

теоретичні переваги, емпіричні порівняння часто показують, що вони не завжди систематично перевершують надійні статичні факторні моделі (Stock & Watson) [4], які зазвичай реалізуються за допомогою аналізу головних компонент (PCA), з точки зору точності прогнозування.

Отримані результати вказують на можливість істотного поліпшення якості короткострокових економічних прогнозів (Nowcasting) для ВВП України та її регіонів. Моделі MF-VAR/MF-FAVAR мають фундаментальні переваги над традиційними моделями, оскільки вони здатні обробляти змішані частоти даних і великі незбалансовані набори даних за допомогою ефективного використання спільних факторів. Це особливо важливо для регіонального прогнозування, де для отримання своєчасних оцінок необхідна інтеграція неоднорідних і високочастотних показників.

Список джерел:

- [1] Koop G., McIntyre S., Mitchell J., Poon A., Wu P. Incorporating Short Data into Large Mixed-Frequency VARs for Regional Nowcasting // Federal Reserve Bank of Cleveland Working Paper Series. – 2023. – No. 23-09
- [2] Зомчак Л. М., Ракова А. С. U-MIDAS-модель прогнозування ВВП України на даних змішаної частоти // Східна Європа: Економіка, Бізнес та Управління. – 2020. – Вип. 2 (25). – С. 436–441.
- [3] Constantinescu M., Karpner K., Szumilo N. The Warcast Index: Estimating Economic Activity without Official Data during the Ukraine War in 2022 // NBU Working Papers. – 2024. – № 03/2024.
- [4] Stock J. H., Watson M. W. Forecasting Using Principal Components From a Large Number of Predictors // Journal of the American Statistical Association. – 2002. – Vol. 97, № 460. – P. 1167–1179.

**ФРАКТАЛЬНО-ДИFUЗІЙНІ ГЕНЕРАТИВНІ МОДЕЛІ: ІЄРАРХІЧНИЙ ПІДХІД ДО СИНТЕЗУ ЗОБРАЖЕНЬ / FRACTAL-DIFFUSION GENERATIVE MODELS: A HIERARCHICAL APPROACH TO IMAGE SYNTHESIS**

*Шалімов А. В., Авраменко О. В. / Shalimov A., Avramenko O.*

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики  
E-mail: a.shalimov@ukma.edu.ua, o.avramenko@ukma.edu.ua

This study extends the Fractal Generative Model by replacing its autoregressive generator with a diffusion-based generator at the top level of the hierarchy. FractalGen builds images recursively: each level invokes a similar generative module on subregions, enforcing self-similar structure across scales. In the original model, this module is autoregressive, generating pixels sequentially and enabling exact likelihoods. We introduce a new generator type that uses a conditional diffusion U-Net to produce a coarse global canvas in parallel via denoising. Lower levels of the hierarchy remain unchanged and further refine local detail. This hybrid fractal-diffusion architecture preserves the recursive, self-similar assembly of FractalGen while replacing pixel-by-pixel sampling with parallel diffusion at the latent level. We evaluate a baseline DDPM against our approach on ImageNet using visual-fidelity and distributional metrics.

Сучасні методи генерації зображень ґрунтуються на масштабних авторегресійних і дифузійних моделях, здатних створювати реалістичні та семантично узгоджені зображення. Дифузійні підходи [1] формують зображення шляхом ітеративного знешумлення латентного вектора, забезпечуючи високу якість, але потребуючи значних обчислень. Авторегресійні моделі факторизують розподіл за регіонами, що забезпечує узгодженість між підобластями, проте залишаються повільними через послідовну природу генерації. Для подолання цих обмежень запропоновано фрактальні генеративні моделі (Fractal Generative Models, FGM) [2], які рекурсивно застосовують один і той самий генератор на різних масштабах, формуючи самоподібну ієрархію, що концептуально узгоджується з фрактальною природою [3]. У початковій реалізації FGM використовували авторегресійні модулі, які забезпечували узгодженість між підобластями, але зберігали низьку швидкість через послідовну генерацію.

