DOI: https://doi.org/10.20535/kpisn.2025.4.343191
UDC 004.855.5

V.O. Nikitin[1]*, V.Ya. Danilov[1]
[1]Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
*corresponding author: nvo63911@gmail.com

DEEP LEARNING-BASED MELANOMA CLASSIFICATION ENHANCED BY FRACTAL DIMENSION ANALYSIS

Background. Melanoma is a malignant skin lesion that is prone to metastasise aggressively, leading to an almost guaranteed lethal outcome if left unchecked. In contrast, early-stage detection allows for the tumour to be removed via a harmless surgical procedure that may not even leave a scar. However, the availability of competent diagnostics are often limited due to a shortage of healthcare specialists and technologies. Deep Learning models such as Visual Transformer (ViT) have demonstrated strong performance, but researchers continuously seek to improve the results by incorporating new features. Since human skin exhibits fractal-like characteristics, it is theorised that metrics quantifying this complexity can act as valuable supplementary features for DL models, leading to increased classification accuracy.
Objective. We investigated the impact of the integration of fractal dimension (FD) on a Vision Transformer deep learning model used for melanoma classification. A comparison was conducted between the model that was exposed to random noise and the models that were provided with computed FD values.
Methods. Vision Transformer was used as a feature-extracting backbone pre-trained on the ImageNet dataset. Fine-tuning was done on this backbone in combination with a classification head targeted to distinguish melanoma vs. nevus classes. Along with extracted features, the classification head received FD value. An identical model received random noise instead of FD. Statistical testing and FD impact analysis were conducted to validate the significance of the new feature.
Results. Integrating FD into ViT showed noticeable improvement in test metrics. SHAP analysis confirmed the meaningfulness of the new feature. McNemar's test validated that the difference in model predictions was statistically significant.
Conclusions. The results suggest that FD can serve as a valuable supplementary feature for DL models, and the integration of biomarkers such as FD provides a basis for more robust melanoma classification.
Keywords: deep learning; vision transformer; melanoma; fractal dimension; XAI; skin cancer.

Introduction

Malignant skin lesions, such as squamous cell carcinoma and melanoma, can metastasise at advanced stages, significantly reducing the chances of successful treatment. Among those, melanoma is considered the most aggressive skin cancer type. Late-stage melanoma has a low likelihood of a positive outcome. In contrast, early-stage lesions can often be surgically removed with just minimal or no scarring [1]. However, the availability of competent diagnostics is often limited due to a shortage of healthcare specialists and technologies. Consequently, there is ongoing research focused on developing robust computer-aided diagnostic (CAD) systems leveraging deep learning (DL) techniques, which is being undertaken by various teams, including ISIC [2].

Data feature engineering is a significant part of any machine learning (ML) pipeline. Since human skin exhibits fractal-like characteristics, it is hypothesized that Fractal Dimension (FD) may serve as a valuable feature for enhancing DL-based skin lesion classification models [3]. FD is a metric that quantifies the complexity of fractal-like structures.

To investigate this hypothesis, we employed the Vision Transformer (ViT) [4] model as a feature extractor, as it has demonstrated strong performance in skin cancer classification tasks [5].

Пропозиція для цитування цієї статті: В.О. Нікітін, В.Я. Данилов, "Модель глибокого навчання для класифікації меланоми, покращена за допомогою фрактальної розмірності", *Наукові вісті КПІ*, № 4, с. 40–45, 2025. doi: https://doi.org/10.20535/kpisn.2025.4.343191

Offer a citation for this article: V.O. Nikitin, V.Ya. Danilov, "Deep learning-based melanoma classif ication en-hanced by fractal dimension analysis", *KPI Science News*, no. 4, pp. 40–45, 2025. doi: https://doi.org/10.20535/kpisn.2025.4.343191

© The Autor(s).
The article is distributed under the terms of the license CC BY 4.0

Contributions

We hypothesise that integrating FD as a feature to DNN Skin Cancer Classifier can improve the results. The main contributions of this study are as follows: We developed a deep neural network (DNN) classifier that combines ViT-extracted features with FD as an additional input. We conducted statistical analysis, including McNemar's test, to confirm the significance of the observed performance improvements after incorporating FD. Evaluation of the FD impact was performed using SHAP (SHapley Additive exPlanations).

