

DOI: <https://doi.org/10.20535/kpissn.2025.4.343202>

UDC 004.85.032.26:004.93'1:616(045)

О.О. Зарицький^{1*}, В.Я. Данилов¹¹КПІ ім. Ігоря Сікорського, Київ, Україна

*Відповідальний автор: zaritskiy.alexey@gmail.com

МЕТОД ФРАКТАЛЬНО-КЕРОВАНОЇ РЕГУЛЯРИЗАЦІЇ АВТОЕНКОДЕРІВ ДЛЯ НАПІВКЕРОВАНОГО НАВЧАННЯ В ЗАДАЧАХ КЛАСИФІКАЦІЇ МЕДИЧНИХ ЗОБРАЖЕНЬ

Проблематика. Класифікація медичних зображень за допомогою глибокого навчання є критично важливою задачею, однак її ефективність обмежується дефіцитом розмічених даних, збір яких є дорогим. Методи напівкерованого навчання (НН) вирішують цю проблему, залучаючи нерозмічені дані. Поширені підходи, що ґрунтуються на автоенкодерах (АЕ), використовують реконструкцію як навчальний сигнал. Утім, стандартна мінімізація втрат реконструкції не гарантує, що отриманий латентний простір буде оптимально структурований для вирішення завдання класифікації, оскільки модель може фокусуватися на нерелевантних для діагностики ознаках.

Мета дослідження. Розробка та експериментальна перевірка нового методу регуляризації латентного простору — фрактально-керованої регуляризації (FDR). Мета полягає у покращенні метричних показників класифікації медичних зображень в умовах гострого дефіциту розмічених даних (5 %) за допомогою інтеграції фрактальної розмірності (ФР) як додаткового, апіорного навчального сигналу.

Методика реалізації. Запропонована модель FDR-AE ґрунтується на архітектурі АЕ, доповненій двома повнозв'язними шарами, що приєднані до латентного простору: класифікаційним та регресійним. Регресійний шар навчається прогнозувати ФР вхідного зображення, обчислену заздалегідь методом «box-counting». Загальна функція втрат є комбінацією трьох компонент: втрат класифікації на 5 % розмічених даних і втрат реконструкції та фрактальної регресії на 100 % даних. Ефективність методу перевірено на трьох наборах даних різної модальності (ISIC2024, COVID-19 Radiology, Brain Tumor MRI) порівняно з базовою згортковою мережею Base-CNN і стандартним напівкерованим АЕ SSL-AE.

Результати дослідження. Експерименти показали стабільну перевагу запропонованого методу. На датасеті ISIC2024 модель FDR-AE досягла F1-Score 0.508 для класу «malignant» проти 0.431 у SSL-AE та 0.304 у Base-CNN. На датасеті COVID-19, F1-Score для класу «covid19» склав 0.722 для FDR-AE проти 0.695 для SSL-AE. У 4-класовій задачі Brain Tumor модель FDR-AE продемонструвала покращення F1-Score щодо всіх класів, причому найбільший приріст +0.079 та +0.054, відповідно, спостерігався для класів 0 та 3, що мали найбільшу взаємну статистичну відмінність у ФР.

Висновки. Фрактально-керована регуляризація доводить, що ФР є цінним апіорним сигналом для навчання більш якісних і структурно обґрунтованих представлень у задачах НН. Метод особливо ефективний на простих архітектурах в умовах сильного дефіциту даних. Перспективи подальших досліджень включають використання FDR як методу попереднього навчання (pre-training) або впровадження динамічного коефіцієнта для регресійного компонента функції втрат.

Ключові слова: напівкероване навчання; фрактальна розмірність; автоенкодер; регуляризація латентного простору; медичні зображення; класифікація зображень; box-counting.

Вступ

Автоматизований аналіз медичних зображень на основі глибоких нейронних мереж є важливим напрямом досліджень, оскільки може

допомогти у ранній діагностиці патологій. Однією з основних проблем у впровадженні таких систем є дефіцит розмічених даних, отримання й розмітка яких вимагає висококваліфікованих спеціалістів і суттєвих витрат грошей і часу.

Пропозиція для цитування цієї статті: О.О. Зарицький, В.Я. Данилов, “Метод фрактально-керованої регуляризації автоенкодерів для напівкерованого навчання в задачах класифікації медичних зображень”, *Наукові вісті КНУ*, № 4, с. 31–39, 2025. doi: <https://doi.org/10.20535/kpissn.2025.4.343202>

Offer a citation for this article: O.O. Zarytskyi, V.Y. Danilov, “A method for fractal-driven regularization of autoencoders in semi-supervised medical image classification”, *KPI Science News*, no. 4, pp. 31–39, 2025. doi: <https://doi.org/10.20535/kpissn.2025.4.343202>

Для вирішення цієї проблеми активно розвиваються методи НН, що дають можливість використовувати одночасно розмічені й нерозмічені дані для навчання моделей.

