DOI: https://doi.org/10.26565/2304-6201-2025-66-07

УДК (UDC) 004.852:519.87:61

Sudakov Dmytro
Master student,
V. N. Karazin Kharkiv National University, Svobody Square, 4, Kharkiv, Ukraine, 61077
e-mail: demorsud@gmail.com
https://orcid.org/0009-0003-0060-7451

Shmatkov Segii
Professor;
V. N. Karazin Kharkiv National University, Svobody Square, 4, Kharkiv, Ukraine, 61077
e-mail: s.shmatkov@karazin.ua;
https://orcid.org/0000-0002-0298-7174

Using fractal analysis in neural network optimization algorithms in medical diagnostics

Relevance. The development of optimization methods for neural networks in medical tasks is limited by data noisiness and imbalance, which complicates the application of classical algorithms. The use of fractal analysis makes it possible to create new approaches for improving the robustness, stability, and accuracy of models.
Goal. To improve the convergence and stability of training deep neural networks in medical diagnostics through a new optimization algorithm based on fractal self-similarity.
Methods. The proposed algorithm extends the Adam by introducing fractal modulation of gradient moments through multiscale averaging. Two temporal moments are maintained: a short-term component reflecting local gradient trends and a long-term component that accumulates fractal-smoothed information over multiple scales. The update rule incorporates a fractal coefficient which controls the balance between local adaptability and global stability. This design allows the optimizer to perform gradient corrections in a self-similar manner, analogous to fractional-order dynamics.
Results. Experimental results showed that the FractalMomentAdam optimizer achieves superior performance across several key metrics. The algorithm reached a validation accuracy of 96.44%, exceeding the baseline Adam by 2.5%, while also demonstrating smoother convergence and reduced loss oscillations between epochs. The multiscale fractal smoothing contributed to better noise resistance and more stable training dynamics in the presence of data imbalance. The combination of adaptive moment estimation and fractal modulation effectively enhanced both convergence speed and final model quality.
Conclusions. The research confirms that the fractal approach to optimization provides a robust and efficient alternative to traditional gradient-based methods. By incorporating self-similar structures into moment estimation, FractalMomentAdam enhances the stability, reliability, and adaptability of neural network training in medical tasks. These findings open prospects for further research in the field of adaptive fractal optimizers, including dynamic parameter tuning, hybridization with metaheuristic strategies, and application to broader classes of medical datasets.

Keywords: *fractal analysis, neural networks, medical diagnostics, optimization algorithm, machine learning*.

Як цитувати: Sudakov D. and Shmatkov S. Using fractal analysis in neural network optimization algorithms in medical diagnostics. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2025. вип. 66. С.73-80. https://doi.org/10.26565/2304-6201-2025-66-07

How to quote: D. Sudakov and S. Shmatkov, “Using fractal analysis in neural network optimization algorithms in medical diagnostics”, *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 66, pp. 73-80, 2025. https://doi.org/10.26565/2304-6201-2025-66-07

1. Introduction

The rapid development of modern neural networks has profoundly influenced virtually all domains of human activity, with medicine being one of the most critical areas of machine learning applications. Deep learning methods are currently employed to process X-ray images, MRI results, streaming data of intracranial pressure and heart rate, historical patient records, and many other types of medical data. Contemporary neural networks have already reached a level at which they are capable of making simple diagnoses autonomously, thereby opening up enormous prospects for enhancing the efficiency and accessibility of healthcare delivery.

© Судаков Д. Г., Шматков С. І., 2025
This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0.

However, the large-scale integration of these technologies into clinical practice remains constrained by several fundamental challenges specific to medical data. Most current neural network architectures are unable to achieve the level of accuracy required for autonomous operation, being limited by computational cost and, most critically, by the acute shortage of high-quality annotated training data. Medical datasets are typically characterized by high heterogeneity and noise. In X-ray imaging, for instance, common issues include film artifacts, variability in tonal range and contrast, all of which contribute to gradient non-uniformity and introduce natural noise that complicates the learning process [1]. A major obstacle lies in the insufficient handling of medical noise sources, which often exhibit complex correlation structures and cannot be adequately captured by standard statistical models. Moreover, class imbalance is almost always present: in rare disease detection tasks, positive samples may account for less than 1% of the dataset. Additional critical challenges include the limited availability of expert-labeled data – annotation requiring the involvement of highly skilled and expensive medical specialists and substantial inter-patient variability, where identical pathologies may present differently depending on patient sex, age, or ethnic background.

