applications, restricted to the FM broadcast band, and based on measurements from two monitoring stations. Consequently, further validation is required before conclusions are extended to other spectrum bands, sensing environments, or anomaly categories.

## VI. Conclusions

This study comparatively evaluated three autoencoder architectures—AE, CAE, and VAE—for RF spectrum anomaly detection using FM-band waterfall measurements. Standalone reconstruction-based autoencoders achieved moderate detection performance, whereas substantially stronger results were obtained when encoder-derived latent representations were combined with supervised classifiers.

Among the evaluated configurations, the CAE + ExtraTrees pipeline achieved the best overall performance, suggesting that convolutional feature extraction is particularly well suited to capturing discriminative patterns in waterfall-image data.

These findings indicate that hybrid encoder–classifier approaches may serve as practical decision-support tools for spectrum monitoring environments where labeled anomalies are limited. However, the results are restricted to the investigated FM-band dataset, whose size is relatively limited for deep learning applications, and to the specific measurement conditions considered in this study.

Future work should include broader multi-band datasets, external validation, stronger end-to-end deep learning baselines, and deployment-oriented real-time testing.

## VII. Data Availability Statement

The RF spectrum measurement data used in this study were obtained from the Hungarian National Spectrum Monitoring Network, operated under the national regulatory authority. Due to regulatory and confidentiality restrictions, the full raw measurement data cannot be made publicly available.

- A limited set of 20 representative waterfall spectrum images [14] is openly available at Zenodo: S. L. Takács. (2025). Spectrum images in the FM frequency band. Zenodo. https://doi.org/10.5281/zenodo.17174395
- The program codes and Jupyter Notebook implementing the preprocessing and training pipeline, including autoencoder models and supplementary scripts [15], are openly available at Zenodo: S. L. Takács. (2025). Autoencoder Models and Supplementary Code for RF Spectrum Anomaly Detection (FM Band). Zenodo. https://doi.org/10.5281/zenodo.17174452

Additional derived data may be shared upon reasonable request, subject to regulatory approval.

## References

[1] Hungarian National Assembly, "Act C of 2003 on Electronic Communications," 2003. [Online]. Available: https://net.jogtar.hu/jogszabaly?docid=a0300100.tv. Accessed: Aug. 26, 2025.

[2] S. Rajendran, W. Meert, V. Lenders, and S. Pollin, "Unsupervised wireless spectrum anomaly detection with interpretable features," *IEEE Trans. Cogn. Commun.* Netw., vol. 5, pp. 637–647, 2019. DOI: 10.1109/TCCN.2019.2911524

[3] Q. Cheng, Z. Shi, D. N. Nguyen, and E. Dutkiewicz, "Deep learning network based spectrum sensing methods for OFDM systems," *arXiv preprint arXiv:1807.09414*, 2018. [Online]. Available: https://arxiv.org/abs/1807.09414. Accessed: Aug. 26, 2025.

[4] J. Zhao, L. Hu, Y. Dong, L. Huang, S. Weng, and D. Zhang, "A combination method of stacked autoencoder and 3D deep residual network for hyperspectral image classification," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 102, Art. no. 102 459, 2021. DOI: 10.1016/j.jag.2021.102459

[5] Z. Lin, Y. Chen, X. Zhao, and G. Wang, "Spectral–spatial classification of hyperspectral image using autoencoders," in *Proc. 9th Int. Conf. Inf., Commun. Signal Process. (ICICS)*, Tainan, Taiwan, Dec. 2013, pp. 1–5. DOI: 10.1109/ICICS.2013.6782778

[6] N. Tandiya, A. Jauhar, V. Marojevic, and J. H. Reed, "Deep predictive coding neural network for RF anomaly detection in wireless networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Kansas City, MO, USA, May 2018, pp. 1–6. DOI: 10.1109/ICCW.2018.8403654

[7] S. Ahmad, K. Styp-Rekowski, S. Nedelkoski, and O. Kao, "Autoencoder-based condition monitoring and anomaly detection method for rotating machines," in *Proc. IEEE Int. Conf. Big Data*, Atlanta, GA, USA, Dec. 2020, pp. 4093–4102. DOI: 10.1109/BigData50022.2020.9378015

