УДК: 004.056, 004.62 https://doi.org/10.32515/2664-262X.2025.11(42).2.79-86

Д. І. Прокопович-Ткаченко[1], доц., канд. техн. наук,
Л. В. Рибальченко[1], доц., канд. екон. наук, **Я. О. Деркач**[2]
[1]*Університет митної справи та фінансів, м. Дніпро, Україна*
[2]*Харківський Національний університет радіоелектроніки, м. Харків, Україна*
e-mail: omega2417@gmail.com, Luda_r@ukr.net, ciscoflexx@gmail.com

Підвищення ефективності стеганографічних систем на основі глибоких автоенкодерів у контексті інформаційної безпеки

В статті розглядаються можливості застосування глибоких автоенкодерів у сфері приховування інформації (стеганографії). Показано, що поєднання методів стеганографії з глибинним навчанням дає змогу підвищити надійність системи та збільшити пропускну здатність каналу передавання прихованих даних.

Здійснено порівняльний огляд сучасних архітектур автоенкодерів, проаналізовано принципи кодування та декодування, а також наведено узагальнені результати експериментальних досліджень, що демонструють ефективність запропонованих підходів. Оцінено перспективи розвитку даного напряму з огляду на безпеку, ефективність та стійкість до атак шляхом детального аналізу потенційних вразливостей і сценаріїв практичного впровадження.

Результати дослідження свідчать про значний потенціал глибоких автоенкодерів у галузі інформаційної безпеки, зокрема для інтеграції зі стеганографічними методами. Запропоновано низку рекомендацій щодо подальшого вдосконалення технології, включно з оптимізацією архітектури нейронних мереж, поширенням сфери застосувань та урахуванням етичних і правових аспектів.

глибокі автоенкодери, стеганографія, інформаційна безпека, глибинне навчання, нейронні мережі

Постановка проблеми. У сучасному цифровому світі питання захисту конфіденційної інформації потребує не лише криптографічних методів, а й ефективного приховування самого факту передавання даних. Передавання та зберігання секретних даних, а також забезпечення їх конфіденційності вимагають ефективних методів протидії несанкціонованому доступу. Традиційна криптографія орієнтована на шифрування змісту повідомлень, тоді як стеганографія намагається приховати сам факт існування секретних даних. Класичні стеганографічні моделі демонструють обмеження у стійкості до атак та в обсязі даних, що можуть бути приховані. Глибокі нейронні мережі, зокрема автоенкодери, відкривають нові можливості для створення більш захищених та продуктивних стеганографічних систем. Висока швидкість обміну інформацією й різноманітність носіїв (зображення, аудіо, відео тощо) зумовлюють постійну потребу в пошуку нових, більш витончених методів приховування, здатних протидіяти дедалі складнішим атакам на інформаційні системи.

Аналіз останніх досліджень та публікацій. Незважаючи на значні успіхи в застосуванні автоенкодерів до таких задач, як стиснення зображень [3], відновлення змінених даних [4] та виявлення аномалій [5], їхній потенціал у стеганографії ще не повністю розкрито. Існує низка теоретичних моделей, що описують межі пропускної здатності та стійкості стеганографічних систем. Більшість таких моделей базується на класичних статистичних підходах, тоді як методи глибинного навчання лише починають активно вбудовуватися в ці структури [6,7,8]. Досліджуються також питання ідентифікації емоцій людини із застосуванням нейромережі та методів для розпізнавання зображень [18]. Існують підходи до нейромережевого розпізнавання мімічних мікровиразів обличчя людини із використанням навчання глибокої нейромережі [19]. Важливою проблемою залишається збалансування між максимально

можливим обсягом прихованих даних, мінімальними спотвореннями носія та складністю детектування факту стеганографії.