Поширеним підходом у НН є використання АЕ, які вивчають корисні ознаки через задачу реконструкції вхідних зображень. Такий підхід довів свою ефективність у таких методах, як Adversarial Autoencoders та Ladder Networks. Утім, стандартна мінімізація втрат реконструкції не гарантує, що латентний простір моделі буде оптимально структурованим для вирішення завдання класифікації. Наше припущення ґрунтується на тому, що його можна покращити, використовуючи апіорні знання про структуру і складність даних, на яких навчається модель.

Постановка задачі

Метою цієї роботи є розробка та експериментальна перевірка нового методу НН – FDR. Ми пропонуємо використовувати ФР як додатковий навчальний сигнал для покращення якості латентного простору в задачах НН. Ефективність методу продемонстровано на трьох медичних датасетах в умовах гострого дефіциту розмічених даних (5 % вибірки), де запропонована модель АЕ із фрактальною регуляризациєю FDR-AE порівнюється з базовою керованою моделлю Base-CNN і стандартним напівкеруваним АЕ SSL-AE.

Фрактальна розмірність як класифікаційна ознака у медичних зображеннях

Виділення інформативних ознак, що описують складні біологічні структури, є ключовою проблемою під час аналізу медичних зображень. Фрактальна розмірність – це числовий показник, що оцінює структурну складність і нерегулярність об'єкта. Цей показник широко використовують як ключову ознаку в багатьох медичних дослідженнях, оскільки доведено, що ФР корелює з діагностичними станами [1–5]. Наприклад, у нейроонкології ФР корелює з показником агресивності гліобластом [3], а в мамографії доведено чітку кореляцію між зляксісністю утворення та його ФР [5]. Утім, у більшості досліджень ФР використовують як основну ознаку для аналізу. Відповідно, якщо статистичні розподіли значень ФР суттєво накладаються один на одного, – використання цього показника як основного сигналу для класифікації є неможливим. У цьому дослідженні ми пропонуємо метод використання ФР як сигналу для НН (фрактальна регуляри-

зація латентного простору мережі) і покажемо, що навіть в разі суттєвого накладання статистичних розподілів ця ознака може суттєво покращувати якість моделі.

Вибір та адаптація методу обчислення фрактальної розмірності

Є багато методів обчислення ФР зображень, що адаптовані під різні задачі та структуру самих зображень, зокрема основані на базовому методі box-counting [6–9], морфології об'єкта [10] чи навіть Фур'є-спектра [11–12]. Для цього дослідження було обрано один із найбільш поширених методів визначення ФР: box-counting. Він полягає у накладанні на зображення сіток з різними розмірами комірок r та підрахунку кількості комірок N_r , що містять пікселі вимірюваного об'єкта. Фрактальну розмірність D_{box} вимірюють як нахил прямої графіка залежності N_r від $1/r$:

$$D_{box} = \lim_{r \rightarrow 0} \frac{\log(N_r)}{\log(\frac{1}{r})}.$$

На практиці вищезгаданий нахил графіка найчастіше вимірюють за допомогою лінійної регресії [9]. Вибір цього методу був зумовлений простотою обчислення. Також було експериментально визначено, що у розглянутих нами наборах даних значення ФР, підраховані цим методом, мають найбільшу різницю у середніх значеннях для різних класів. Наприклад, для набору даних ISIC2024 (див. «Опис наборів даних») box-counting показав $\Delta Mean = 0.149$, тоді як найближчий за цим показником метод Minkovsky-Bouligand [10] показав різницю всього у 0.084.

Стандартний метод підрахунку N_r вимагає бінаризованого зображення. Це є проблемою для медичних зображень, оскільки діагностично важливою є інформація, що може міститися у відтінках сірого. Для вирішення цієї проблеми у нашій роботі ми використовуємо адаптивну бінаризацію замість фіксованого порогу. Для кожного зображення пороговим значенням буде середнє (mean) по зображенню.