У цьому дослідженні пропонується модифікація FGM із використанням дифузійної нейронної мережі на архітектурі U-Net як альтернативного генератора. Такий підхід підвищує паралельність процесу, скорочує час синтезу й інтегрує ієрархічний самоподібний принцип у стандартну дифузійну модель. Конфігурація працює в латентному просторі [4], що зменшує ймовірність артефактів, спричинених відсутністю явного авторегресійного обумовлення. Верхній рівень ініціалізує гаусівський шум, який знешумлюють дочірні вузли, що діють на

маскованих підобластях із контекстним урахуванням глобального рівня [5]. Узгодженість між підобластями забезпечується спільною ініціалізацією шуму та контекстним навчанням.

Порівняння роботи моделі проводимо з використанням датасетів MNIST та ImageNet у роздільності  $64 \times 64$ . Для оцінки якості генерації застосовуємо метрику FID (Fréchet Inception Distance), яка відповідає за візуальну якість зображень. Також виконано експерименти з варіюванням глибини фрактала, що відповідає за збільшення рівнів ієрархії та кількості дочірніх вузлів.

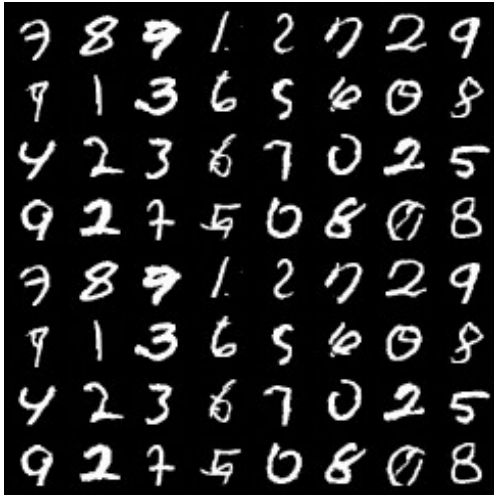


Рис. 1 Згенеровані зображення при  $N = 50$ .



Рис. 2 Згенеровані зображення при  $N=20$ .

Під час експериментів метрику FID оцінювали для різних глибин фрактала  $N \in \{20, 30, 50, 80, 100, 150\}$ . Результати показали немонотонну залежність якості від глибини: FID зменшується від  $81.2 (N=20)$  до  $72.4 (N=50)$ , після чого знову зростає до  $\approx 78$  при  $N \geq 100$ . Така поведінка свідчить про наявність оптимальної кількості рекурсивних рівнів, за якої досягається баланс між глобальною структурою та локальною деталізацією. Як видно з рис. 1–2, глибші ієрархії ( $N=50$ ) формують узгоджену глобальну композицію та плавні переходи між підобластями, тоді як менша глибина ( $N=20$ ) спричиняє локальні артефакти й втрату семантичної цілісності сцени.

Отримані результати підтверджують, що рекурсивний фреймворк, описаний у роботі [2], може ефективно поєднуватися з дифузійними модулями без послідовної авторегресії. Збережено ієрархічну композицію сцени та покращено можливості паралельної генерації підрегіонів. Наступні кроки включають теоретичний аналіз ініціалізації первинного шуму в латентному просторі та дослідження взаємозв'язку між спектральними й візуальними характеристиками згенерованих зображень залежно від глибини фрактала.

У перспективі такий фрактально-дифузійний підхід може бути поширений на інші типи даних, зокрема тривимірні сцени, відеопослідовності та мультимодальні моделі типу “текст  $\rightarrow$  зображення”. Завдяки ієрархічній структурі, подібні архітектури потенційно здатні поєднувати глобальну семантичну цілісність з локальною адаптивністю, що відкриває можливості для побудови більш універсальних генеративних систем нового покоління.

5. Ho, J., Jain, A., Abbeel, P. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*. 2020. Vol. 33. P. 6840–6851.
6. Li, T., Sun, Q., Fan, L., He, K. Fractal Generative Models. *arXiv preprint arXiv:2502.17437*. 2025.
7. Mandelbrot, B. B. *The Fractal Geometry of Nature*. Revised and enlarged edition. New York: W. H. Freeman, 1983.
8. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. High-Resolution Image Synthesis with Latent Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. P. 10684–10695.
9. Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., Van Gool, L. RePaint: Inpainting Using Denoising Diffusion Probabilistic Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. P. 11461–11471.