

Д. В. ЯКОВЛЕВ, студент, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна; e-mail: yakovlevdv31@gmail.com; ORCID: <https://orcid.org/0009-0005-4785-6361>

М. С. ГОЛІКОВ, аспірант, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна; e-mail: maksym.holikov@karazin.ua; ORCID: <https://orcid.org/0000-0003-4842-7823>

В. Є. СТРИЛЕЦЬ, кандидат технічних наук (PhD), доцент кафедри комп'ютерних систем та робототехніки, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна; e-mail: viktoria.strilets@karazin.ua; ORCID: <https://orcid.org/0000-0002-2475-1496>

АНАЛІЗ ВПЛИВУ ПОПЕРЕДНЬОГО ВІДНОВЛЕННЯ ЗАШУМЛЕНИХ ЗОБРАЖЕНЬ АВТОЕНКОДЕРОМ НА ТОЧНІСТЬ КЛАСИФІКАЦІЇ CNN

У роботі досліджено вплив попереднього відновлення зображень за допомогою денойзингового автоенкодера (DAE) на точність класифікації згортковою нейронною мережею (CNN) при різних типах шумів. Актуальність теми зумовлена тим, що в реальних умовах оптичні зображення часто містять спотворення, спричинені зміною освітлення, вібраціями, рухом камер та іншими факторами, що істотно ускладнює завдання розпізнавання об'єктів. Традиційні фільтри не завжди забезпечують достатню якість очищення та можуть призводити до втрати важливих структурних ознак. У зв'язку з цим використання глибоких нейронних мереж, зокрема автоенкодерів, постає перспективним напрямом підвищення стійкості алгоритмів комп'ютерного зору до шумів різної природи. У дослідженні використано датасет CIFAR-10 та реалізовано двокомпонентну модель: автоенкодер для попереднього очищення та CNN для класифікації. Навчений автоенкодер відновлює структуру зображення після впливу гауссівського, імпульсного, пуассонівського або спекл шумів. Було проведено три серії експериментів: класифікація чистих зображень, класифікація зашумлених даних без очищення та класифікація після попереднього відновлення автоенкодером. Результати показали, що на чистих даних CNN демонструє точність 70,37%, проте при внесенні шумів точність знижується до 30–59% залежно від типу спотворення. Після застосування автоенкодера точність класифікації зросла до 56–60% на всіх видах шумів, а найбільше покращення відзначено для гауссівського шуму з високою дисперсією. Отримані результати підтверджують, що використання автоенкодера як етапу попереднього відновлення є ефективним методом підвищення точності класифікації та зменшення вразливості CNN до шумів. Такий підхід забезпечує краще узагальнення та стабільність роботи системи, що особливо важливо для застосувань у реальному часі – зокрема в динамічних системах, робототехніці, автономному транспорті та навігаційних комплексах, де якість оптичних даних часто є нестабільною.

Ключові слова: автоенкодер, згорткова нейронна мережа, шум, класифікація, комп'ютерний зір, інтелектуальний аналіз зображень.

Вступ.

У сучасному світі цифрові зображення є невід'ємною частиною практично всіх сфер людської діяльності – від медицини й транспорту до промисловості, освіти та розваг. З розвитком технологій комп'ютерного зору та машинного навчання зображення стали не лише засобом передачі інформації, але й ключовим джерелом даних для інтелектуальних систем. Завдяки нейронним мережам комп'ютерні системи сьогодні здатні ідентифікувати об'єкти, визначати їхні межі, розпізнавати сцени, виявляти аномалії та генерувати нові візуальні дані.

Проте на практиці зображення часто містять спотворення: шум, артефакти, розмитість тощо. Оптичні зображення, отримані під час управління рухомими динамічними системами (дрони, роботизовані платформи, автономні транспортні засоби), часто містять шум, спричинений рухом, коливанням камери, зміною освітлення та зовнішніми завадами, що ускладнює розпізнавання об'єктів у режимі реального часу. Ці спотворення можуть бути виявлені, наприклад, у процесі зйомки, передавання або обробки фото та відео, через що вони зазнають спотворень. Однією з ключових проблем є шум, що ускладнює автоматичне розпізнавання об'єктів на цих зображеннях. Тому створення моделей, здатних ефективно працювати у зашумлених умовах, є важливою задачею сучасного машинного навчання.