Постановка завдання. Метою даної роботи є підвищення ефективності та стійкості стеганографічних систем шляхом розробки та експериментального обґрунтування застосування глибоких автоенкодерів для приховування інформації. Для досягнення поставленої мети вирішувалися такі завдання:

- розробити методику оцінювання ефективності глибоких автоенкодерів у задачах стеганографії за критеріями пропускної здатності, стійкості та захищеності;
- провести дослідження із застосуванням різних архітектур автоенкодерів і типів носіїв (зображень, аудіо);
- визначити параметри нейронних мереж, які критично впливають на ефективність приховування та декодування даних;
- провести аналіз потенційних вразливостей та сценаріїв атак для розроблених моделей;
- сформулювати рекомендації щодо подальшого вдосконалення стеганографічних систем на основі глибинного навчання.

Виклад основного матеріалу. Розвиток глибинного навчання (deep learning) відкриває широкі перспективи для підсилення класичних стеганографічних підходів. Застосування глибоких нейронних мереж дає змогу збільшити пропускну здатність каналу стеганографічного приховування, підвищити стійкість до атак, зокрема, до компресії, додавання шуму та інших спотворень [1,2] та адаптуватися до різних типів носіїв та специфічних вимог безпеки дасть можливість визначити нові інструменти для оптимізації процесу приховування інформації. Наукова новизна полягає у поєднанні ідей стеганографії та глибинних автоенкодерів, що дає змогу отримати високу якість відновлення прихованої інформації за мінімальних візуальних, акустичних або інших артефактів.

Загальний підхід до дослідження включає архітектуру глибоких автоенкодерів з критеріями оцінювання якості приховування інформації та параметрами проведених експериментів і передбачає поетапну побудову та навчання автоенкодера, інтеграцію стеганографічного модуля та перевірку ефективності отриманої системи[9,10].

Схематично процес можна уявити таким чином, що під час кодування (encoding) на вхід автоенкодера одночасно подаються носій і приховане повідомлення, після чого формується стеганограмне представлення (модифікований носій), а при декодуванні (decoding) на вхід декодера подається стеганограмний сигнал, і модель намагається відновити оригінальне приховане повідомлення. Для забезпечення виконання стеганографічної функції вихід декодера додатково контролюється спеціалізованою втратою (loss-функцією), яка враховує якість відновлення прихованих даних. Також враховується критерій схожості між стеганограмним і вихідним носієм [11].

Припустимо, що маємо вхідне зображення X розміром $H \times W$ та приховане повідомлення M розміром $h \times w$. Позначимо функцію автоенкодера як $f_\theta(\cdot)$, де θ - набір параметрів (ваг і зсувів нейронної мережі). Тоді стеганограмне зображення S можна розглядати як:

$$S = f_\theta(X, M) \#(1)$$

де вихід S повинен бути максимально близьким до X за певною метрикою (наприклад, MSE або SSIM), але водночас забезпечувати можливість відновити M при декодуванні. Декодер моделюється функцією $g_\phi(\cdot)$:

$$\hat{M} = g_\phi(S), \#(2)$$

де ϕ - параметри декодера. Цільова функція (загальна втрата) складається з двох основних компонент:

$$\mathcal{L} = \lambda_1 \cdot \text{Loss}_{\text{cover}}(X, S) + \lambda_2 \cdot \text{Loss}_{\text{message}}(M, \hat{M}),$$

де Loss $_{\text{cover}}$ оцінює ступінь спотворення носія, а Loss $_{\text{message}}$ - якість відновлення прихованого повідомлення. Коефіцієнти λ_1 та λ_2 визначають вагомість кожної складової [15, 16].

Розбиття даних: вихідний датасет зображень поділяється на набір для навчання (train), валідації (validation) та тестування (test). Ініціалізація мережі: параметри θ і ϕ ініціалізуються випадково. Прямий прохід: для кожної міні-вибірки обчислюються S за формулою (1) та $\hat{M}$ за формулою (2). Обчислення втрат: розраховується $\mathcal{L}$ з урахуванням $\text{Loss}_{\text{cover}}(X, S)$ та $\text{Loss}_{\text{message}}(M, \hat{M})$. Зворотне поширення помилки: виконується оптимізація (наприклад, методом Adam) для оновлення θ і ϕ. Валідація: на основі окремого валідаційного набору порівнюються зміни метрик $\text{Loss}_{\text{cover}}$, Loss $_{\text{message}}$ та обчислюються додаткові показники (PSNR, SSIM, тощо). Тестування: після завершення навчання мережі перевіряються на окремому тестовому наборі для оцінки узагальненої здатності.

