



WAVELETS IN ARTIFICIAL INTELLIGENCE: A REVIEW

Mathematics

Mamta Rani*

Assistant Professor Chandigarh University, Mohali, India-140413 *Corresponding Author

ABSTRACT

Wavelet transforms originally developed for signal processing and time-frequency analysis have increasingly found utility in artificial intelligence (AI). Their ability to capture both spatial and frequency information at multiple resolutions makes them ideal for tasks involving non-stationary data. This article reviews the role of wavelets in AI, including their integration with machine learning, deep learning, and hybrid models, and outlines emerging research trends and challenges.

KEYWORDS

Wavelets, artificial intelligence, Deep learning.

INTRODUCTION

Wavelets are mathematical functions that decompose data into different frequency components while retaining spatial (or temporal) information. In contrast to traditional Fourier analysis, wavelets offer localized analysis and are well-suited for non-stationary signals. As AI systems increasingly deal with complex, high-dimensional, and noisy data, wavelets provide an efficient preprocessing and feature extraction tool.

Wavelet transform can be considered as a refinement of Fourier transform. Fourier transform is localized in only one domain either time or frequency hence it is not useful for non-stationary signals. Wavelets can examine signals simultaneously in both time and frequency domain. Wavelets have ability to analyze aperiodic, transient and non-stationary signals. It has applications in the field of data compression, computer graphics, numerical analysis, signal processing, pattern recognition, image processing, data mining and other medical image technology like EEG, ECG etc [1-2].

Wavelet Transforms Come In Two Major Forms:

- Continuous Wavelet Transform (CWT): Offers a continuous analysis across scales and translations; ideal for fine-grained analysis but computationally expensive.
- Discrete Wavelet Transform (DWT): More efficient; often used in practical AI applications. Popular bases include Haar, Daubechies, and Symlets.

Wavelets are particularly powerful for multiresolution analysis (MRA), enabling decomposition of a signal into coarse and detail components [3-5].

Applications In AI Feature Extraction

Wavelet transforms are widely used to extract discriminative and compact features from various types of data due to their ability to localize information in both time (or space) and frequency domains. In image processing, wavelets are applied for tasks such as edge detection, texture analysis, and object recognition, where high-frequency coefficients highlight fine details like edges, while low-frequency coefficients capture global shape and structure. These features are especially valuable in applications such as medical imaging (e.g., MRI, CT scans), remote sensing, and biometric authentication.

In the audio domain [6], wavelet-based features effectively represent time-varying spectral content, making them ideal for speech and speaker recognition, music genre classification, and audio event detection [7]. Wavelets outperform traditional Fourier-based methods in capturing transient acoustic phenomena due to their multiresolution nature.

For time-series data, such as electrocardiogram (ECG) or electroencephalogram (EEG) signals, wavelets enable precise identification of anomalies like arrhythmias or epileptic spikes [8-9]. In financial data analysis, they facilitate the extraction of features reflecting trends, volatility, or abrupt market changes, which are critical for tasks like algorithmic trading and fraud detection.

Once extracted, wavelet features are typically input into machine

learning classifiers, including support vector machines (SVMs), decision trees, random forests, or artificial neural networks (ANNs). These features often result in improved classification accuracy, reduced model complexity, and greater robustness to noise compared to raw input data or conventional feature sets [10].

Moreover, feature selection or dimensionality reduction techniques such as principal component analysis (PCA) are often applied to wavelet-transformed features to enhance model performance and generalizability, particularly in high-dimensional datasets.

Signal And Image Denoising

AI models often suffer from degraded performance or overfitting when trained on data contaminated by noise. Denoising is therefore a critical preprocessing step, and wavelet-based techniques have proven particularly effective due to their ability to separate noise from meaningful signal content at different scales. Unlike traditional linear filters, which may blur edges or distort important structures, wavelet denoising maintains high fidelity to the original signal by selectively attenuating noise components in the wavelet domain.