Завдання з обробки таких зображень має особливе значення у сфері машинного навчання, адже якість

вхідних даних безпосередньо впливає на точність моделі. Традиційні фільтри, такі як медіанний або гауссівський, мають обмежену ефективність і часто призводять до втрати важливих деталей. Саме тому сучасні дослідження зосереджуються на використанні глибоких нейронних мереж, здатних самостійно навчатися виділяти ознаки й відновлювати структуру зображення навіть при значному рівні шуму.

Отже, проблема покращення обробки зашумлених зображень є актуальною як з теоретичної, так і з практичної точки зору. Вона охоплює питання фільтрації, реконструкції, сегментації та класифікації даних у складних умовах

Теоретичні основи та постановка задачі.

До моделей розпізнавання зображень висуваються вимоги не тільки виявити об'єкт, але й відокремити його від шуму. Умовно цю задачу можна розділити на два етапи: очищення зображення (denoising) та його класифікацію/сегментацію.

У залежності від типу шуму можна визначити характер спотворення, а також складність подальшої обробки. Шуми розподіляються на:

- гауссівський шум (gaussian noise) містить нормальний розподіл та додає до кожного пікселя випадкове значення, через що зображення стає менш чітким. Він часто виникає через електронні перешкоди або слабе освітлення;
- імпульсний шум (salt-and-pepper) проявляється у вигляді поодиноких чорних і білих пікселів, зазви-

© Яковлев Д. В., Голіков М. С., Стрілець В. Є., 2025



Дослідницька стаття: Цю статтю опубліковано видавництвом **НТУ «ХПІ»** у збірнику «Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології». Ця стаття поширюється за міжнародною ліцензією [Creative Commons Attribution \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). **Конфлікт інтересів:** Автор/и заявив/или про відсутність конфлікту.



чай спричинених втратою даних під час передавання або пошкодженням сенсорів;

- спекл-шум (speckle) з'являється при роботі з когерентним випромінюванням (лазерна зйомка, ультразвукова, радарна) і має мультиплікативний характер, що ускладнює його видалення;

- пуассонівський шум (poisson noise) характерний для низького рівня освітленості або коротких експозицій, коли інтенсивність світла має стохастичний розподіл [1–4].

У рухомих динамічних системах оптична інформація з камери надходить у реальному часі, що призводить до появи шумів різної природи: гауссівський шум від низького освітлення, імпульсного шуму через втрати кадрів, спекл-шуму під дією вібрацій та мультиплікативних впливів. Тому моделі обробки зображень, які забезпечують стійкість до шуму, є критично важливими для систем навігації, стабілізації, розпізнавання перешкод і цілей.

Кожен із цих типів шумів потребує конкретного індивідуального підходу до обробки, а універсальні методи на основі глибоких нейронних мереж дозволяють досягати стабільних результатів навіть без попереднього знання природи шуму.

Згорткові нейронні мережі (Convolutional Neural Network, CNN) – один з основних інструментів для класифікації та розпізнавання зображень, який працює завдяки механізму згортки і дозволяє виділяти локальні ознаки. CNN працюють на основі ієрархічного виділення ознак. Спочатку виявляються низькорівневі ознаки або характеристики, наприклад грані, потім, просуваючись глибше у мережу розпізнаються особливості вищого рівня, такі як форми та об'єкти. Перевагою CNN є висока точність класифікації, особливо при роботі з великими наборами даних. Проте при значному рівні шуму моделі можуть плутати текстурні артефакти з ознаками об'єкта, що призводить до зниження точності.

Таким чином, задача дослідження полягає у визначенні впливу попередньої обробки зображення, а саме очищення від шумів, на точність моделей класифікації на основі згорткової нейронної мережі.

Глибокі нейромережеві моделі для обробки зображень із шумом.

Для боротьби з шумом часто використовують попередню фільтрацію або поєднання CNN з автоенкодерами, які виконують очищення вхідних даних [5].

Автоенкодер – це тип нейронної мережі, що навчається стискати вхідні дані у компактне представлення (encoding) і потім відновлювати їх назад (decoding). Такий підхід дозволяє мережі вчитися витягати найважливіші ознаки. Окремий тип – denoising autoencoder (автоенкодер для видалення шуму) – спеціально навчається на парах “зашумлене – чисте” зображення, щоб відновлювати початкову структуру [6].