Related Work

Fractal Dimension in Skin Lesion Analysis. Fractal Dimension has been explored as a quantitative metric to capture the complexity and irregularity of skin lesion boundaries. Studies have demonstrated that FD can potentially serve as a discriminative feature in differentiating between benign and malignant lesions [3]. For instance, research utilising the Higuchi's method for computing surface FD, combined with colour features, achieved classification accuracy of approximately 80 % [6]. Despite these promising results, the integration of FD into DL architectures for skin lesion classification remains fairly underexplored. We previously made efforts on the integration of FD to DL models such as Vision Transformer. The study showed that FD can positively impact a model's output [7]. Another of our studies explored approximating FD for skin lesions. In this study, the box counting dimension and its modification were compared against the Hausdorff dimension of real fractals [8].

Vision Transformers in Skin Cancer Classification. ViTs leverage self-attention mechanisms to capture global contextual information, which is particularly beneficial in analysing complex skin lesion patterns. Recent studies have demonstrated the efficacy of ViTs in skin cancer classification tasks, achieving high accuracy rates [5]. For example, a study employing a ViT model on the HAM10000 dataset reported an accuracy of 96.15 % [9].

Explainable AI in Medical Image Classification. The integration of explainable AI (XAI) techniques in medical image classification has become increasingly important to enhance model transparency and trustworthiness. SHapley Additive exPlanations (SHAP) is one such technique that provides insights into feature contributions towards model predictions. In the context of skin lesion classification, SHAP has been utilised to interpret model decisions, thereby aiding in the validation and acceptance of AI systems in clinical settings [10].

Methods

Model Selection and Study Design. The Vision Transformer architecture was selected as the base model, building upon our previous investigation into integrating FD in [7], where preliminary results indicated potential benefits. In this study, the methodology was improved through more rigorous data selection and splitting, while hyperparameters were set based on prior research. We also reduced the problem to binary classification, leaving only malignant (melanoma) and benign (nevus) classes. We compared models that receive random noise as an additional feature vs. fractal dimension. The aim was to identify if FD is useful to a DL classifier.

Data. We utilised ISIC 2019 dataset [2]. It is composed of HAM10000 (ViDIR Group, Department of Dermatology, Medical University of Vienna) [11], BCN20000 (Department of Dermatology, Hospital Clнnic de Barcelona) [12], and MSK Dataset (anonymous).

The data was reduced to only two classes: melanoma and nevus. Class imbalance was addressed by taking a number of nevi samples equal to the amount of melanoma samples. All the data was resized to 224Ч224 pixels. In order to achieve more robust training, augmentation was applied: random flipping (horizontal and vertical), random sharpness adjustment, random rotations (0°, 90°, 180°, or 270°), color jitter (brightness, contrast, and saturation varied within [0.9, 1.1]), and random resized cropping (scale: 0.9, 0.9). Training split represents the one from the ISIC 2019 Challenge [2]. Test and validation datasets were taken from ISIC 2019 Test data in a ratio of 4:1.

Fractal Dimension Calculation. We defined two approaches for calculating the fractal dimension (FD) of a lesion.

In the first approach, the FD is calculated from the curve representing the lesion's border, for which we used the box-counting dimension [13]. The lesion was preprocessed as follows (results can be seen on Fig. 1):

1. Hair was removed using the Dull Razor algorithm [14].
2. The image was converted to grayscale.
3. The image was cropped, assuming the lesions are centered.
4. The Canny edge detection algorithm [15] was applied.

5. To remove small artefacts picked up by Canny, we used openCV.findContours [16] selected the largest contour, eroded it, and applied it as a mask to Canny results.