The most commonly used wavelet-based denoising approach involves thresholding, where small wavelet coefficients—typically associated with noise—are either set to zero (hard thresholding) or shrunk toward zero (soft thresholding), while larger coefficients, which represent significant signal features, are preserved [11]. This allows for noise suppression without compromising critical signal characteristics such as edges in images or sharp transitions in time-series data [12].

Wavelet Denoising Is Widely Employed In:

- Biomedical signal processing, including the cleaning of ECG, EEG, and EMG signals to improve the reliability of downstream diagnostic algorithms (e.g., arrhythmia detection, seizure prediction, or sleep stage classification) [13-14].
- Remote sensing and satellite imaging, where wavelets enhance spatial resolution and remove atmospheric or instrumental noise, thereby improving land cover classification or object detection [15-16].
- Speech and audio processing, where wavelet denoising improves intelligibility and recognition accuracy in noisy environment [17-18].
- Digital photography and surveillance, to reduce sensor or compression artifacts while preserving image detail critical for facial recognition and object tracking tasks.
- Additionally, wavelet-based denoising can be integrated into deep learning pipelines, where it acts as a preprocessor to stabilize training, or even embedded directly within architectures, such as wavelet-based convolutional neural networks (Wavelet CNNs), to enable learned noise suppression [19-20].

Deep Learning Integration

- Recent research has increasingly explored the integration of wavelet transforms into deep neural networks to enhance their efficiency, interpretability, and generalization capabilities. Traditional deep learning architectures often rely on convolution and pooling layers to extract hierarchical features and reduce spatial dimensions [21-22]. However, these methods may lead to information loss, particularly in the high-frequency components

that are crucial for detailed pattern recognition. Wavelet transforms offer a more structured and information-preserving alternative.

- **Wavelet Convolutional Neural Networks (Wavelet CNNs):** In Wavelet CNNs, the Discrete Wavelet Transform (DWT) is employed in place of standard pooling layers (e.g., max-pooling) for spatial downsampling. This substitution retains both high- and low-frequency information at each level of decomposition, leading to improved feature richness and better generalization [23]. Empirical results have demonstrated that Wavelet CNNs outperform conventional CNNs in various image classification and segmentation tasks, particularly when training data is limited or noisy. Additionally, the inverse wavelet transform (IWT) can be used for upsampling in decoder parts of networks, making these models well-suited for autoencoders, generative models, and image super-resolution [24-25].
- **Wavelet Scattering Networks (WSNs):** WSNs are built using predefined wavelet filters instead of learned convolutional kernels, forming a non-trainable yet highly stable feature extractor. Scattering networks exploit the mathematical properties of wavelets—such as invariance to small translations and stability to deformations—to produce hierarchical features that are robust to signal variations [26]. These features are particularly useful in domains like texture classification, speech processing, and biomedical image analysis. WSNs offer interpretability and theoretical guarantees regarding invariance and stability, making them appealing for applications requiring transparency and explainability [27].
- Beyond these core architectures, researchers have also proposed hybrid models where wavelet coefficients are used as input features to deep networks, or where wavelet-based attention mechanisms are integrated into transformer models to provide multi-scale context awareness. Moreover, wavelet transforms can serve as a regularization strategy or be used to constrain feature learning in deep models, thereby reducing overfitting and enhancing robustness [28-30].
- Overall, integrating wavelets into deep learning not only improves model performance but also enhances interpretability and efficiency, particularly for applications involving structured, hierarchical, or non-stationary data.

Compression And Model Efficiency

Wavelets play a vital role in enhancing compression and model efficiency, particularly in scenarios where computational resources and memory are constrained, such as edge computing, mobile devices, IoT nodes, and embedded AI systems [31-32]. Their capacity to represent signals with high sparsity in the wavelet domain makes them highly effective for both input data compression and model parameter reduction [33].

On the input side, wavelet-based dimensionality reduction allows AI systems to focus on the most informative features while discarding redundant or noise-prone components. By decomposing high-dimensional data into approximation and detail coefficients, wavelets enable the selection or truncation of less significant components, thus reducing the input size without significant loss of critical information [34-35]. This preprocessing step is particularly useful for real-time inference in applications such as video analytics, wearable health monitoring, and remote sensing, where bandwidth and storage are limited.