In the context of such a deficit of high-quality data, optimization algorithms play a pivotal role in the training process, as they are responsible for adjusting the weight coefficients of the network. Over the past decades, a wide range of optimization methods has been developed, ranging from the classical Stochastic Gradient Descent (SGD) to more sophisticated adaptive schemes. Currently, the most widely used optimization algorithm is Adam, which combines the principles of momentum and RMSProp by maintaining separate exponentially smoothed estimates of the first and second moments of the gradient [2]. Due to its versatility, Adam has become the de facto standard in many practical applications.

Nevertheless, it is not without shortcomings: Adam is prone to stagnation in flat regions of the loss landscape (plateaus) and, in certain scenarios, may exhibit inferior generalization performance compared to classical SGD [3]. To overcome these limitations, a substantial number of Adam variants have been proposed, such as AMSGrad, which seeks to ensure monotonic decrease of the loss function, or AdamW, which decouples weight decay from gradient-based parameter updates.

Despite this wide variety of existing solutions, the choice of an optimizer remains largely empirical. None of the currently available algorithms provides consistently superior performance across all tasks. This limitation becomes particularly evident when working with noisy and imbalanced medical datasets, where classical adaptive approaches often exhibit training degradation, poor generalization, and a tendency toward overfitting [4]. Most widely used optimizers, including Adam, demonstrate limited robustness to noise, local gradient minima, and flat regions of the loss landscape – constraints that are especially critical in medical applications, where optimization errors may ultimately lead to incorrect diagnostic outcomes.

2. Review of contemporary approaches to gradient noise mitigation.

The optimization of neural network training on noisy data remains an active area of research, particularly in the domain of medical applications, where errors may have direct clinical consequences. In recent years, several promising directions have emerged that aim to overcome the limitations of classical optimization algorithms. One of the most common strategies involves modifying standard optimizers to improve their performance under noisy conditions [5]. For instance, adaptive methods such as Adam often exhibit instability on noisy datasets due to their sensitivity to gradient fluctuations. In response, researchers have proposed various modifications of the base Adam algorithm [6]. Among them is AdaBound, which combines the advantages of adaptive methods with constraints on the learning rate, thereby preventing excessive parameter oscillations during later stages of training.

An alternative line of research focuses on gradient smoothing techniques. Methods such as Stochastic Weight Averaging (SWA) or the application of moving averages to model parameters have shown promising results when dealing with noisy data. Particularly noteworthy are hybrid approaches that integrate smoothing techniques with loss landscape geometry analysis, enabling the adaptive adjustment of the smoothing level in accordance with the intensity of noise present in the data.

The challenges of data heterogeneity and noise are also being addressed beyond optimization algorithms themselves, through modifications of neural network architectures and data preprocessing strategies. Among these, Focal Loss stands out as it reduces the weight of easy examples and focuses on harder or rarer ones, which is particularly important when working with imbalanced medical datasets. The Dropout technique involves the random “deactivation” of neurons during training, thereby reducing dependence on specific input features and lowering the risk of overfitting. Label Smoothing, where probability distributions are used instead of hard labels, helps mitigate the adverse impact of incorrectly annotated data [7]. Another effective approach is Deep Ensembles, in which multiple models are trained independently and their predictions averaged, allowing for the estimation of model uncertainty.