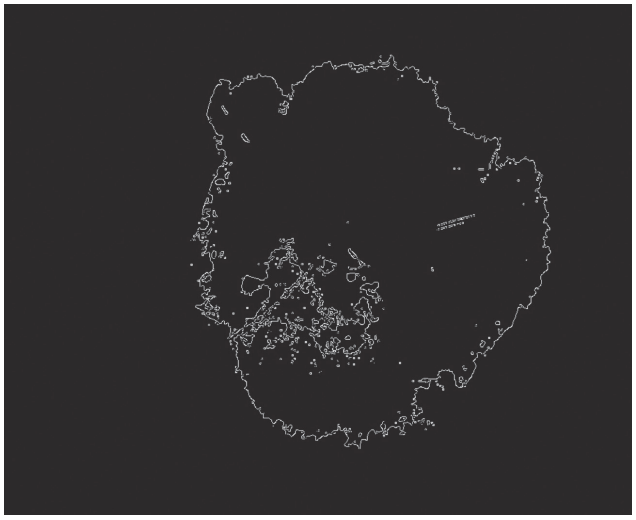

Fig. 1. Borders of a lesion

The second approach assumes the lesion is a 2D plane. We used two methods to assess the FD for this lesion representation here: a modified intensity-based dimension presented in [3], and a 2D version of the Higuchi fractal dimension [17]. Preprocessing for this approach included only cropping and grayscaling.

Model Architecture and Training. The ViT variant vit_b_32 was employed with ImageNet weights. The models were trained using the categorical cross-entropy loss function and optimised with the Adam optimiser, initialised with a learning rate of 1e-5 for classification layers and 1e-6 for fine-tuning pre-trained ViT layers. torch.optim.lr_scheduler. CosineAnnealingLR(eta_min=1e-7) was applied to mitigate overfitting. Each model was trained for a maximum of 30 epochs, with early stopping applied if the validation loss failed to improve for five consecutive epochs. FD was integrated into the model by concatenating it as an additional input to the classifier head. Prior to integration, the FD values underwent a preprocessing pipeline consisting of a linear layer (torch.nn.Linear(1, 16)). The final classifier head was structured as follows: Normalisation, Linear, GELU, Dropout, Normalisation, Linear.

Evaluation Metrics and Statistical Analysis. Model performance was assessed using accuracy, precision, recall, and F1-score. Given that the dataset, after modifications, was balanced, these metrics provided a meaningful basis for comparison. Performance was evaluated across all 3 datasets (train, validation and test). To test the significance of performance differences between the baseline and FD-enhanced models, McNemar's test – a paired chi-square test frequently employed in machine learning studies – was applied to the predictions on the test set. Finally, SHAP values were computed to quantify the contribution of FD to individual model predictions, providing insight into the actual impact of FD as an auxiliary feature.

Experiments and results

The experiment results are presented in Table 1. RN represents a model that received random noise (our control model), BC FD – box counting dimension, IFD – intensity-based box counting dimension and HFD – Higuchi dimension.

Statistical Analysis. To assess whether the performance differences between the baseline model (without FD) and the FD-enhanced model were statistically significant, McNemar's test was applied. We specifically tested BC FD as it showed the best test results. Predictions from both models were compared using predictions on the test dataset. A 2×2 contingency table was constructed, recording the number of samples where both models were correct, both were incorrect, or where one model outperformed the other. McNemar's test was performed using the standard χ^2 approximation with continuity correction applied. A significance level of 0.05 was used to determine whether the performance differences were statistically significant.

McNemar's test revealed a statistically significant difference in the performance of the two models ($\chi^2 > 3.841$, $p < 0.05$). Specifically, the FD-enhanced model made significantly different predictions compared to the baseline model. Given that the FD-enhanced models also achieved higher accuracy, F1-score, recall, and precision, the results support the positive contribution of the FD feature to improved classification performance.

SHAP Analysis. SHAP (SHapley Additive exPlanations) values were computed using GradientExplainer for all samples in the test set. The plots below summarise the impact of random noise and FD in predicting malignant classes (Fig. 2).