Статистичне обґрунтування фрактальної регуляризації

Основною гіпотезою запропонованого методу є те, що обчислена ФР може слугувати ефективним джерелом «слабкого» сигналу для регуляризації латентного простору мережі. Для цього

потрібно, щоб середні значення ФР різнилися для різних класів. При цьому перевагою використання сигналу саме як регуляризаційного, а не як основного сигналу, для класифікації полягає в тому, що накладання розподілів ФР не є таким критичним для цієї задачі. Щоб підтвердити вищезгадану гіпотезу, ми провели статистичний аналіз на трьох обраних датасетах, щоб перевірити, чи існує статистично значуща різниця між класами. З результатів дослідження, наведених у табл. 1 та 2, можна побачити, що всі три набори даних мають статистично значущу різницю між класами (детальний опис наборів даних та обґрунтування їх вибору викладено у розд. «Матеріали та методологія досліджень»). Наявність різниці між середніми значеннями вимірів ФР ($\Delta Mean$), а також співвідношення сигнал/шум ($\Delta Mean/\max(\sigma)$), що не близькі до нуля, свідчать про релевантність застосування такої ознаки.

Регуляризація за допомогою фрактальної регуляризації

Запропонований метод полягає у структурній регуляризації латентного простору. Ми вводимо регресійний шар f_{fd} , що приєднується

до латентного вектора z і навчається прогнозувати ФР зображення, порівнюючи свої прогнози $f_{fd}(z)$ із значеннями d , визначеними апіорним методом (у нашому випадку box-counting). Формулу для обчислення втрат регресії L_{fd} для батчу розміру N описують таким чином:

$$\hat{d}_i = f_{fd}(z_i);$$

$$L_{fd} = \frac{1}{N} \sum_{i=1}^N (\hat{d}_i - d_i)^2,$$

де $\hat{d}_i$ та d_i – прогнозоване та апіорне значення ФР для i -го вектора у батчі, а z_i – латентне представлення i -го вхідного зображення.

Оскільки статистичний аналіз підтвердив, що ФР є релевантною ознакою, її явне кодування у латентному просторі може покращувати його структуру у наступній класифікації.

Щоб запобігти втраті іншої цінної інформації під час регресії ФР, ми застосуємо цей метод як модифікацію архітектури АЕ. Головне завдання АЕ – мінімізація функції втрат реконструкції L_{rec} . Вона гарантує, що вектор z зберігатиме максимум інформації для відновлення зображення. Цю регуляризаційну властивість

Таблиця 1. Статистичний аналіз фрактальної розмірності, підрахованої методом box-counting (датасети із двома класами)

Набір даних	Mean FD (class 0)	Mean FD (class 1)	$\Delta Mean$	σ (class 0)	σ (class 1)	$\Delta Mean/\max \sigma$
ISIC2024 Challenge	1.375	1.524	0.149	0.498	0.389	0.299
COVID-19 Radiology Database	1.851	1.828	0.023	0.024	0.023	0.941

Таблиця 2. Статистичний аналіз фрактальної розмірності, підрахованої методом box-counting (датасети з чотирма класами)

Набір даних	Mean FD (class 0)	Mean FD (class 1)	Mean FD (class 2)	Mean FD (class 3)
Brain Tumor MRI Dataset	1.88 518	1.88 538	1.8851	1.87 125
	σ (class 0)	σ (class 1)	σ (class 2)	σ (class 3)
	0.0 188 859	0.0 211 904	0.0 318 129	0.0 140 467

Набір даних	Клас	Клас	$\Delta Mean$	$\max \sigma$	$\Delta Mean/\max \sigma$
Brain Tumor MRI Dataset	0	1	0.0002	0.0 211 904	0.00 943 824
	0	2	0.00 008	0.0 318 129	0.0 025 147
	0	3	0.01 393	0.0 188 859	0.7 375 873
	1	2	0.00 028	0.0 318 129	0.00 880 146
	1	3	0.01 413	0.0 211 904	0.66 681 139
	2	3	0.01 385	0.0 318 129	0.43 535 798

АЕ широко використовують у НН для навчання на нерозмічених даних, наприклад в adversarial autoencoders [13] та ladder network архітектурі [14]. Таким чином, оптимізація двох вищевказаних функцій втрат діє одночасно як регуляризація повноти інформації (L_{rec}) і як структурна регуляризація (L_{fd}) у латентному просторі.

Для порівняння реконструйованого зображення з оригінальним ми використали MSE (mean squared error). Для навчання на розмічених даних було додано класифікаційний шар, що використовує cross-entropy loss функцію (L_{CE}). Загальну функцію втрат можна описати так:

$$L = L_{CE} + \lambda_{rec} L_{rec} + \lambda_{fd} L_{fd}.$$

Гіперпараметри λ_{rec} та λ_{fd} контролюють внесок некерованого навчання. Оптимізація цієї функції втрат змушує енкодер створювати простір z , що є одночасно повним, структурно чутливим та оптимізованим для класифікації. На рис. 1 зображено модель, що реалізує запропонований метод. Стрілками позначено зв'язки між компонентами моделі, а пунктирними лініями — умовні зв'язки, які демонструють, як обчислюються компоненти функції втрат.