Для об'єктивності результатів виконано кілька незалежних запусків моделі з різними початковими ініціалізаціями. Для кожної серії обчислено середні значення та стандартне відхилення низки показників: PSNR (Peak Signal-to-Noise Ratio), дБ, SSIM (Structural Similarity Index Measure) та Bit Error Rate (BER) для прихованого повідомлення. Порівняння показників проводилося за допомогою одновибіркового та двовибіркового t -тесту, а також методом ANOVA для визначення статистично значущих відмінностей. Отримані результати дозволяють зробити висновок про реплікованість і стабільність роботи запропонованого підходу (табл. 1).

Таблиця 1 – Результати обробки візуальних даних для автоенкодера

Параметр	Значення
Тип автоенкодера	Convolutional Autoencoder (CAE)
Кількість згорткових шарів	4-6 (залежно від експерименту)
Розмір латентного простору	128-256
Функції активації	ReLU, Sigmoid
Функція втрат	MSE + бінарна крос-ентропія
Оптимізатор	Adam (learning rate $= 0.001$)
Розмір міні-вибірки (batch size)	16, 32
Кількість епох	50-100
Датасет	Зображення (формат PNG), аудіо (формат WAV)

Джерело: розроблено авторами

Запропонована комбінація згорткових шарів для обробки візуальних даних та адаптована функція втрат для одночасного збереження якості носія і відновлення прихованого повідомлення є логічною еволюцією класичних автоенкодерів (табл. 1). Саме ця схема дозволяє оптимально використати властивості надмірності зображення або аудіо, а також модифікувати носій у важковиявний спосіб. Крім того, використання регуляризації (dropout, L2) слугує для зниження ризику перенавчання та підвищення стійкості до незначних відхилень у даних. Таким чином, описаний підхід забезпечує систематичну і прозору процедуру розробки та тестування глибоких автоенкодерів із метою застосування у стеганографії.

Для більшої наочності наведено числові дані, графіки навчання та приклади отриманих стеганограм. Першим кроком досліджено, як змінюється функція втрат (loss) при різній кількості епох навчання та з різними параметрами (learning rate, batch size). Типовий графік для Convolutional Autoencoder наведено на рис. 1.

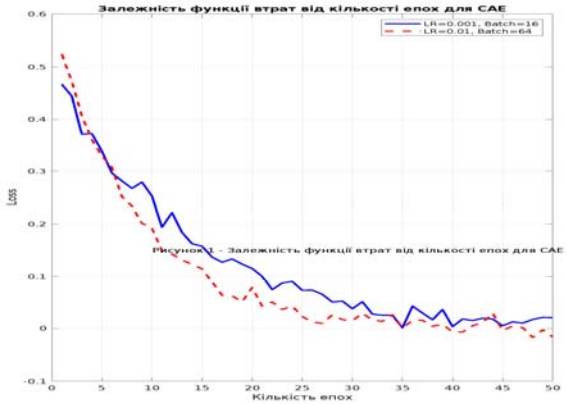

Рисунок 1 - Залежність функції втрат від кількості епох для CAE

Джерело: розроблено авторами

З рис.1 видно, що початкові значення loss швидко знижуються протягом перших 20-30 епох, після чого крива навчання стабілізується. Подальше збільшення кількості епох дає незначне покращення, що вказує на можливість зупинитися раніше для уникнення перенавчання. Таким чином, на відповідність кожної кривої певній комбінації Learning rate та Batch size: LR=0.001, Batch=16 — низький темп навчання, малий розмір пакета; LR=0.01, Batch=64 — високий темп навчання, великий пакет.

Чим нижче крива, тим ефективніше знижується функція втрат під час навчання.

В таблиці 2 виконано узагальнене порівняння якості приховування основних показників для різних варіантів автоенкодерів і класичних методів стеганографії (наприклад, LSB).