Nevertheless, despite considerable progress, most contemporary methods still exhibit significant limitations when applied to real-world medical data [8]. These constraints underscore the necessity of developing new, specialized optimization techniques that account not only for the statistical properties of noise but also for the structural characteristics of medical data. There is a growing need for novel, more adaptive optimization strategies capable of operating effectively under conditions of uncertainty, noise, and limited sample sizes. A particularly promising research direction involves the development of hybrid approaches that combine the strengths of adaptive optimizers with methods of nonlinear dynamics, specifically the use of fractal analysis to improve convergence on heterogeneous data.

3. Implementation of the Fractal Optimization Algorithm.

The theoretical foundation for the development of the new optimization algorithm lies in fractal analysis, which provides a unique opportunity to simultaneously capture the structure of the loss function across multiple scales – from global patterns to local details. Fractals, as mathematical objects, are characterized by the property of self-similarity, meaning that their structure repeats itself at different scales [9]. This property carries profound practical significance in the natural sciences, biology, physics, and signal processing. In the context of optimization algorithms, it enables the detection of patterns and structures hidden within complex dependencies.

A large proportion of medical data processed by neural networks exhibits fractal characteristics. For example, EEG and ECG signals contain patterns recurring across different temporal scales, where fast oscillations may correspond to instantaneous changes (such as motion artifacts), while slower ones reflect long-term physiological states. Medical images (MRI, CT) often display fractal-like textures due to the hierarchical organization of tissues – for instance, vascular networks replicate their structure across multiple scales. When loss function gradients are computed on such data, they inherently inherit this multiscale nature, creating the necessity for a multiscale approach to gradient processing.

Traditional optimizers such as Adam or RMSProp rely on exponential smoothing of gradients and their squared values, implicitly assuming a single dominant temporal scale for weight update dynamics. While this approach can be effective for tasks involving regular or well-structured data, it proves less robust when the data vary simultaneously across multiple temporal or spatial scales – a situation commonly encountered in real-world medical datasets. In contrast, fractal-based methods operate not on a single scale but across a range of scales, thereby enabling the optimization process to account for both short-term fluctuations and long-term trends in the gradients.

3.1 Mathematical description of the algorithm.

An original mathematical formalization of the fractal approach was developed on the basis of the Adam algorithm, reinterpreting its operational principle. Instead of the traditional use of a single pair of exponentially smoothed moments, we introduce multiple moments with distinct smoothing scales. In its classical form, Adam maintains two exponentially smoothed moments – the mean of the gradients m_t and the mean of the squared gradients v_t, updated according to formulas 3.1 and 3.2, where g_t is the gradient at iteration *t*, and $\beta_1, \beta_2 \in [0,1)$ are smoothing coefficients.

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \quad (3.1)$$
$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (3.2)$$

Although this formulation ensures stable and rapid convergence in many tasks, it exhibits limitations when dealing with data characterized by multiscale properties or high variability (for instance, medical signals, clinical records, or pathological time series). A single-scale exponential filter cannot simultaneously capture both short-term and long-term trends.

To address this issue, the algorithm was extended through the introduction of two pairs of moments: fast moments (updated using the classical coefficients β_1 and β_2, as in Adam) and slow moments (updated with larger coefficients β_{1s} and β_{2s}, corresponding to a longer temporal horizon).

The resulting algorithm, FractalMomentAdam, extends standard Adam by incorporating multiscale smoothing, which reflects the fundamental principle of fractals – self-similarity across scales. Instead of a single pair of moments (m_t, v_t), the algorithm employs a collection of moments with different smoothing factors.

Steps:

Update of fast moments (fast scale):

$$m_t^{(f)} = \beta_1 \cdot m_{t-1}^{(f)} + (1-\beta_1) \cdot g_t \tag{3.3}$$
$$v_t^{(f)} = \beta_2 \cdot v_{t-1}^{(f)} + (1-\beta_2) \cdot g_t^2 \tag{3.4}$$

Update of slow moments (slow scale):

$$m_t^{(s)} = \beta_{1s} \cdot m_{t-1}^{(s)} + (1-\beta_{1s}) \cdot g_t \tag{3.5}$$
$$v_t^{(s)} = \beta_{2s} \cdot v_{t-1}^{(s)} + (1-\beta_{2s}) \cdot g_t^2 \tag{3.6}$$