Матеріали і методологія дослідження

Опис наборів даних

Для експериментальної демонстрації ефективності запропонованого методу було обрано три відкриті датасети медичних зображень, що різняться за типом діагностичної задачі, візуальними характеристиками, а також статис-

тичними властивостями значень ФР. У табл. 3 надано опис кожного датасету із поясненнями трансформацій та обґрунтування вибору кожного датасету.

Архітектура моделей для експериментів

Експериментальну частину було виконано у середовищі Python за допомогою бібліотеки Pytorch. Для практичної демонстрації результатів дослідження було розроблено й реалізовано три архітектури нейронних мереж, кожна з яких відіграє певну роль у дослідженні.

Базова модель — класична згорткова нейронна мережа Base-CNN, яка слугує еталонною моделлю для керованого навчання. Її архітектура складається із чотирьох послідовних шарів Conv2d→ReLU. Після згорткових шарів подання передається послідовно у три повнозв'язні шари Linear. Усі інші моделі у цьому експерименті були побудовані на основі цієї моделі з мінімальними модифікаціями, потрібними для реалізації відповідної архітектури. Цю модель навчали тільки на розміченій частині даних (5 % даних). Порівняння метрик інших моделей із базовою моделлю дають можливість відслідкувати, наскільки використання нерозмічених даних покращує класифікаційні якості моделі на тестовій вибірці, тобто наскільки запропонований метод НН ефективний.

Для порівняння запропонованого методу з уже відомими методами НН було обрано базовий АЕ як найближчу за структурою модель. Автоенкодер довели свою ефективність у цьому класі задач за рахунок їхньої здатності вивчати змі-

стовні ознакові подання (feature representations) з нерозмічених даних [19, 14]. У цій роботі модель напівкерованого АЕ (SSL-AE) включає енкодер, що повністю повторює архітектуру енкодера базової моделі, а також декодер, що дзеркально відтворює його структуру, використовуючи шари ConvTranspose2d→ReLU для реконструкції зображення. Слід зазначити, що у відкритому доступі не було знайдено публікацій, присвячених НН на обраних наборах даних з ідентичною експериментальною конфігурацією (зокрема, 5 % розмічених даних). Через це для забезпечення коректного і чесного порівняння

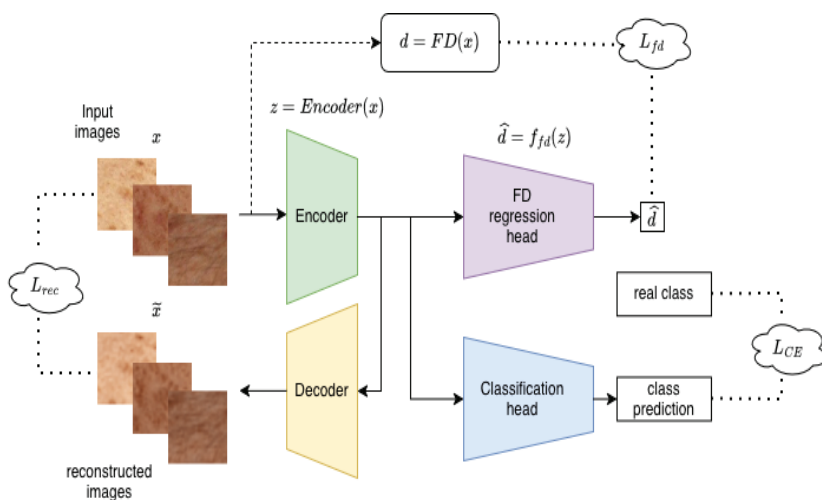


Рис. 1. Модель автоенкодера з фрактальною регуляризцією

базова модель SSL-AE була реалізована нами самостійно з тією ж базовою архітектурою, що й запропонована модель FDR-AE.