In addition to input compression, wavelets contribute to neural network model compression by transforming and pruning internal representations. For example, weight matrices in dense or convolutional layers can be transformed using wavelet bases, and then thresholded or quantized to eliminate insignificant coefficients [36-38]. This results in sparser networks with fewer parameters, reduced memory footprint, and lower energy consumption, all while preserving or even enhancing model accuracy. This is particularly important for deploying AI models on resource-limited platforms like microcontrollers or FPGA-based accelerators.

Moreover, wavelet transforms can be incorporated into neural network architecture design, where layers perform wavelet-like filtering or downsampling. Such approaches reduce the need for parameter-heavy convolutional layers, making models more efficient in both training and inference [39-40]. In some cases, wavelet transforms are used in conjunction with model pruning, knowledge distillation, or

quantization-aware training to further optimize performance for on-device AI.

Overall, wavelets offer a principled and effective framework for balancing model complexity, computational efficiency, and representational power, enabling scalable AI deployment in environments where resources are severely limited.

Wavelets assist in reducing input dimensionality and compressing model weights (especially in edge computing or embedded AI systems) [41-42].

Hybrid Models

Combining wavelet transforms with machine learning techniques and evolutionary algorithms has given rise to hybrid systems that exhibit strong performance in a wide range of tasks, especially where data is noisy, complex, or limited. These hybrid approaches leverage the multiresolution analysis capability of wavelets for feature extraction, and pair it with powerful classification, regression, or optimization models for enhanced accuracy, robustness, and interpretability [43-44].

Typical Application Domains Include:

Fault Detection: Wavelet-based features capture transient patterns and subtle anomalies in industrial signals such as vibration, current, or acoustic emissions. When combined with classifiers like Support Vector Machines (SVMs) or Artificial Neural Networks (ANNs), these features enable early and accurate fault diagnosis in rotating machinery, power transformers, and aerospace systems [45-46].

Anomaly Detection: In time-series data (e.g., from sensors or cybersecurity logs), wavelets help isolate sudden, localized changes that signify anomalies [47]. When fused with clustering techniques or decision trees, they enable unsupervised or semi-supervised anomaly detection frameworks with low false-positive rates.

Classification Under Resource Constraints: In embedded systems and edge AI environments, hybrid wavelet models can reduce both input size and model complexity. For instance, wavelet-transformed features fed into lightweight classifiers such as k-NN or rule-based systems provide high accuracy with minimal computational cost [48].

Examples Of Hybrid Architectures Include:

- **Wavelet-Genetic Algorithms (Wavelet-GA):** Genetic algorithms are used to optimize wavelet parameters (e.g., wavelet type, decomposition level, or thresholding strategy), as well as to select the most discriminative features for classification [48]. This is particularly useful in applications like biometric authentication, remote sensing, and speech processing [49-50].
- **Wavelet-SVM systems:** These combine wavelet-transformed inputs with SVMs, capitalizing on wavelets for noise suppression and on SVMs for high-margin classification. Such systems have demonstrated strong results in ECG-based arrhythmia detection, image-based disease diagnosis, and structural health monitoring [51-52].
- **Wavelet-ANN models:** Artificial Neural Networks trained on wavelet features exhibit faster convergence, better generalization, and improved resilience to noisy inputs. These models are widely applied in weather forecasting, financial time-series prediction, and energy load estimation [53].
- In addition, researchers are exploring more advanced hybrids, such as wavelet-convolutional models and wavelet-transformer combinations, which combine the structural efficiency of wavelets with the expressive power of deep learning. These systems open new avenues for interpretable AI, energy-efficient inference, and domain-specific intelligence in fields ranging from healthcare diagnostics to smart manufacturing [54-56].