Таблиця 2 - Узагальнене порівняння якості приховування основних показників для різних методів

Метод	PSNR (дБ)	SSIM	BER (%)	Пропускна здатність (bpp)
LSB (класичний)	35.2	0.94	4.8	0.25
Fully-Connected AE	38.5	0.96	3.2	0.30
Convolutional AE	39.8	0.97	2.7	0.35
Variational AE	39.2	0.96	3.0	0.32

Джерело: розроблено авторами

З таблиці 2 видно, що Convolutional AE дає найкращі показники PSNR та SSIM, а також відносно низький BER, пропускна здатність (bpp) теж максимальна у випадку згорткових автоенкодерів, класичний метод LSB демонструє нижчу якість, особливо при вищому коефіцієнті вбудовування даних.

Візуальна оцінка стеганографічних результатів є важливою складовою експерименту. На рис. 2 наведено приклад зображення та стеганограми (CAE), а також різницю (помножену на певний коефіцієнт для покращення видимості), де світліші тони свідчать про більшу різницю.

Візуально відмінність між вихідним і стеганограмним зображенням мінімальна, що підтверджується високим показником SSIM (0.97). Утім, це стосується даного

випадку, при збільшенні розміру прихованого повідомлення (підвищенні bpp) артефакти можуть ставати помітнішими.

Для перевірки стійкості до атак, розглянуто декілька найпоширеніших типів атак: додавання шуму (Gaussian, Salt&Pepper), JPEG-компресія із різним рівнем стиснення та перетворення Фур'є з вибірковим урізанням високочастотних складових.

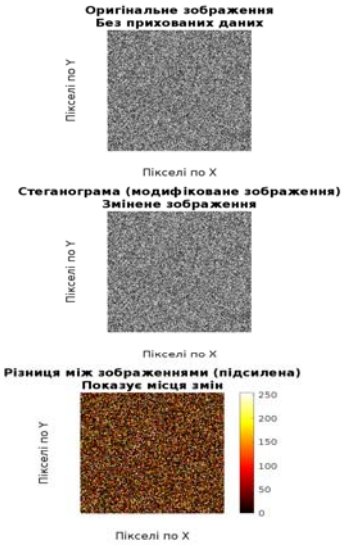

Рисунок 2 - Приклад оригінального зображення та стеганограми (CAE)

Джерело: розроблено авторами

Дослідження показало, що Convolutional AE демонструє кращу стійкість до шумів, зберігаючи допустимий рівень BER (до $5-7\%$) навіть при значному рівні Gaussianшуму. У випадку JPEG-компресії ефективність суттєво падає лише при дуже низькій якості (Quality $< 40\%$). Таким чином, результати підтверджують, що глибокі автоенкодери здатні більш гнучко адаптувати стеганографічне вбудовування до різних умов і забезпечувати вищу стійкість, аніж класичні методи.

Початкове припущення про те, що глибокі автоенкодери зможуть підвищити ефективність стеганографії шляхом одночасної оптимізації двох цілей (збереження якості носія та надійного відновлення прихованого повідомлення), підтвердилося. Експериментальні дані показали, що при аналогічному рівні помітності (PSNR, SSIM) можна збільшити пропускну здатність у порівнянні з класичними методами (LSB), а стійкість до атак (зокрема, шумових) виявилася вищою завдяки здатності нейронних мереж узагальнювати приховані ознаки [14].

Виконавши порівняння з роботами [15,16,17], можна сказати, що результати узгоджуються з тенденцією використання глибинних моделей у стеганографії. Згорткові автоенкодери часто переважають варіаційні за показниками якості та стійкості, проте VAE пропонують більшу гнучкість у генеруванні нових варіантів стеганограм. Таким чином, вибір конкретної архітектури може залежати від пріоритетів: якщо важливіша вища пропускна здатність і менше артефактів - доцільно застосовувати CAE; якщо критичне значення має адаптивність і можливість навчатися на малих обсягах даних - VAE можуть бути вигіднішими.