Третьою є запропонована модель AE із фрактальною регуляризацією FDR-AE. Вона розширює архітектуру SSL-AE і включає регресійну частину, що намагається на основі вихідного шару енкодера визначити фрактальну розмір-

ність початкового зображення. Регресор складається із двох послідовних повнозв'язних шарів $\text{Linear} \rightarrow \text{ReLU}$. Таким чином функція втрат формується як комбінація втрат класифікації, реконструкції та втрат регресійної частини. На рис. 2 зображено внутрішню структуру компонентів моделі, а також яким чином обчислюють різні компоненти функції втрат під час навчання

Таблиця 3. Опис наборів даних, використаних для експериментів

Набір даних	Призначення	Характеристики	Попередня обробка
ISIC2024 Challenge	Бінарна класифікація уражень на шкірі для виявлення раку	Датасет містить 400 000 зображень уражень шкіри, що були отримані з 3d-знімків усього тіла (технологія 3d TBP) [15]	Цей датасет характеризується суттєвим дисбалансом класів. Клас «benign» становить більше 99 % всіх даних, тому для ефективного навчання датасет був збалансований за рахунок видалення із навчальної вибірки більшості екземплярів. Усі зображення були нормалізовані, а також зменшені до розміру 96×96 для зниження витрат ресурсів в експерименті
Містить два класи, що позначають злоякісні та доброякісні утворення – «malignant» та «benign»	Класифікація і сегментація захворювань за радіологічними знімками легень пацієнтів	Набір даних, зібраний дослідниками з Університету Катару та Дакки [16, 17]. Датасет складається із трьох класів: 1) 11 956 знімків пацієнтів із COVID-19, 2) 11 263 знімки вірусної пневмонії (інші інфекції), 3) 10 701 знімок здорових легень. Цей датасет є важливим ресурсом для демонстрації ефективності запропонованого методу, оскільки візуальні ознаки COVID-19 можуть бути менш помітними й вимагати аналізу складних текстурних особливостей знімку	Для спрощення експериментів у межах цього дослідження було використано підмножину, що складається із класів «COVID-19 positive» та «Normal». Таким чином у нашому дослідженні ми розглядали бінарну класифікацію. Усі зображення були нормалізовані, а також зменшені до розміру 96×96 для зниження витрат ресурсів в експерименті
Brain Tumor MRI Dataset	Багатокласова класифікація пухлин мозку за МРТ-зображеннями	Цей набір даних є поєднанням трьох джерел (figshare, SPARTAJ, Br35H) і містить 7023 знімки МРТ людського мозку [18]. Набір даних розподілено на чотири класи: гліома (glioma), менингіома (meningioma), гіпофізарна пухлина (pituitary) і відсутність пухлини (no tumor). Класи подані у збалансованих пропорціях. Датасет було обрано для демонстрації застосування запропонованого методу на багатокласовому датасеті	Для спрощення експериментів у межах цього дослідження було використано підмножину, що складається із класів «COVID-19 positive» та «Normal». Таким чином у нашому дослідженні ми розглядали бінарну класифікацію. Усі зображення були нормалізовані, а також зменшені до розміру 96×96 для зниження витрат ресурсів в експерименті

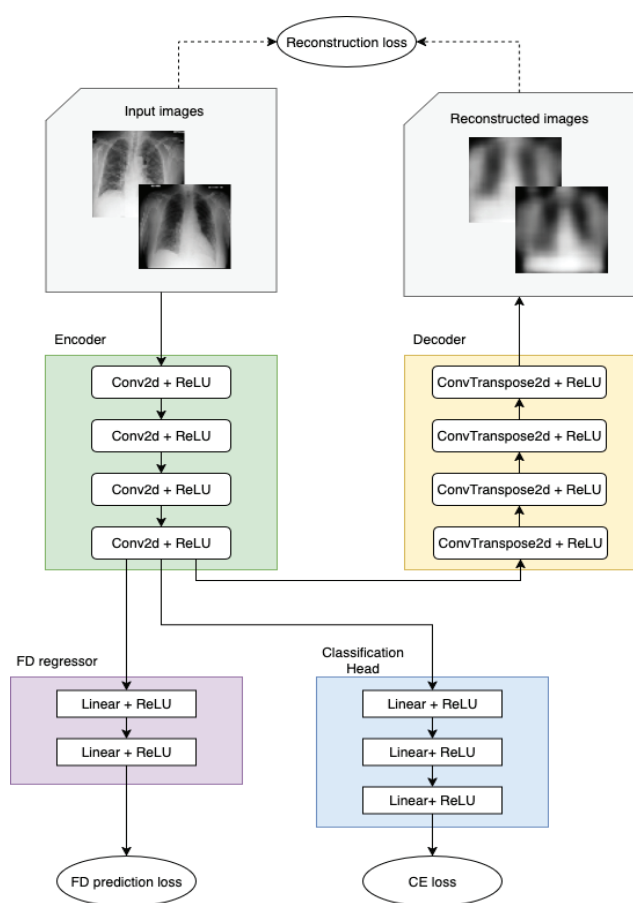


Рис. 2. Модель автоенкодера із фрактальною регуляризцією.
Внутрішня структура компонентів моделі

(компоненти функції втрат на рис. 2 обведено еліпсами).