As expected, SHAP analysis revealed that the FD feature has a monotonic relationship between the feature value and its impact on the "Malignant" class prediction. This finding is significant as it aligns with the clinical hypothesis that malignant lesions exhibit more complex, irregular boundaries,

***Table* 1.** Performance comparison of the control model (RN) and FD-enhanced models (BC FD, IFD, HFD) across train, validation, and test sets

Dataset	Metric	RN	BC FD	IFD	HFD
Train	Precision	0.901988	0.942235	0.923132	0.876098
Train	Recall	0.859372	0.828839	0.867172	0.844662
Train	Accuracy	0.882995	0.889013	0.897482	0.862603
Train	F1-Score	0.880164	0.881907	0.894277	0.860093
Validation	Precision	0.771875	0.800000	0.810559	0.795597
Validation	Recall	0.953668	0.908367	0.970260	0.958333
Validation	Accuracy	0.838095	0.847619	0.868571	0.855238
Validation	F1-Score	0.853195	0.850746	0.883249	0.869416
Test	Precision	0.772270	0.801292	0.773220	0.768995
Test	Recall	0.947719	0.935849	0.958733	0.957020
Test	Accuracy	0.833572	0.850262	0.839771	0.835002
Test	F1-Score	0.851046	0.863359	0.856041	0.852766

which are captured by a higher fractal dimension. At the same time, analysis from the control model trained on a random noise feature demonstrates a near null impact with SHAP values randomly scattered around zero.

Results of SHAP analysis suggest that the FD feature provides a genuine and interpretable predictive signal, which the model successfully learned to exploit, whereas it correctly ignored the irrelevant control feature.

Discussion

The results of this study demonstrate that incorporating FD into a DL pipeline for skin cancer classification yields consistent improvements across key performance metrics, including accuracy, precision, recall, and F1-score. The FD-enhanced models outperformed the baseline ViT model across all evaluation datasets (train, validation, and test). These findings support the initial hypothesis that FD can serve as a valuable auxiliary feature by capturing the inherent fractal characteristics present in skin lesion morphology.

The application of McNemar's test confirmed that the observed performance improvements were statistically significant, reducing the likelihood that these gains were due to random variation. Furthermore, SHAP analysis provided insight into the role of FD within the model's decision-making process. SHAP values indicated that FD had a direct contribution to individual

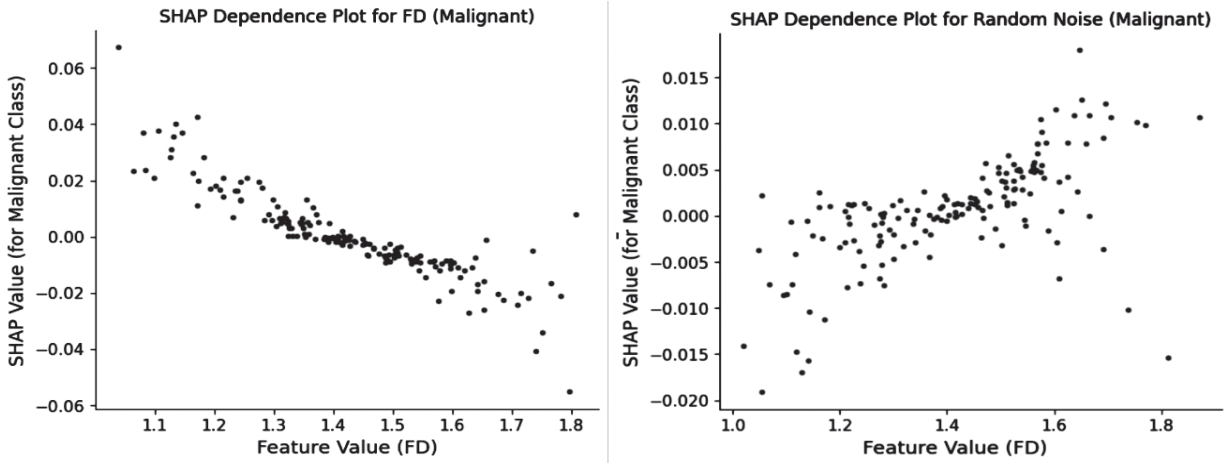

Fig. 2. SHAP Dependence plot for FD and Random Noise

predictions. This effect may promote better model generalisation.

However, limitations must be acknowledged. The study focused exclusively on the ISIC 2019dataset, which, while comprehensive, consists primarily of dermoscopic images. This reliance does not guarantee generalizability to real-world clinical settings, where images may be captured by non-dermoscopic tools. Furthermore, the task was simplified to a binary classification between melanoma and nevus, which are often visually distinct. The contribution of FD may differ in a more complex, multi-class scenario involving other, more similar-looking lesion types, especially ones in intermediate stages between benign and malignant.