[8] C. Zhou and R. C. Paffenroth, "Anomaly detection with robust deep autoencoders," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, Halifax, NS, Canada, Aug. 2017, pp. 665–674. DOI: 10.1145/3097983.3098052

[9] J. Batson, C. G. Haaf, Y. Kahn, and D. A. Roberts, "Topological obstructions to autoencoding," *J. High Energy Phys.*, vol. 2021, no. 4, pp. 1–43, 2021. DOI: 10.1007/JHEP04(2021)280

[10] L. Beggel, M. Pfeiffer, and B. Bischl, "Robust anomaly detection in deep autoencoders using adversarial training," *Mach. Learn.*, vol. 109, pp. 1611–1630, 2020. DOI: 10.48550/arXiv.1901.06355

[11] I. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015. [Online]. Available: https://arxiv.org/abs/1412.6572. Accessed: Aug. 26, 2025.

[12] A. Selim, F. Paisana, J. A. Arokkiam, Y. Zhang, L. Doyle, and L. A. DaSilva, "Spectrum monitoring for radar bands using deep convolutional neural networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6. DOI: 10.1109/GLOCOM.2017.8254105

[13] E. V. Vijay and K. Aparna, "Spectrum sensing using deep learning for proficient data transmission in wireless sensor networks," *Ain Shams Eng. J.*, 2024. DOI: 10.1016/j.asej.2024.102784

[14] S. L. Takács, "Spectrum Images in the FM Frequency Band," *Zenodo*, 2025. DOI: 10.5281/zenodo.17174395

[15] S. L. Takács, "Autoencoder Models and Supplementary Code for RF Spectrum Anomaly Detection (FM Band)," *Zenodo*, 2025. DOI: 10.5281/zenodo.17174452

**Szilárd László Takács** was born in Szombathely, Hungary, on July 25, 1997. He received the B.Sc. degree in electrical engineering in 2019 and the M.Sc. degree in electrical engineering in 2022 from Széchenyi István University, Győr, Hungary. He also received the M.Sc. degree in physics in 2022 from the Budapest University of Technology and Economics, Budapest, Hungary, and the postgraduate diploma in mathematics expert in data analytics and machine learning in 2023 from Eötvös Loránd University, Budapest, Hungary. He began his Ph.D. studies in 2022 at Széchenyi István University, Győr, Hungary, in the Multidisciplinary Doctoral School of Engineering Sciences, specializing in informatics. His major field of study is artificial intelligence.

He is currently working at the National Media and Infocommunications Authority (NMHH), Hungary, as a Technology Analyst. He also serves as an Assistant Lecturer at Széchenyi István University, Győr, Hungary, where he teaches courses in Microwave Technology I–II, Signal Processing and Quantum Communication. He previously held teaching positions at the Budapest University of Technology and Economics as a Demonstrator and Teaching Assistant in Mathematics between 2021 and 2022, and he was a Research Assistant at the Wigner Research Centre for Physics, Department of Applied and Nonlinear Optics, in the Ultrafast Nanooptics "Lendület" Research Group. His research focuses on the application of artificial intelligence in signal processing. He has published works in this field, including "Applicability of Vanilla Autoencoder and Convolutional Autoencoder in the Field of Radio Spectrum Monitoring," presented at the XXIX International PhD Conference of Professors for the European Hungary Association, Budapest, Hungary, in 2025.

Mr. Takács is a member of several national professional communities in the fields of telecommunications and information technology. His current professional interests include artificial intelligence, data-driven technologies, and signal processing. He is actively involved in academic education and research collaboration projects aimed at integrating machine learning techniques into modern communication systems.

**Konrád Kajdy** was born in Nagyatád, Hungary in 2004. He is currently studying computer science at Széchenyi István University in Győr, and will soon receive his B.Sc. degree in 2026.

In 2024 he spent a semester abroad in Seoul, South Korea at Chung-Ang University. Currently he is working at NMHH in Budapest, mainly focusing on researching modern technologies.