У табл. 4, 5 та 6 наведено метричні показники запропонованих моделей для трьох вищезгаданих датасетів.

Результати експериментів (табл. 4–6) демонструють перевагу запропонованого методу на всіх трьох наборах даних. На датасеті ISIC2024 (табл. 4) FDR-AE модель досягла F1-Score 0.508 для малорепрезентованого класу «malignant», значно перевершивши як Base-CNN (0.304), так і SSL-AE (0.431). Аналогічна ситуація спостерігається також і в датасеті COVID-19 (табл. 5), де FDR-AE покращив якість F1-Score для класу «covid-19» на 0.027 та на 0.112 відносно моделей SSL-AE та Base-CNN. Особливо показовими є результати на багатокласовому датасеті Brain Tumor (табл. 6), де метод суттєво покращив F1-score для всіх класів, причому найбільший приріст відносно SSL-AE отримали класи 0 (0.079) та 3 (0.054), які мали найбільшу взаємну статистичну відмінність у ФР ($\Delta Mean/\max(\sigma) = 0.738$). Таким чином, у всіх трьох сценаріях фрактальна регуляризація виявилася ефективнішою за стандартну реконструкцію (SSL-AE), підтверджуючи, що ФР є цінним сигналом для навчання більш якісних представлень у задачах НН.

Висновки

Наявні рішення для класифікації медичних зображень в умовах дефіциту даних ґрунтуються на НН (SSL), зокрема на AE. Утім, стандартна реконструкція не гарантує створення латентного простору, оптимального для класифікації. У цій статті запропоновано новий метод – фрактально-керовану регуляризацію (FDR-AE) – для вирішення цієї проблеми. На відміну від відомих AE, наш метод вводить додатковий навчальний сигнал – регресійний шар, який навчається прогнозувати фрактальну розмір-

Таблиця 4. Метричні показники моделей, навчених на 5 % розмічених даних ISIC2024

Модель	Precision (benign)	Precision (malignant)	Recall (benign)	Recall (malignant)	F1 Score (benign)	F1 Score (malignant)
Base-CNN	0.951	0.240	0.897	0.414	0.923	0.304
SSL-AE	0.596	0.524	0.737	0.367	0.659	0.431
FDR-AE	0.615	0.517	0.632	0.500	0.623	0.508

Таблиця 5. Метричні показники моделей, навчених на 5 % розмічених даних COVID-19 Radiology Database

Модель	Precision (normal)	Precision (covid19)	Recall (normal)	Recall (covid19)	F1 Score (normal)	F1 Score (covid19)
Base-CNN	0.799	0.682	0.873	0.552	0.834	0.610
SSL-AE	0.684	0.878	0.920	0.575	0.785	0.695
FDR-AE	0.702	0.896	0.930	0.605	0.800	0.722

Таблиця 6. Метричні показники моделей, навчених на 5 % розмічених даних Brain Tumor MRI Dataset

Модель	Precision (class 0)	Precision (class 1)	Precision (class 2)	Precision (class 3)
Base-CNN	0.683	0.447	0.703	0.598
SSL-AE	0.720	0.500	0.848	0.716
FDR-AE	0.704	0.560	0.848	0.797
Модель	Recall (class 0)	Recall (class 1)	Recall (class 2)	Recall (class 3)
Base-CNN	0.460	0.359	0.773	0.833
SSL-AE	0.447	0.572	0.810	0.927
FDR-AE	0.570	0.549	0.867	0.940
Модель	F1 Score (class 0)	F1 Score (class 1)	F1 Score (class 2)	F1 Score (class 3)
Base-CNN	0.550	0.399	0.736	0.696
SSL-AE	0.551	0.534	0.828	0.808
FDR-AE	0.630	0.554	0.857	0.862

ність зображення з його латентного вектора. Перевага цього методу полягає у примусовому збереженні структурно важливої інформації.

Експерименти з трьома наборами даних (ISIC2024, COVID-19, Brain Tumor) за умови 5 % маркованих даних показали, що FDR-AE стабільно перевершує як базову повністю керовану модель (Base-CNN), так і стандартний AE (SSL-AE). Наприклад, у завданні з чотирма класами Brain Tumor (табл. 6) FDR-AE поліпшив показник F1-Score для всіх класів, причому найбільше зростання (+0,079 і +0,054) було для класів 0 і 3. Сферою застосування методу є системи комп'ютерної діагностики на основі нейронних мереж. Конкретно цей метод може бути застосований як для НН, так і для переднавчання моделей.