Перевага використання автоенкодерів полягає у здатності ефективно зменшувати вплив шуму без втрати дрібних деталей. Недоліком є те, що сам по собі автоенкодер не виконує класифікацію, тому часто використовується як попередній етап з CNN або іншими моделями.

Мережі U-Net використовуються для сегментації зображень і можуть точно локалізувати об'єкти, навіть при наявності шуму. Завдяки симетричній структурі та механізмам збереження деталей (skip connections), U-Net ефективно відновлює структуру об'єкта і чудово працює навіть на невеликих наборах даних і відзначається високою точністю локалізації об'єктів.

Варто відзначити, що завдяки здатності відновлювати структуру навіть із пошкоджених зображень, U-Net часто використовується для сегментації та очищення зашумлених медичних зображень, проте є певний недолік, а саме значна обчислювальна складність та потреба у великих ресурсах для навчання, особливо при роботі з великими зображеннями [7].

Архітектура моделі глибокого навчання Vision Transformer (ViT) переносить ідеї трансформерів, успішних в обробці текстів, у галузь комп'ютерного зору. Замість згорток, ViT використовує механізм самоуваги (self-attention), який дозволяє враховувати глобальні залежності між частинами зображення. Ця архітектура показує високу точність на великих наборах даних, проте, так само як і для U-Net, для навчання потрібні значні ресурси. При обмежених даних ViT може поступатися CNN через недостатню локалізацію ознак [8].

Також не можна обійти стороною той факт, що протягом останніх років значного поширення набули моделі дифузії (Denoising Diffusion Models). Вони працюють за принципом поступового додавання шуму до зображення, а потім навчаються відновлювати його у зворотному напрямку. Цей процес дозволяє моделі “навчитися” очищати будь-які типи спотворень. Моделі дифузії демонструють надзвичайно високу якість реконструкції та сьогодні використовуються не лише для очищення, а й для генерації нових зображень (наприклад, у системах Stable Diffusion), однак через високу обчислювальну складність може бути обмеженою у використанні в реальному часі [9].

Реалізація процесу очищення та класифікації зображень.

Для проведення дослідження була використана мова програмування Python [10] та бібліотеки для роботи з зображеннями і нейронними мережами:

- TensorFlow – для побудови та навчання нейронної мережі;
- NumPy – для обробки даних;
- Matplotlib – для візуалізації результатів;
- Scikit-image – для генерації шуму на зображеннях [11, 12].

Для навчання і тестування моделей було використано датасет CIFAR-10, який містить 60.000 зображень розміром 32x32 пікселі, розділених на 10 класів за типами об'єктів на зображеннях.

Було побудовано двокомпонентну модель обробки зображень, яка складається з денойзингового автоенкодера (denoising autoencoder, DAE) для попереднього очищення оптичних даних та згорткової нейронної мережі (CNN) для подальшої класифікації.

Автоенкодер використовує оптимізатор Adam і функцію втрат MSE (через завдання реконструкції).

Його архітектура – симетрична згорткова структура, що складається з двох частин: енкодера та декодера. Енкодер перетворює вхідне зображення $32 \times 32 \times 3$ у компактне представлення $8 \times 8 \times 64$, щоб витягнути компактне та стійке до шуму латентне представлення. А декодер у свою чергу реконструює очищене зображення з латентного представлення.

CNN використовується як класифікатор після очищення зображень автоенкодером. Архітектура згорткової нейронної мережі складається з двох згорткових блоків (Conv2D $\rightarrow$ MaxPooling), які формують багаторівневе представлення зображення, та двох повнозв'язних шарів для класифікації. Модель має 32 та 64 фільтри в згорткових шарах відповідно, використовує функцію активації ReLU, а на завершальному етапі застосовує softmax для розподілу ймовірностей між 10 класами. Додавання шару Dropout зі значенням 0.3 дозволило суттєво зменшити ризик перенавчання.

Експериментальні результати та їх аналіз. На першому етапі було проведено базове навчання згорткової нейронної мережі без жодного шуму у вхідних даних. Мета цього етапу – визначити базову точність класифікації для подальшого порівняння з результатами на зашумлених даних.