Advantages And Limitations

Advantages	Limitations
Multiscale analysis	Selection of appropriate wavelet basis is non-trivial
Time-frequency localization	Computational cost for large-scale data
Noise robustness	Less effective on highly abstract tasks without domain-specific tuning

Future Directions

As wavelet transforms continue to integrate with cutting-edge

machine learning and AI models, several exciting research directions are emerging that promise to expand the capabilities and impact of wavelets in the AI landscape:

Wavelet-transform Layers In End-to-end Trainable Models: One of the most promising areas for future research is the integration of wavelet transforms directly into deep learning architectures, allowing wavelet-based layers to be trained alongside other layers in an end-to-end fashion. This integration enables models to learn the optimal wavelet transform and filter bank for specific tasks, eliminating the need for manual feature extraction. Such models could lead to more efficient and effective learning, particularly in areas like image processing, signal denoising, and video analysis, where multiscale information is crucial. This would make wavelet transforms more adaptable to various input modalities, creating models that can learn both spatial and temporal representations simultaneously.

Integration With Transformers For Spatiotemporal Data: The popularity of transformer models in handling sequential data, such as in natural language processing (NLP) and time-series forecasting, has opened up new opportunities for integrating wavelet transforms. Transformers, with their ability to capture long-range dependencies, can benefit from wavelet-based feature extraction to better capture spatiotemporal patterns in data. For example, in climate modelling, video analysis, or traffic prediction, wavelet transforms can be used to decompose input data at different scales before being fed into transformer models, improving the model's ability to capture local and global features simultaneously. This could result in more efficient training and better performance on tasks involving complex, high-dimensional spatiotemporal data.

Wavelet-based Explainability Tools In Interpretable AI: As AI models become more complex, the need for interpretability and explainability has grown. Wavelet transforms, with their multiresolution representation, offer a natural framework for generating interpretable features. Wavelet-based explainability tools could provide insights into which frequency components or scales are most influential in a model's decision-making process. This is particularly useful in fields like healthcare, where decisions made by AI models must be understandable by human experts. By identifying the specific wavelet components that correspond to meaningful patterns in the data (such as tumour boundaries in medical imaging or heart arrhythmias in ECG signals), these tools could facilitate the development of trustworthy and transparent AI systems.

Quantum wavelet transforms for next-gen AI models: The field of quantum computing holds tremendous potential for revolutionizing AI, particularly in handling large datasets and performing computations more efficiently. Wavelet transforms, with their ability to capture multiscale information and compress data, could play a critical role in the development of quantum machine learning algorithms. Quantum wavelet transforms could leverage quantum computing's parallelism and speed to perform real-time wavelet decompositions on massive datasets. This could enable next-generation AI models to handle previously intractable problems in areas like genomics, drug discovery, and real-time language translation. Additionally, quantum wavelets may offer novel approaches to quantum signal processing, providing new insights into both classical and quantum data analysis.

CONCLUSION

Wavelets are a powerful tool for enhancing AI models, especially in tasks involving structured signals, time-series data, and complex patterns. Their ability to decompose data at multiple scales enables efficient feature extraction, denoising, and compression, making them invaluable in applications like image recognition, biomedical analysis, and speech processing. As AI evolves towards more efficient and interpretable systems, wavelets will continue to be a key component, driving improvements in model performance and robustness. Their integration with deep learning, evolutionary algorithms, and quantum computing promises to further expand their potential in tackling complex, real-world problems. Ultimately, wavelets will remain essential in advancing the capabilities and interpretability of AI technologies.

REFERENCES

- Chui, C. K. (1992). *An Introduction to wavelets*. Academic Press, San Diego.
- Chui, C. K. (1997). *A mathematical tool for signal analysis*. Philadelphia: SIAM.
- Daubechies, I. (1992). *Ten Lectures on wavelets*, CBMS-NSF, SIAM, Philadelphia.
- Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic Press.