Fractal combination (effective moments):

$$m_t^{(eff)} = \gamma \cdot m_t^{(f)} + (1-\gamma) \cdot m_t^{(s)} \tag{3.7}$$
$$v_t^{(eff)} = \gamma \cdot v_t^{(f)} + (1-\gamma) \cdot v_t^{(s)} \tag{3.8}$$

The key advantage of the fractal approach lies in its universality. Unlike methods that require manual tuning of learning rates or adaptation schedules, FractalMomentAdam inherently accounts for multiple data scales. This feature is particularly important in tasks involving non-stationary data (e.g., time series with changing statistical properties) or gradients with heavy tails – large, rare variations frequently observed in medical data, especially in the presence of rare pathologies or in multi-class classification problems.

3.2 Software implementation.

The FractalMomentAdam algorithm was implemented as a programmatic class within the TensorFlow/Keras framework. The class FractalMomentAdam was derived from the base class tf.keras.optimizers.Optimizer, which ensured full compatibility with the existing API and facilitated ease of use. The class architecture provides initialization with a parameter set that includes standard configurations such as the learning rate (learning_rate), as well as unique parameters specific to the fractal approach: coefficients for fast moments (beta_1, beta_2), coefficients for slow moments (beta_1_slow, beta_2_slow), and the scale-mixing parameter (gamma). All moments are stored as slots for each model parameter, as required by the TensorFlow API, ensuring correct functionality of optimizer state-saving and restoration mechanisms.

A critical stage in the optimizer's workflow is the update_step method, which implements the parameter update logic. At each iteration, this method computes the gradients and subsequently updates two pairs of moments separately: fast moments (m_fast, v_fast) which respond more sensitively to local variations, and slow moments (m_slow, v_slow) which integrate information over a longer horizon. The next step is the mixing procedure, where effective moments are calculated as a weighted sum of fast and slow components using the parameter gamma. Finally, analogous to the classical Adam algorithm, bias correction is applied to obtain unbiased estimates of the effective moments (m_eff_hat, v_eff_hat).

4. Comparative experimental analysis of optimizer efficiency on the PathMNIST medical dataset

To objectively evaluate the effectiveness of the developed optimizer, a comprehensive comparative experimental analysis was conducted on the real-world medical dataset PathMNIST from the MedMNIST collection. PathMNIST is considered one of the most complex and informative subsets of MedMNIST [10], designed for the task of multiclass classification of histopathological images. It is based on the dataset from the publication by Kather et al., 2019, "Predicting survival from colorectal cancer histology slides using deep learning" [11], which utilized histopathological data from patients diagnosed with colorectal cancer. From the original Whole-Slide Images (WSI), image patches of size 224×224 pixels were extracted, and for PathMNIST these patches were downsampled to 28×28 pixels and stored in RGB format. Each image represents a prepared tissue section stained with hematoxylin and eosin.

The dataset is characterized by the following parameters: image size of 28×28×3 (RGB), number of classes – 9, and task type – multiclass classification. The dataset contains 89,996 training samples, 7,180 test samples, and 10,004 validation samples. The tissue classes included in the dataset are: epithelial tissue, connective tissue, muscle tissue, adipose tissue, tumor tissue, necrotic tissue, lymphocytes, plasma cells, and background/other.

The rationale for choosing this dataset lies in its direct medical and clinical relevance, as the data reflect real histological slides used in oncological diagnostics – a field where automation is particularly needed due to the high workload on pathologists. A significant class imbalance creates a realistic and challenging scenario for testing optimizer robustness. Furthermore, unlike the clearly separated digits of the standard MNIST, the different tissue types exhibit subtle morphological differences, demanding high sensitivity to details from both the model and the optimizer. The presence of artifacts – such as variations in staining, inconsistencies in tissue cutting, and differing sample quality – introduces a degree of "noise" that is highly typical of real-world medical data. The small image size (28x28x3) offers the practical advantage of enabling local experimentation without the need for cloud computing, while the high quality of the annotations, with class labels established by experts based on original microscopic images, ensures scientific credibility.