Результати показали, що модель досягає точності 0.7037 (70.37%) на тестовій вибірці без шуму. Це є типовим значенням для базової CNN, натренованої на CIFAR-10 без спеціальної оптимізації. Дане значення в рамках дослідження було вибрано як еталонна точка для подальших експериментів.

На другому етапі було проведено серію експериментів для визначення, як різні типи шуму впливають на точність класифікації. Для цього до тестових зображень CIFAR-10 було штучно додано кілька типів шумів:

- Gaussian noise – випадкові коливання яркості пікселів ($\sigma = 0.1, 0.2$);
- Salt-and-Pepper noise – поодинокі чорні та білі пікселі;
- Speckle noise – мультиплікативний шум, який додає зернистість;

- Poisson noise – стохастичні флуктуації інтенсивності, характерні для слабкого освітлення.

Після додавання шумів модель CNN тестувалася без жодного попереднього очищення. Отримані результати наведено нижче в табл. 1:

Таблиця 1 – Показники точності для різних типів шумів

Тип шуму	Точність
Gaussian ($\sigma=0.1$)	0.5295
Gaussian ($\sigma=0.2$)	0.2968
Salt & Pepper	0.4993
Speckle	0.5942
Poisson	0.5890

Як видно з табл. 1, усі типи шуму призвели до зниження точності, причому найбільший негативний вплив мав гауссівський шум із високою дисперсією ($\sigma=0.2$), де точність впала майже удвічі. Найменше погіршення відбулося при пуассонівському та спекл шумі, оскільки їхній вплив менш руйнівний для локальних контурів і текстур.

Отримані результати свідчать про те, що згорткові мережі, незважаючи на здатність виділяти локальні ознаки, залишаються вразливими до спотворень даних, якщо попередньо не проводиться фільтрація або нормалізація вхідного зображення.

На третьому етапі було проведено дослідження з використанням DAE – нейронної мережі, здатної навчатися відновлювати чисте зображення із зашумленого.

Автоенкодер було навчено на парах “зашумлене – чисте” зображення з гауссівським шумом ($\sigma=0.1$). Після навчання модель використовувалася для очищення тестових зображень з різними типами шуму перед подачею їх до CNN, що залишалася незмінною. Результати тестування CNN показані на рис. 1.

Порівняно з другим етапом дослідження, видно суттєве покращення точності у всіх випадках. Наприклад, для гауссівського шуму з $\sigma=0.2$ точність зросла з 0.2968 до 0.5635, тобто майже удвічі. Це

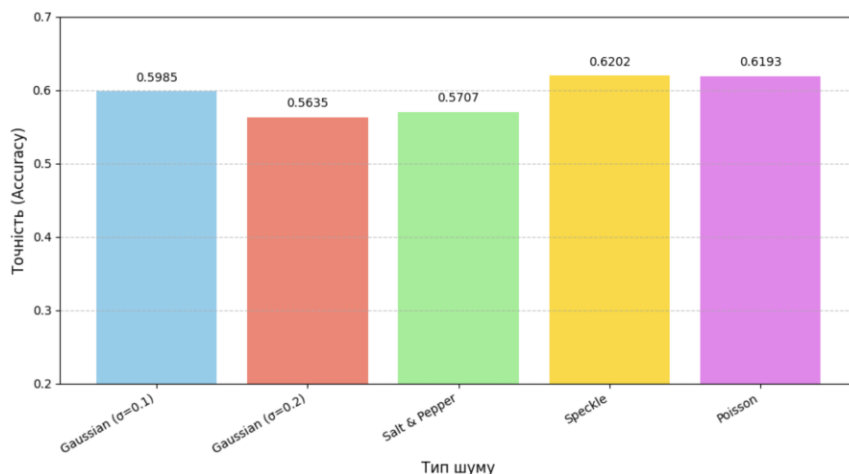


Рис. 1. Порівняння точності CNN після очищення для різних видів шумів

підтверджує, що попереднє використання автоенкодера може ефективно зменшувати вплив шуму, покращуючи роботу згорткової мережі.

Загалом результати після очищення стабілізувалися в межах 0.56–0.62, як показано на рис. 2. Це свідчить про зменшення розкиду точності між різними типами шумів і підвищення узагальнюючої здатності моделей.