Regarding the feature itself, the SHAP analysis revealed that though the absolute magnitude of the FD's contribution is often subtle, it confirmed a clear correlation. This suggests its impact could be secondary to the primary features extracted by the ViT. Our approach to calculating 2D FD on a simple crop instead of a lesion bounding box or segmentation mask impacted the accuracy of the value. It would be valuable to investigate the fusion of FD with other state-of-the-art architectures, such as traditional CNNs or Hybrid Attention CNN models. It would potentially depend on whether the received results are specific to Transformers or a more universal feature.

Nevertheless, this contribution remains significant, as we not only improved the ViT's classification metrics but also used SHAP to confirm that the fractal dimension feature was responsible for this improvement in a clinically coherent manner.

Conclusions

This study introduced and evaluated the integration of FD as an auxiliary feature in a ViT-based skin cancer classification model. Empirical results demonstrated that FD-enhanced models consistently outperformed baseline models across multiple performance metrics. Statistical testing confirmed the significance of these improvements, while explainable AI techniques (SHAP) provided additional interpretability regarding FD's role in model predictions.

The findings underscore the potential of incorporating geometric complexity measures, such as FD, to enhance DL models in medical image analysis. This approach contributes to the growing body of research on explainable and reliable AI for healthcare applications. Future work will focus on expanding the methodology to diverse datasets, exploring alternative fractal metrics, and investigating the potential of FD in other medical imaging tasks beyond skin cancer classification.

References

[1] World Cancer Research Fund International, "Skin cancer statistics," 2022. [Online]. Retrieved from: https://www.wcrf.org/cancer-trends/skin-cancer-statistics/. [Accessed: Nov. 7, 2025].

[2] International Skin Imaging Collaboration (ISIC), "ISIC Archive." [Online]. Retrieved from: https://www.isic-archive.com/. [Accessed: Nov. 7, 2025].

[3] P. Popecki *et al.*, "Fractal dimension analysis of melanocytic nevi and melanomas in normal and polarised light-A preliminary report," Life, Vol. 12, no. 7, p. 1008, 2022. doi: 10.3390/life12071008

[4] A. Dosovitskiy *et al.*, "An image is worth 16×16 words: Transformers for image recognition at scale," arXiv preprint arXiv:2010.11929, 2021. [Online]. Retrieved from: https://doi.org/10.48550/arXiv.2010.11929

[5] S. Khan *et al.*, "Identifying the role of vision transformer for skin cancer – A scoping review," *Frontiers in Artificial Intelligence*, Vol. 6, Jul. 2023, doi: 10.3389/frai.2023.1202990

[6] S. Moldovanu *et al.*, "Skin lesion classification based on surface fractal dimensions and statistical colour cluster features using an ensemble of Machine Learning Techniques," Cancers, Vol. 13, no. 21, p. 5256, 2021. doi: 10.3390/cancers13215256

[7] V. Nikitin and V. Danilov, "Integration of fractal dimension in vision transformer for skin cancer classification," *Electronics and Control Systems*, no. 76, 2023. doi: 10.18372/1990-5548.76.17662

[8] V. Nikitin and V. Danilov, "Fractal dimension calculation techniques for skin lesion characterisation," in *Матеріали конференцій МЦНД*, 2024. doi: 10.62731/mcnd-26.07.2024.005

[9] G.M. S. Himel *et al.*, "Skin cancer segmentation and classification using vision transformer for automatic analysis in dermatoscopy-based non-invasive digital system," arXiv preprint arXiv:2401.04746, 2024. [Online]. Retrieved from: https://doi.org/10.48550/arXiv.2401.04746.

[10] T. Khater *et al.*, "Skin cancer classification using explainable artificial intelligence on pre-extracted image features," *Intelligent Systems with Applications*, Vol. 21, p. 200275, Feb. 2024. doi: 10.1016/j.iswa.2023.200275

[11] P. Tschandl *et al.*, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, Vol. 5, p. 180161, 2018. doi: 10.1038/sdata.2018.161

[12] M. Combalia *et al.*, "BCN20000: Dermoscopic lesions in the wild," arXiv preprint arXiv:1908.02288, 2019. [Online]. Retrieved from: https://doi.org/10.48550/arXiv.1908.02288. [Accessed: Nov. 8, 2025].