- Addison, P. S. (2002). *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*. CRC Press.
- Unser, M. (1995). "Texture classification and segmentation using wavelet frames." *IEEE Transactions on Image Processing*, 4(11), 1549–1560.
- Tewfik, A. H., & Kim, M. (1992). "Wavelet transform as a tool for speech recognition." *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 123–126.
- Subasi, A. (2005). "EEG signal classification using wavelet feature extraction and a mixture of expert model." *Expert Systems with Applications*, 29(4), 713–720.
- Addison, P. S., Watson, J. N., Clegg, G. R., et al. (2005). "Evaluating arrhythmias in ECG signals using wavelet transforms." *IEEE Engineering in Medicine and Biology Magazine*, 24(1), 104–109.
- Guyon, I., & Elisseeff, A. (2003). "An introduction to variable and feature selection." *Journal of Machine Learning Research*, 3, 1157–1182.
- Donoho, D. L. (1995). "De-noising by soft-thresholding." *IEEE Transactions on Information Theory*, 41(3), 613–627.
- Donoho, D. L., & Johnstone, I. M. (1995). "Adapting to unknown smoothness via wavelet shrinkage." *Journal of the American Statistical Association*, 90(432), 1200–1224.
- Subasi, A. (2007). "EEG signal classification using wavelet feature extraction and a hybrid model." *Computer Methods and Programs in Biomedicine*, 87(2), 112–121.
- Addison, P. S., Watson, J. N., Clegg, G. R., et al. (2005). "Evaluating arrhythmias in ECG signals using wavelet transforms." *IEEE Engineering in Medicine and Biology Magazine*, 24(1), 104–109.
- Starck, J. L., Murtagh, F., & Fadili, J. M. (2010). *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*. Cambridge University Press.
- Gonzalez-Audicana, M., Saleta, J. L., Catalan, R. G., & Garcia, R. (2004). "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition." *IEEE Transactions on Geoscience and Remote Sensing*, 42(6), 1291–1299.
- Tewfik, A. H., & Kim, M. (1992). "Wavelet transform as a tool for speech recognition and denoising." *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 123–126.
- Ko, H., & Lee, C. (1999). "Speech enhancement using wavelet packet decomposition and soft-thresholding." *Signal Processing*, 79(1), 45–55.
- Portilla, J., Strela, V., Wainwright, M. J., & Simoncelli, E. P. (2003). "Image denoising using scale mixtures of Gaussians in the wavelet domain." *IEEE Transactions on Image Processing*, 12(11), 1338–1351.
- Buades, A., Coll, B., & Morel, J. M. (2005). "A non-local algorithm for image denoising." *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2, 60–65.
- Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian Processes for Machine Learning*. MIT Press.
- Fujieda, S., Takayama, K., & Hasegawa, T. (2018). "Wavelet Convolutional Neural Networks for Texture Classification." 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 988–997.
- Fujieda, S., Takayama, K., & Hasegawa, T. (2018). "Wavelet Convolutional Neural Networks for Texture Classification." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 988–997.
- Williams, J., Li, X., Fang, C., & Do, M. N. (2018). "Wavelet Pooling for Convolutional Neural Networks." *International Conference on Learning Representations (ICLR)*.
- Bruna, J., & Mallat, S. (2013). "Invariant Scattering Convolution Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1872–1886.
- Andén, J., & Mallat, S. (2014). "Deep Scattering Spectrum." *IEEE Transactions on Signal Processing*, 62(16), 4114–4128.
- Oyallon, E., & Mallat, S. (2015). "Deep Roto-Translation Scattering for Object Classification." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2865–2873.
- Liu, Y., Zhang, X., Zhang, S., & Li, Q. (2020). "Wavelet-based convolutional neural network for image denoising." *Neurocomputing*, 383, 47–58.
- Wu, Y., & Huang, Y. (2022). "Wavelet Attention: A New Attention Mechanism Using Wavelet Transform." *Pattern Recognition Letters*, 159, 33–40.
- Zhou, Y., Bai, J., Zhang, Q., & Ding, G. (2021). "Waveletformer: Multi-resolution time-series forecasting using wavelets and transformers." *Knowledge-Based Systems*, 212, 106548.
- Shao, Z., & Cai, J. (2021). "Wavelet-based multi-resolution deep learning for time-series forecasting." *Knowledge-Based Systems*, 212, 106548.
- Xu, X., Hu, Z., Zhang, H., & Wang, H. (2019). "Multi-resolution convolutional neural networks for EEG-based emotion recognition." *IEEE Transactions on Affective Computing*.
- Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic Press.
- Strang, G., & Nguyen, T. (1996). *Wavelets and Filter Banks*. Wellesley-Cambridge Press.
- Sifuzzaman, M., Islam, M. R., & Ali, M. Z. (2009). "Application of Wavelet Transform and Its Advantages Compared to Fourier Transform." *Journal of Physical Sciences*, 13, 121–134.
- Chakraborty, S., & Ghosh, S. (2019). "Wavelet-based convolutional neural network for image classification." *Neural Networks*, 118, 208–218.
- Wang, Z., Liu, B., & Liu, Y. (2020). "Energy-efficient neural network inference for edge computing with wavelet-based compression." *IEEE Internet of Things Journal*, 7(3), 2194–2205.
- Gandhi, V., Wu, Z., & Kim, J. (2021). "Wavelet transform for model compression and adversarial robustness in neural networks." *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(6), 4834–4842.
- Motamedi, A., & Ko, A. J. (2019). "Parameter pruning using wavelet transforms for deep neural networks."
- Han, S., Mao, H., & Dally, W. J. (2016). "Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding." *International Conference on Learning Representations (ICLR)*.
- Chen, T., et al. (2016). "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks." *IEEE Journal of Solid-State Circuits*, 52(1), 127–138.
- Patil, D., & Panigrahi, B. K. (2020). "Wavelet-based feature compression and model optimization for wearable health monitoring systems." *IEEE Transactions on Instrumentation and Measurement*, 69(10), 7927–7936.
- Hou, Y., Bai, J., Zhang, Q., & Ding, G. (2021). "Waveletformer: Multi-resolution time-series forecasting using wavelets and transformers."
- Liu, Y., et al. (2021). "Towards compact and efficient deep learning models via wavelet domain sparsity." *Pattern Recognition*, 113, 107813.
- Li, C., & He, D. (2012). "Rotating machinery fault diagnosis using adaptive wavelet packets and support vector machines." *Mechanical Systems and Signal Processing*, 25(5), 1557–1569.
- Subasi, A. (2007). "EEG signal classification using wavelet feature extraction and a mixture of expert model." *Computer Methods and Programs in Biomedicine*, 87(2), 113–123.
- Samarasinghe, T., & Alahakoon, D. (2013). "Machine learning for anomaly detection in