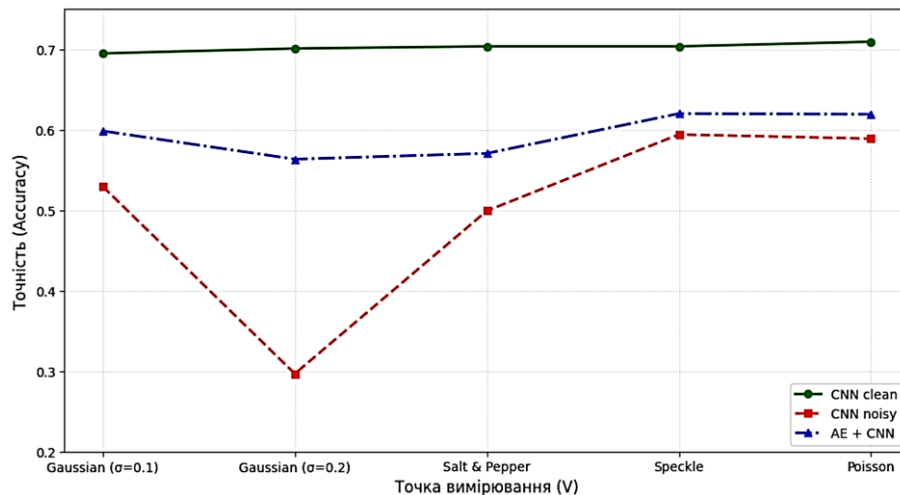


Рис. 2. Порівняння точності моделей

З огляду на отримані результати, інтеграція автокодувальника як етапу попередньої обробки є ефективною для побудови надійних систем класифікації зображень, зокрема у реальних умовах, де якість даних не може бути гарантована. Такий підхід забезпечує баланс між високою точністю (наближеною до чистої моделі) та необхідною стійкістю до різноманітних шумів.

Висновки. У дослідженні було проведено три експерименти, спрямовані на аналіз впливу попереднього відновлення зображень автоенкодером на точність класифікації згортковою нейронною мережею (CNN) при різних типах шумів.

Перший експеримент визначив базову точність роботи CNN на чистих даних CIFAR-10, що становила 70,37%. Це значення було використано як еталон для подальшого порівняння результатів на зашумлених зображеннях.

Другий експеримент показав, що додавання шуму суттєво знижує точність класифікації: залежно від типу шуму вона коливалася у межах 30–68%. Таке падіння демонструє високу вразливість CNN до спотворень у візуальних даних та підтверджує необхідність підсилення стійкості моделей комп'ютерного зору в умовах деградованої інформації.

У третьому експерименті зображення перед класифікацією проходили попереднє відновлення за допомогою денойзингового автоенкодера. Отримані результати показали суттєве підвищення точності – до 56–60% на всіх типах шумів. Це свідчить про те, що автоенкодер здатен ефективно компенсувати негативний вплив спотворень та відновити частину початкової

структури зображення, необхідної для коректної роботи CNN. Таким чином, проведений аналіз підтверджує, що попереднє відновлення зображень автоенкодером є дієвим механізмом підвищення точності класифікації в умовах шумів різної природи.

Отримані результати показують, що комбінація автоенкодера та CNN може бути використана для підвищення стійкості систем комп'ютерного зору, які

працюють у складі динамічних систем. Це дозволить покращити якість ідентифікації об'єктів у реальному часі навіть за умови значних спотворень оптичних сигналів, що є важливим для задач навігації, керування рухом, супроводу об'єктів та уникнення зіткнень.