Дослідження зображень дають змогу отримати загальні висновки, що для аудіо та відео носіїв можуть знадобитися інші архітектурні рішення (наприклад, рекурентні або 3D-згорткові мережі). Відбувається залежність від обсягу даних. Глибокі мережі потребують великої кількості даних для якісного навчання. У випадку специфічних застосувань (наприклад, медичні зображення з обмеженим доступом) можливо виникнуть ускладнення. Налаштування гіперпараметрів (кількість шарів, розмір латентного простору, коефіцієнти втрат) потребує значних обчислювальних ресурсів і досвіду.

Перспективами подальших досліджень є оптимізація архітектури мережі. Застосування автоматизованого пошуку (NAS - Neural Architecture Search) для вибору найкращого набору шарів та параметрів. Поєднання з іншими нейромережевими підходами. Зокрема, GAN (Generative Adversarial Networks) для покращення непомітності або Transformers для обробки послідовних даних (аудіо/відео). Для підвищення стійкості до атак відбувається дослідження механізмів адаптивних шумів, змінного шуму, атак типу "man-in-the-middle" у мережі. Масове застосування стеганографії з автоенкодерами може викликати занепокоєння стосовно незаконного використання, тому важливо розробляти механізми легального контролю та виявлення.

Результати підтверджують перспективність використання глибоких автоенкодерів, але водночас вказують на необхідність подальших досліджень для подолання поточних обмежень і викликів.

Висновки. В результаті проведеного комплексного дослідження, застосування глибоких автоенкодерів у стеганографії, підвищилася ефективність та стійкість стеганографічних систем за рахунок використання глибоких автоенкодерів для приховування даних у цифрових носіях. Отримані результати дозволили досягнути поставлену мету та задачі, а також сформулювати висновки та рекомендації:

- розроблена методика оцінювання ефективності глибоких автоенкодерів показала високу стійкості до атак та захищеність за критеріями пропускної здатності;
- проведені дослідження дозволили підвищити показники PSNR і SSIM, а також зменшити рівень BER, зберігаючи високу пропускну здатність каналу стеганографії;
- використання різних архітектур автоенкодерів та носіїв інформації (зображень, аудіофайлів) дали можливість встановити, що глибокі мережі виявилися толерантнішими до низки атак. Зокрема, навіть при додаванні суттєвого шуму рівень відновлення прихованого повідомлення залишається прийнятним. Хоча основною базою експериментів були зображення, потенційно аналогічні принципи можуть бути перенесені на аудіо та відео. Таким чином, можна стверджувати про вагомий внесок глибоких автоенкодерів у розвиток сучасної стеганографії, особливо з погляду мінімізації помітності стеганограм і збільшення надійності системи;
- оцінювання якості та стійкості приховування та визначення найбільш критичних параметрів, які впливають на ефективність. Запропонована концепція спільного навчання кодувальника та декодувальника під стеганографічним контролем розширює модельні уявлення про можливості автоенкодерів у задачах приховування інформації.

Рекомендаціями та напрямами подальших досліджень вважається застосування та впровадження методів adversarial training з метою підвищення надійності декодування в умовах активних атак. В якості розширення архітектур пропонується використання 3D-згорткових автоенкодерів для відео або рекурентних мереж для аудіо, що покращить результати у спеціалізованих сценаріях.

Застосування глибоких автоенкодерів дозволило істотно підвищити якість приховування інформації при мінімальному погіршенні якості носія та забезпечити вищу пропускну здатність стеганографічного каналу. У порівнянні з класичними методами, модель продемонструвала вищу стійкість до атак і здатність до генералізації. Таким чином, мету дослідження досягнуто в повному обсязі.

Результати дослідження можуть бути використані в різних галузях інформаційної безпеки, зокрема як: цифрові водяні знаки для захисту авторських прав; приховане передавання повідомлень у каналах зв'язку, стійких до перехоплення; безпека IoT-пристроїв, де важливо мінімізувати обсяг даних і водночас зберегти секретність. Запропонована методика може бути додатково вдосконалена шляхом інтеграції криптографічного шифрування прихованого повідомлення, що сприятиме підвищенню загального рівня безпеки.