The experimental environment was implemented in Python using the TensorFlow and Keras libraries. The model is a deep Convolutional Neural Network (CNN) [12] consisting of three main image-processing blocks and two fully connected layers for classification. Each block includes two convolutional layers with a 3×3 kernel, ReLU activation, L2 regularization, a BatchNormalization layer, MaxPooling (2×2), and Dropout with a progressively increasing rate. This architecture was chosen as sufficiently complex to solve the task, yet not overly cumbersome, ensuring fast experimental execution. All experiments were conducted under identical conditions: batch size of 128 and 10 training epochs. For each optimizer – including FractalMomentAdam with different γ values (0.2 and 0.6), Adam, SGD, SGD with momentum, and RMSProp – accuracy metrics and loss functions were recorded on both the training and validation sets. This allowed us to evaluate not only the final performance but also the training dynamics, stability, and generalization capability. The learning curves on the training and validation sets were also evaluated, focusing on the smoothness of error reduction and accuracy improvement, as well as the stability between epochs. This stability, quantified numerically and visualized graphically, reveals "collapses" in the model's learning process between epochs, indicating potential performance regressions during training

4.1. Training analysis

On the training set (Fig. 4.1), all optimizers demonstrated stable learning; however, the most indicative criterion was the validation accuracy (Fig. 4.2), which reflects the model's ability to generalize.