Список використаної літератури

1. Cui Z.-X., Fan Q. A “Nonconvex+Nonconvex” approach for image restoration with impulse noise removal. *Applied Mathematical Modelling*. 2018. Vol. 62. P. 254–271. DOI: 10.1016/j.apm.2018.05.035.
2. Dong W., Paz-Silva G. A., Viola L. Resource-efficient digital characterization and control of classical non-Gaussian noise. *Applied Physics Letters*. 2023. Vol. 122. Art. 244001. DOI: 10.1063/5.0153530.
3. Liu J., Ni A., Ni G. A nonconvex $l_1(l_1-l_2)$ model for image restoration with impulse noise. *Journal of Computational Applied Mathematics*. 2020. Vol. 378. Art. 112934. DOI: 10.1016/j.cam.2020.112934.
4. Hernandez-Hernandez D., Ricalde-Guerrero J. H. Conditional McKean-Vlasov Differential Equations with Common Poissonian Noise: Propagation of Chaos. *arXiv, Mathematics*. 2023. DOI: 10.48550/arXiv.2308.11564.
5. Purwono P., Ma'arif A., Rahmian W., Fathurrahman H. I. K., Frisky A. Z. K., Haq Q. M. U. Understanding of convolutional Neural Network (CNN): A Review. *International Journal of Robotics and Control Systems*. 2022. Vol. 2. No. 4. P. 739–748. DOI: 10.31763/ijrcs.v2i4.888.
6. Michelucci U. An introduction to autoencoders. *arXiv, Computer Science*. 2022. DOI: 10.48550/arXiv.2201.03898.
7. Schonfeld E., Schiele B., Khoreva A. A U-Net Based Discriminator for Generative Adversarial Networks. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (13–19 June 2020, Seattle, WA, USA)*. 2020. P. 8207–8216. DOI: 10.1109/CVPR42600.2020.00823.
8. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv, Computer Science*. 2020. DOI: 10.48550/arXiv.2010.11929.

9. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*. 2020. Vol. 33. P. 6840–6851. DOI: 10.48550/arXiv.2006.11239.
10. Sharma V. K., Kumar V., Sharma S., Pathak S. *Python Programming: A Practical Approach*. CRC Press, 2021. 310 p. DOI: 10.1201/9781003185505.
11. Ramchandani M., Khandare H., Singh P., Rajak P., Suryawanshi N., Jangde A. S., Sahu M. Survey: Tensorflow in Machine Learning. *Journal of Physics: Conference Series*. 2022. Vol. 2273. Art. 012008. DOI: 10.1088/1742-6596/2273/1/012008.
12. Harris J. R., Millman K. J., van der Walt S. J., Gommers R., Virtanen P., Cournapeau D., Wieser E., Taylor J., Berg S., Smith N. J., Kern R., Picus M., Hoyer S., van Kerkwijk M. H., Haldane A., Fernández del Río J., Wiebe M., Peterson P., Gérard-Marchant P., Sheppard K., Reddy T., Weckesser W., Abbasi H., Gohlke C., Oliphant T. E. Array programming with NumPy. *Nature*. 2020. Vol. 585. P. 357–362. DOI: 10.1038/s41586-020-2649-2.
5. Purwono P., Ma'arif A., Rahmaniar W., Fathurrahman H. I. K., Frisky A. Z. K., Haq Q. M. U. Understanding of convolutional Neural Network (CNN): A Review. *International Journal of Robotics and Control Systems*. 2022, vol. 2, no. 4, pp. 739–748. DOI: 10.31763/ijrcs.v2i4.888.
6. Michelucci U. An introduction to autoencoders. *arXiv, Computer Science*. 2022. DOI: 10.48550/arXiv.2201.03898.
7. Schonfeld E., Schiele B., Khoreva A. A U-Net Based Discriminator for Generative Adversarial Networks. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (13–19 June 2020, Seattle, WA, USA)*. 2020, pp. 8207–8216. DOI: 10.1109/CVPR42600.2020.00823.
8. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv, Computer Science*. 2020. DOI: 10.48550/arXiv.2010.11929.
9. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*. 2020, vol. 33, pp. 6840–6851. DOI: 10.48550/arXiv.2006.11239.
10. Sharma V. K., Kumar V., Sharma S., Pathak S. *Python Programming: A Practical Approach*. CRC Press, 2021. 310 p. DOI: 10.1201/9781003185505.
11. Ramchandani M., Khandare H., Singh P., Rajak P., Suryawanshi N., Jangde A. S., Sahu M. Survey: Tensorflow in Machine Learning. *Journal of Physics: Conference Series*. 2022, vol. 2273, article 012008. DOI: 10.1088/1742-6596/2273/1/012008.
12. Harris J. R., Millman K. J., van der Walt S. J., Gommers R., Virtanen P., Cournapeau D., Wieser E., Taylor J., Berg S., Smith N. J., Kern R., Picus M., Hoyer S., van Kerkwijk M. H., Haldane A., Fernández del Río J., Wiebe M., Peterson P., Gérard-Marchant P., Sheppard K., Reddy T., Weckesser W., Abbasi H., Gohlke C., Oliphant T. E. Array programming with NumPy. *Nature*. 2020, vol. 585, pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