- structural health monitoring: A wavelet-based approach." *Structural Control and Health Monitoring*, 20(7), 1143–1156.
48. Sahin, Y., & Duman, E. (2011). "Detecting credit card fraud by decision trees and support vector machines." *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS)*, 1, 442–447.
49. Najafi, M., Ghaemi, M., & Soltani, R. (2015). "A hybrid wavelet-GA model for speech emotion recognition." *Expert Systems with Applications*, 42(14), 6235–6247.
50. Kavitha, M., & Ramakrishnan, S. (2013). "Biometric authentication system using wavelet-based fusion and genetic algorithm." *Procedia Computer Science*, 19, 692–699.
51. Zhou, J., Huang, H. Z., & Xu, H. (2015). "ECG classification using wavelet packet entropy and SVM." *Expert Systems with Applications*, 38(7), 6579–6585.
52. Singh, B. N., & Tiwari, A. K. (2006). "Optimal selection of wavelet basis function applied to ECG signal denoising using soft computing." *Applied Soft Computing*, 9(1), 265–276.
53. Yu, L., Wang, S., & Lai, K. K. (2008). "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm." *Energy Economics*, 30(5), 2623–2635.
54. Fujieda, S., Takayama, K., & Hasegawa, T. (2018). "Wavelet Convolutional Neural Networks for Texture Classification." *CVPR*, 988–997.
55. Zhou, Y., Bai, J., Zhang, Q., & Ding, G. (2021). "Waveletformer: Multi-resolution time-series forecasting using wavelets and transformers."
56. Chakraborty, S., & Ghosh, S. (2019). "Wavelet-based convolutional neural network for image classification." *Neural Networks*, 118, 208–218.