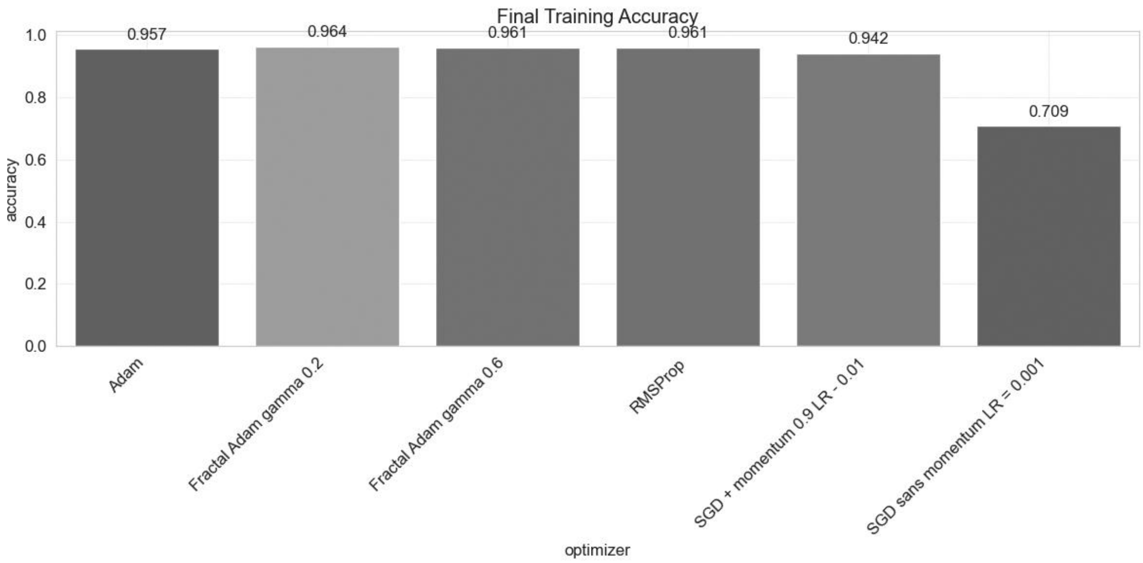

Fig. 4.1 Final training accuracy
Рис. 4.1 Кінцева тренувальна точність

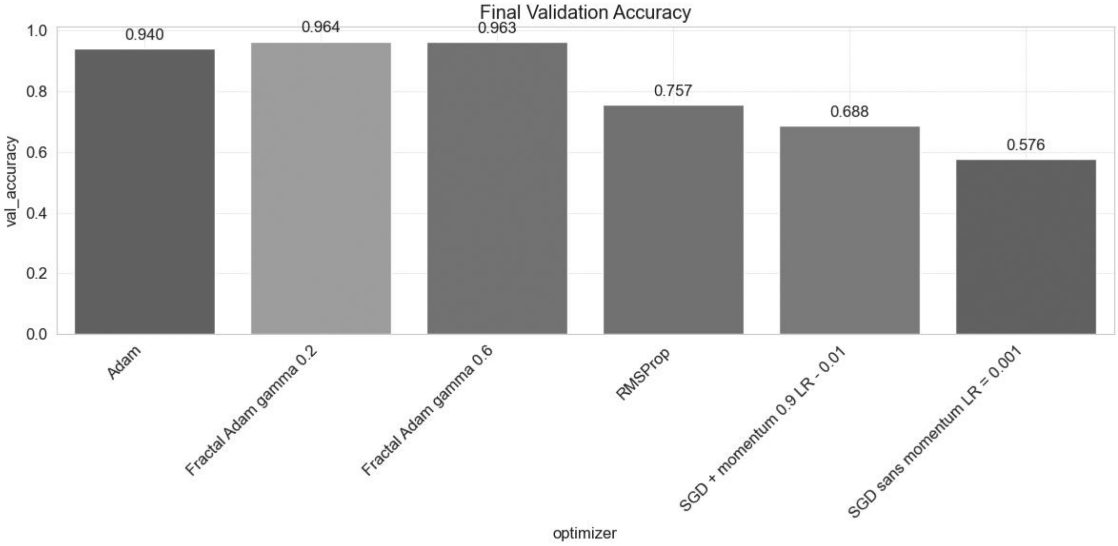

Fig. 4.2 Final validation accuracy
Рис. 4.2 Кінцева валідаційна точність

- FractalMomentAdam (γ=0.6) achieved outstanding results, reaching a peak validation accuracy of 96.27%. Its advantage was a fast start – already by the 3rd epoch the validation accuracy reached 66.49%, indicating effective exploration of the parameter. The algorithm showed stable growth without sharp fluctuations, with a deviation of ±13.8%, which highlights its robustness to noise and its ability to avoid local minima.
- Classical Adam achieved good but less stable results with a peak validation accuracy of 94.01%. It exhibited pronounced fluctuations between epochs and a slower start (59.97% at the 3rd epoch), indicating sensitivity to gradient noise and difficulties in generalization on complex data.
- FractalMomentAdam (γ=0.2) demonstrated the highest stability with a deviation of ±11.3% and a smooth increase up to a peak value of 96.44%. This version of the algorithm exhibited the lowest variance on validation – the losses decreased gradually without abrupt jumps, reflecting high generalization capability and effective regularization.
- SGD with momentum yielded the weakest results, with sharp spikes in validation accuracy. Pure SGD demonstrated the slowest learning, with a maximum validation accuracy of 69.77%, confirming its inefficiency for complex tasks with multimodal loss landscapes.
- The most problematic behavior was observed for RMSProp, which exhibited catastrophic validation collapses, with accuracy dropping to 25.76% at the 6th epoch, despite high training accuracy (96.10%). This vividly demonstrates the phenomenon of catastrophic forgetting and instability in the presence of noisy data.

The analysis of validation loss curves confirmed the advantages of the fractal approach. Both versions of FractalMomentAdam demonstrated stable and smooth loss reduction, whereas Adam exhibited oscillations over a wide range with sudden spikes in loss by several factors.The convergence speed of FractalMomentAdam was the highest among all tested algorithms. By the 5th epoch, both versions had already reached consistently high accuracy levels, while other methods required more time to stabilize.