References (transliterated)

1. Cui Z.-X., Fan Q. A “Nonconvex+Nonconvex” approach for image restoration with impulse noise removal. *Applied Mathematical Modelling*. 2018, vol. 62, pp. 254–271. DOI: 10.1016/j.apm.2018.05.035.
2. Dong W., Paz-Silva G. A., Viola L. Resource-efficient digital characterization and control of classical non-Gaussian noise. *Applied Physics Letters*. 2023, vol. 122, article 244001. DOI: 10.1063/5.0153530.
3. Liu J., Ni A., Ni G. A nonconvex $l_1(l_1-l_2)$ model for image restoration with impulse noise. *Journal of Computational Applied Mathematics*. 2020, vol. 378, article 112934. DOI: 10.1016/j.cam.2020.112934.
4. Hernandez-Hernandez D., Ricalde-Guerrero J. H. Conditional McKean-Vlasov Differential Equations with Common Poissonian Noise: Propagation of Chaos. *arXiv, Mathematics*. 2023. DOI: 10.48550/arXiv.2308.11564.

Надійшла (received) 29.11.2025

UDC 004.8

D. V. YAKOVLEV, Student, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine; e-mail: yakovlevdv31@gmail.com; ORCID: <https://orcid.org/0009-0005-4785-6361>

M. S. HOLIKOV, Postgraduate student, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine; e-mail: maksym.holikov@karazin.ua; ORCID: <https://orcid.org/0000-0003-4842-7823>

V. Y. STRILETS, Candidate of Technical Sciences (PhD), V. N. Karazin Kharkiv National University, Associate Professor at the Department of Computer Systems and Robotics, Kharkiv, Ukraine; e mail: viktoria.strilets@karazin.ua; ORCID: <https://orcid.org/0000-0002-2475-1496>

ANALYSIS OF THE IMPACT OF PRELIMINARY NOISY IMAGE RESTORATION BY AUTOCODER ON THE ACCURACY OF CNN CLASSIFICATION

The paper investigates the impact of preliminary image restoration using a denoising autoencoder (DAE) on the classification accuracy of a convolutional neural network (CNN) under various types of noise. The relevance of the topic is due to the fact that in real conditions, optical images often contain distortions caused by changes in lighting, vibrations, camera movement, and other factors, which significantly complicates the task of object recognition. Traditional filters do not always provide sufficient cleaning quality and can lead to the loss of important structural features. In this regard, the use of deep neural networks, in particular autoencoders, is a promising direction for improving the robustness of computer vision algorithms to noise of various nature. The study uses the CIFAR-10 dataset and implements a two-component model: an autoencoder for preliminary cleaning and a CNN for classification. The trained autoencoder restores the image structure after exposure to Gaussian, impulse, Poisson, and speckle noise. Three series of experiments were conducted: classification of clean images, classification of noisy data without cleaning, and classification after preliminary restoration by the autoencoder. The results showed that CNN demonstrates an accuracy of 70.37% on clean data, but when noise is introduced, the accuracy drops to 30–59% depending on the type of distortion. After applying the autoencoder, classification accuracy increased to 56–60% for all types of noise, with the greatest improvement observed for Gaussian noise with high dispersion. The results confirm that using an autoencoder as a preliminary restoration step is an effective method for improving classification accuracy and reducing CNN vulnerability to noise. This approach provides better generalization and stability of the system, which is especially important for real-time applications—in particular, in dynamic systems, robotics, autonomous transport, and navigation systems, where the quality of optical data is often unstable. The study demonstrates the promise of integrating restoration and classification models into a single structure to improve the performance of computer vision systems in challenging conditions.

Keywords: autoencoder, convolutional neural network, noise, classification, computer vision, intelligent image analysis.

Повні імена авторів / Author's full names

Автор 1 / Author 1: Яковлев Данило В'ячеславович / Yakovlev Danylo Viacheslavovych

Автор 2 / Author 2: Голіков Максим Сергійович / Holikov Maksym Serhiiovych

Автор 3 / Author 3: Стрілець Вікторія Євгенівна / Strilets Viktoriia Yevhenivna