Також у процесі дослідження було виявлено суттєве обмеження – метод найбільш ефективний на відносно простих архітектурах та в разі сильного дефіциту даних. На складних моделях з великою кількістю даних ФР може погіршувати кінцеві показники. Це відкриває шляхи для покращення, які стануть предметом наших наступних робіт: використання FDR-AE як методу попереднього навчання (pre-training) і впровадження динамічного коефіцієнта, що автоматично зменшуватиме вплив регуляризації зі зростанням точності моделі. Практична реалізація цих підходів дозволить створити більш стабільні та ефективні моделі НН для підвищення точності автоматизованої діагностики в реальних клінічних умовах.

References

1. C.S. Fortin *et al.*, “Fractal dimension in the analysis of medical images”, *IEEE Engineering in Medicine and Biology Magazine*, Vol. 11, no. 2, pp. 65–71, 1992. Retrieved from doi: <https://doi.org/10.1109/51.139039>
2. S. Kisan *et al.*, “Fractal dimension in medical imaging: a review”, *IRJET*, Vol. 5, pp. 1102–1106, 2017. [Online]. Available: <https://www.irjet.net/archives/V4/i5/IRJET-V4I5363.pdf>. Accessed: Oct. 11, 2025.
3. S. Liu *et al.*, “Relationship between necrotic patterns in glioblastoma and patient survival: fractal dimension and lacunarity analyses using magnetic resonance imaging”, *Scientific reports*, Vol. 7, no. 1, p. 8302, 2017. Retrieved from doi: <https://doi.org/10.1038/s41598-017-08862-6>
4. M.Z.C. Azemin *et al.*, “Robust methodology for fractal analysis of the retinal vasculature”, *IEEE Trans Med Imaging*, Vol. 30, no. 2, pp. 243–250, 2011. Retrieved from doi: <https://doi.org/10.1109/TMI.2010.2076322>
5. R.M. Rangayyan and T. M. Nguyen, “Fractal analysis of contours of breast masses in mammograms”, *Journal of digital imaging*, Vol. 20, no. 3, pp. 223–237, 2007. Retrieved from doi: <https://doi.org/10.1007/s10278-006-0860-9>
6. K. Falconer, “Fractal geometry: mathematical foundations and applications. Chichester: John Wiley”, 1990, pp. 38–47. [Online]. Available: www.wiley.com/en-us/Fractal+Geometry%3A+Mathematical+Foundations+and+Applications%2C+3rd+Edition-p-9781119942399#description-section. Accessed: Oct. 11, 2025.