[13] J. Wu *et al.*, "An effective method to compute the box-counting dimension based on the mathematical definition and intervals," *Results in Engineering*, Vol. 6, p. 100106, 2020. doi: 10.1016/j.rineng.2020.100106

[14] T. Lee *et al.*, "Dullrazor®: A software approach to hair removal from images," *Computers in Biology and Medicine*, Vol. 27, no. 6, pp. 533–543, Nov. 1997. doi: 10.1016/s0010-4825(97)00020-6

[15] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 8, no. 6, pp. 679–698, Nov. 1986.

[16] S. Suzuki and K. Abe, "Topological structural analysis of digitised binary images by border following," *Computer Vision, Graphics, and Image Processing*, Vol. 30, no. 1, pp. 32–46, Apr. 1985.

[17] S. Spasić, "On 2D generalisation of Higuchi's fractal dimension," *Chaos, Solitons & Fractals*, Vol. 69, pp. 179–187, Dec. 2014. doi: 10.1016/j.chaos.2014.09.015

В.О. Нікітін, В.Я. Данилов

МОДЕЛЬ ГЛИБОКОГО НАВЧАННЯ ДЛЯ КЛАСИФІКАЦІЇ МЕЛАНОМИ, ПОКРАЩЕНА ЗА ДОПОМОГОЮ ФРАКТАЛЬНОЇ РОЗМІРНОСТІ

Проблематика. Меланома – це злоякісне ураження шкіри, схильне до агресивного метастазування, що призводить до майже гарантованого летального наслідку, якщо його не лікувати. На противагу цьому, виявлення на ранній стадії дозволяє видалити пухлину, застосувавши безпечну хірургічну процедуру, яка може навіть не залишити шраму. Утім, доступність компетентної діагностики часто обмежена через нестачу медичних фахівців і технологій. Моделі глибокого навчання (ГН), такі як Visual Transformer (ViT), продемонстрували високу ефективність, але дослідники постійно прагнуть покращити результати, включаючи нові ознаки. Оскільки шкіра людини має фракталоподібні характеристики є гіпотеза, що метрики, які кількісно оцінюють цю складність, можуть слугувати цінними додатковими ознаками для моделей ГН, що підвищує точність класифікації.

Мета дослідження. Ми дослідили вплив інтеграції фрактальної розмірності (ФР) у модель глибокого навчання ViT, яку використовують для класифікації меланоми. Було проведено порівняння між моделлю, які отримувала випадковий шум, і моделями, що отримували розраховані значення ФР.

Методика реалізації. Модель ViT використовували як основу для визначення ознак для класифікації, наперед навчивши на наборі даних ImageNet. Цю основу доналаштовували (fine-tuning) в поєднанні із класифікатором (head), призначеним для розрізнення класів меланоми та невуса. Разом із вилученими ознаками класифікаційний модуль отримував значення ФР. Ідентична модель отримувала випадковий шум замість ФР. Для підтвердження значущості нової ознаки було проведено статистичне тестування та аналіз впливу ФР.

Результати дослідження. Інтеграція ФР у ViT показала помітне покращення тестових метрик. Аналіз SHAP підтвердив змістовність нової ознаки. Тест МакНемара підтвердив, що різниця у прогнозах моделі була статистично значущою.

Висновки. Результати свідчать, що ФР може слугувати цінною додатковою ознакою для моделей ГН, а інтеграція біомаркерів, таких як ФР, забезпечує основу для більш надійної класифікації меланоми.

Ключові слова: глибоке навчання; візуальний трансформер; меланома; фрактальна розмірність; пояснювальний штучний інтелект; рак шкіри.

Рекомендована Радою
Навчально-наукового інституту
прикладного системного аналізу
КПІ ім. Ігоря Сікорського

Надійшла до редакції
25 жовтня 2025 року

Прийнята до публікації
19 грудня 2025 року