4.2 Results analysis

The obtained results provide deeper insights into the mechanisms that make the fractal approach effective for optimization on medical data. The advantages of FractalMomentAdam can be attributed to its ability to operate simultaneously across multiple data scales – from fine-grained textural features to global variations. This property emerges from the synthesis of three key principles: adaptive moments, fractal analysis of gradients, and dynamic mixing. In practice, this manifests as smooth and predictable training trajectories, where each successive step not only minimizes current loss but also aligns with the long-term structure of the task. The parameter γ plays a crucial role in balancing local and global information. At γ=0.2 preference is given to the slow moments, resulting in more stable and cautious learning with improved generalization. This is particularly important for medical data, where emphasizing global trends over local fluctuations helps prevent overfitting to noise. At γ=0.6 the contribution of fast moments increases, leading to more aggressive exploration of the parameter space, albeit with a slight reduction in stability. The optimal value of γ=0.2 for the PathMNIST dataset indicates that, for medical images with high inter-class similarity and weakly expressed gradients, focusing on long-term trends is more effective.

The fractal approach is especially advantageous for medical data due to its inherent multiscale nature. Histopathological images in PathMNIST contain structures that recur across different scales – from individual cells to tissue complexes. FractalMomentAdam can detect these patterns through parallel analysis of short-term and long-term trends in gradients, allowing better consideration of the hierarchical nature of medical data. The results also suggest the potential for further research on adaptive selection of the γ parameter during training. Dynamically adjusting the balance between fast and slow moments depending on the training stage and data characteristics may further enhance the algorithm's efficiency for diverse medical tasks.

5. Conclusions

A novel optimization algorithm, FractalMomentAdam, has been successfully developed, implemented, and investigated, demonstrating high effectiveness for training deep neural networks on complex medical data. Experimental evaluation on the PathMNIST dataset showed that the proposed method achieves significantly higher training stability and improved accuracy compared to traditional optimizers, including Adam, SGD, and RMSProp. The best results were obtained with the parameter γ=0.2, achieving a validation accuracy of 96.44%, which is 2.5% higher than the final accuracy of the standard Adam algorithm. This configuration also exhibited minimal fluctuations between epochs and halved the training loss compared to the baseline Adam. These results were achieved by combining the adaptive principles of Adam with the concept of multiscale fractal gradient smoothing.

The proposed algorithm implements the idea of parallel utilization of two pairs of moments – fast and slow, followed by their mixing. This approach allows simultaneous consideration of both short-term and long-term trends in gradients, which is crucial for effective learning on heterogeneous data.

REFERENCES

1. Philip Ward, *MRI artifacts still require significant care and attention*. 2023. URL: https://www.auntminnieeurope.com/clinical-news/article/15657705/mri-artifacts-still-require-significant-care-and-attention (date of access: 18.06.2025).
2. Bodner B. *10 PyTorch Optimizers Everyone Is Using*. 2024. URL: https://medium.com/@benjybo7/10-pytorch-optimizers-you-must-know-c99cf3390899 (date of access: 21.05.2025).
3. Diederik P. Kingma, Jimmy Lei Ba, *ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION*. 2015.
https://doi.org/10.48550/arXiv.1412.6980 (date of access: 10.06.2025).
4. Robin M. Schmidt, Schneider F., Hennig P. *Descending through a Crowded Valley — Benchmarking Deep Learning Optimizers*. 2021. https://doi.org/10.48550/arXiv.2007.01547 (date of access: 21.05.2025).
5. GfG, *Optimization Algorithms in Machine Learning*. 2025. URL: https://www.geeksforgeeks.org/machine-learning/optimization-algorithms-in-machine-learning/ (date of access: 25.05.2025).
6. Reyad M., Amany M. Sarhan, Arafa M. *A modified Adam algorithm for deep neural network optimization*. 2023. https://doi.org/10.1007/s00521-023-08568-z (date of access: 28.05.2025).
7. G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever and R. R. Salakhutdinov *Improving neural networks by preventing co-adaptation of feature detectors*. 2012. URL: https://doi.org/10.48550/arXiv.1207.0580 (date of access: 21.05.2025).
8. Salmi M., Atif D., Oliva D., Abraham A., Ventura S. *Handling imbalanced medical datasets: review of a decade of research*. 2024. URL: https://doi.org/10.1007/s10462-024-10884-2 (date of access: 18.06.2025).
9. Barnsley, Michael F.; Rising Hawley; *Fractals Everywhere.* Boston: Academic Press Professional, 1993. ISBN 0-12-079061-0

10. Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister & Bingbing Ni. *MedMNIST v2 - A large-scale lightweight benchmark for 2D and 3D biomedical image classification.* 2023. https://doi.org/10.1038/s41597-022-01721-8 (date of access: 18.06.2025).
11. Jakob Nikolas Kather. *Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study.* 2019. URL: https://doi.org/10.1371/journal.pmed.1002730 (date of access: 18.06.2025).
12. Convolutional Neural Network (CNN) TensorFlow Guide for CIFAR. 2025. URL: https://www.tensorflow.org/tutorials/images/cnn (date of access: 25.06.2025).

Судаков Дмитро Геннадійович — *магістр, Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61077*
e-mail: *demorsud@gmail.com*
https://orcid.org/0009-0003-0060-7451

Шматков Сергій Ігорович — *д.т.н., професор кафедри комп'ютерних систем та робототехніки, науково-навчального інституту комп'ютерних систем та робототехніки; Харківський національний університет імені В. Н. Каразіна, майдан Свободи,4, Харків, Україна, 61077*
e-mail: *s.shmatkov@karazin.ua;*
https://orcid.org/0000-0002-0298-7174

Використання фрактального аналізу в алгоритмах оптимізації нейромереж у медичній діагностиці

Актуальність. Розвиток методів оптимізації нейромереж для медичних задач обмежується шумністю та дисбалансом даних, що ускладнює застосування класичних алгоритмів. Використання фрактального аналізу дозволяє створити нові підходи до підвищення стійкості, стабільності та точності моделей.
Мета. Покращити збіжність та стабільність навчання глибоких нейронних мереж у медичній діагностиці шляхом створення нового алгоритму оптимізації, заснованого на фрактальній самоподібності.
Методи. Запропонований алгоритм розширює Adam впроваджуючи фрактальну самоподібність моментів градієнта за допомогою багатомасштабного усереднення. Алгоритм використовує два часові моменти: короткострокову компоненту, що відображає локальні тенденції градієнта, та довгострокову компоненту, яка накопичує фрактально-згладжену інформацію на множині масштабів. Правило оновлення включає фрактальний коефіцієнт, що контролює баланс між локальною адаптивністю та глобальною стійкістю. Така конструкція дозволяє оптимізатору виконувати корекції градієнта самоподібним чином, аналогічно до динаміки дробового порядку.
Результати. Експериментальні результати показали, що оптимізатор FractalMomentAdam досягає вищої продуктивності за декількома ключовими метриками. Алгоритм досяг валідаційної точності 96,44%, перевищивши базовий Adam на 2,5%, а також продемонстрував більш плавну збіжність та зменшену амплітуду коливань функції втрат між епохами. Багатомасштабне фрактальне згладжування сприяло кращій стійкості до шуму та стабільнішій динаміці навчання в умовах несбалансованості даних. Комбінація адаптивної оцінки моментів та фрактальної модуляції ефективно покращила як швидкість збіжності, так і фінальну якість моделі.
Висновки. Дослідження підтверджує, що фрактальний підхід до оптимізації є надійною та ефективною альтернативою традиційним методам. Впровадження самоподібних структур в оцінку моментів дозволяє FractalMomentAdam підвищити стабільність, надійність та адаптивність навчання нейронних мереж для медичних завдань. Ці результати відкривають перспективи для подальших досліджень у галузі адаптивних фрактальних оптимізаторів, включаючи динамічне налаштування параметрів, гібридизацію з метаевристичними стратегіями та застосування для більшої кількості класів медичних датасетів.

Ключові слова: *фрактальний аналіз, нейромережі, медична діагностика, оптимізаційний алгоритм, машинне навчання.*

Надійшла у першій редакції 11.04.2025, в останній - 13.05.2025.
The first version has been received on 11.04.2025, the final version - on 13.05.2025.