7. S. Buczkowski *et al.*, “Measurements of fractal dimension by box-counting: a critical analysis of data scatter”, *Physica A: Statistical Mechanics and its Applications*, Vol. 252, no. 1–2, pp. 23–34, 1998. Retrieved from doi: [https://doi.org/10.1016/S0378-4371\(97\)00581-5](https://doi.org/10.1016/S0378-4371(97)00581-5)
8. J. Li *et al.*, “An improved box-counting method for image fractal dimension estimation”, *Pattern recognition*, Vol. 42, no. 11, pp. 2460–2469, 2009. Retrieved from doi: <https://doi.org/10.1016/j.patcog.2009.03.001>
9. C. Panigrahy *et al.*, “Fractal dimension of synthesized and natural color images in lab space”, *Pattern Analysis and Applications*, Vol. 23, no. 2, pp. 819–836, 2020. Retrieved from doi: <https://doi.org/10.1007/s10044-019-00839-7>
10. R.D.O. Plotze *et al.*, “Leaf shape analysis using the multiscale Minkowski fractal dimension, a new morphometric method: a study with Passiflora (Passifloraceae)”, *Canadian Journal of Botany*, Vol. 83, no. 3, pp. 287–301, 2005. Retrieved from doi: <https://doi.org/10.1139/b05-002>
11. H. Ahammer, “Higuchi dimension of digital images”, *PLoS One*, Vol. 6, no. 9, p. e24796, 2011. Retrieved from doi: <https://doi.org/10.1371/journal.pone.0024796>
12. J.B. Florindo and O. Martinez Bruno, “Fractal descriptors in the Fourier domain applied to color texture analysis”, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Vol. 21, no. 4, 2011. Retrieved from doi: <https://doi.org/10.1063/1.3650233>
13. A. Makhzani *et al.*, “Adversarial autoencoders”, arXiv preprint arXiv:1511.05644, 2015. Retrieved from doi: <https://doi.org/10.48550/arXiv.1511.05644>
14. H. Rasmus *et al.*, “Semi-Supervised Learning with Ladder Networks”, in *NIPS'15: Proceedings of the 29th International Conference on Neural Information Processing Systems*, Vol. 2, pp. 3546–3554, 2015. Retrieved from doi: <https://doi.org/10.48550/arXiv.1507.02672>
15. N.R. Kurtansky *et al.*, “The SLICE-3D dataset: 400,000 skin lesion image crops extracted from 3D TBP for skin cancer detection,” *Scientific Data*, Vol. 11, no. 1, p. 884, 2024. Retrieved from doi: <https://doi.org/10.1038/s41597-024-03743-w>
16. M.E. Chowdhury *et al.*, “Can AI help in screening viral and COVID-19 pneumonia?”, *IEEE Access*, Vol. 8, pp. 132665–132676, 2020. Retrieved from doi: <https://doi.org/10.1109/ACCESS.2020.3010287>
17. T. Rahman *et al.*, “Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images”, *Computers in biology and medicine*, Vol. 132, p. 104319, 2021. Retrieved from doi: <https://doi.org/10.1016/j.compbiomed.2021.104319>
18. M. Nickparvar, “Brain Tumor MRI Dataset”, *Kaggle*, 2023. [Online]. Available: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>. Accessed: Oct. 11, 2025.
19. D.P. Kingma *et al.*, “Semi-supervised learning with deep generative models”, in *Advances in neural information processing systems*, Vol. 27, 2014. [Online]. Available: <https://papers.nips.cc/paper/5352-semi-supervised-learning-with-deep-generative-models>. Accessed: Oct. 11, 2025.

V.Y. Danilov, O.O. Zarytskyi

METHOD FOR FRACTAL-DRIVEN REGULARIZATION OF AUTOENCODERS FOR SEMI-SUPERVISED LEARNING IN MEDICAL IMAGE CLASSIFICATION TASKS

Background. Medical image classification using deep learning is a critical task, yet its effectiveness is constrained by the scarcity of labelled data, which is expensive to acquire. Semi-supervised learning (SSL) methods address this by leveraging unlabelled data. Common autoencoder (AE)-based approaches use reconstruction as a training signal. However, standard reconstruction loss minimisation does not guarantee that the resulting latent space will be optimally structured for the classification task, as the model may focus on diagnostically irrelevant features.

Objective. The paper aims to develop and experimentally validate a novel latent space regularisation method: fractal-driven regularisation (FDR). The goal is to improve classification metrics for medical images under conditions of severe labelled data scarcity (5 %) by integrating fractal dimension (FD) as an additional, a priori training signal.

Methods. The proposed model (FDR-AE) is based on an autoencoder architecture, augmented with two heads attached to the latent space: a classification head and a regression head. The regression head is trained to predict the input image’s FD, which is pre-calculated using the “box-counting” method. The total loss function is a combination of three components: classification loss (on 5 % labelled data) and both reconstruction and fractal regression losses (on 100 % of data). The method’s efficacy was validated on three datasets of different modalities (ISIC2024, COVID-19 Radiology, Brain Tumor MRI), comparing it against a baseline convolutional network (Base-CNN) and a standard semi-supervised autoencoder (SSL-AE).

Results. The experiments demonstrated a consistent advantage for the proposed method. On the ISIC2024 dataset, FDR-AE achieved an F1-Score of 0.508 for the “malignant” class, compared to 0.431 for SSL-AE and 0.304 for Base-CNN. On the COVID-19 dataset, the F1-Score for the “covid19” class was 0.722 for FDR-AE versus 0.695 for SSL-AE. In the 4-class Brain Tumor task, FDR-AE showed improved F1-Scores across all classes, with the most significant gains (+0.079 and +0.054) observed for classes 0 and 3, which also had the greatest mutual statistical difference in their FD values

Conclusions. Fractal-driven regularisation demonstrates that FD is a valuable a priori signal for learning higher-quality, structurally grounded representations in SSL tasks. The method is particularly effective on simple architectures under severe data scarcity. Prospects for future research include utilising FDR as a pre-training method or implementing a dynamic coefficient for the regression component of the loss function.

Keywords: semi-supervised learning; fractal dimension; autoencoder; latent space regularisation; medical images; image classification; box-counting

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

Надійшла до редакції
25 жовтня 2025 року

Прийнята до публікації
8 грудня 2025 року