Список літератури

1. Zhu J., Kaplan R., Johnson J., Fei-Fei L. HiDDeN: Hiding data with deep networks. Advances in Neural Information Processing Systems. 2018. Vol. 31.
2. Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016.
3. Zhou C., Paffenroth R. Anomaly detection with robust deep autoencoders. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017. P. 665–674. URL: http://doi.org/10.1145/3097983.3098052.
4. Hayes J., Danezis G. Generating steganographic images via adversarial training. Advances in Neural Information Processing Systems. 2017. Vol. 30.
5. Montgomery D. C. Design and Analysis of Experiments. John Wiley & Sons, 2017.
6. Gupta S., Joshi R. C., Misra M. A. SeGAN: Segment-based image steganography using generative adversarial network. IET Image Processing. 2019. Vol. 13, No. 10. P. 1706–1713. URL: http://doi.org/10.1049/iet-ipr.2018.6233.
7. Lu X., Li B., Huang J. GAN-based image steganography: review and research trends. IEEE Access. 2019. Vol. 7. P. 179097–179110. URL: http://10.1109/ACCESS.2019.2958574.
8. Zhang R., Wang S., Wang L. Robust deep steganography with pixelwise adversarial training. Neural Computing and Applications. 2021. Vol. 33. P. 2357–2369. URL: http://doi.org/10.1007/s00521-020-05047-2.
9. Nissar A. U., Mir A. H. Classification of steganalysis techniques: A study. Digital Signal Processing. 2010. Vol. 20, No. 6. P. 1758–1770. URL: http://doi.org/10.1016/j.dsp.2010.01.017.
10. Li B., Luo X., Liu T., Huang J. A survey on image steganography and steganalysis. Journal of Information Hiding and Multimedia Signal Processing. 2011. Vol. 2, No. 2. P. 142–172.
11. Ker A. D. Steganalysis of LSB matching in grayscale images. IEEE Signal Processing Letters. 2005. Vol. 12, No. 6. P. 441–444. URL: http://doi.org/10.1109/LSP.2005.847889.
12. Petitcolas F. A. P., Anderson R. J., Kuhn M. G. Information hiding - a survey. Proceedings of the IEEE. 1999. Vol. 87, No. 7. P. 1062–1078. URL: http://doi.org/10.1109/5.771065.
13. Reddy A., Acharya N., Mandal J. K. A new approach to transform domain-based robust steganography using wavelet families. Arabian Journal for Science and Engineering. 2018. Vol. 43. P. 5079–5090. URL: http://doi.org/10.1007/s13369-018-3205-0.
14. Liu Z., Su Z., Hou D., Li H. A robust CNN-based method for image steganography and steganalysis. Multimedia Tools and Applications. 2018. Vol. 77. P. 21769–21785. URL: http://doi.org/10.1007/s11042-018-6089-1.
15. Tang S., Wu X. Adaptive steganography based on deep reinforcement learning. Neurocomputing. 2020. Vol. 370. P. 35–46. URL: http://doi.org/10.1016/j.neucom.2019.08.090.
16. Yu W., Chen S. Improved autoencoder-based image steganography. IEEE Access. 2021. Vol. 9. P. 41395–41406. URL: http://doi.org/10.1109/ACCESS.2021.3064091
17. Duan Y., Yang B., Gao H. Deep hiding in video frames with convolutional neural networks // Information Sciences. 2021. Vol. 551. P. 27–43. URL: http://doi.org/10.1016/j.ins.2020.11.038.
18. Мельник К.В., Мельник В.М., Коптюк Ю.Ю. Дослідження методів розпізнавання зображень на основі нейронних мереж. *Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво"*. Луцьк, 2019. Вип. № 35. С. 161-165.
19. Яровий А.А., Кашубін С.Г., Кулик О.О. Розпізнавання мімічних мікровиразів обличчя людини. *Системи технічного зору та штучного інтелекту з обробкою та розшаруванням зображень*. Вінниця: ВНТУ, 2015. С.76-83.

References

1. Zhu J., Kaplan R., Johnson J. & Fei-Fei L. (2018). *HiDDeN: Hiding data with deep networks.* Advances in Neural Information Processing Systems. Vol. 31.
2. Goodfellow I., Bengio Y. & Courville A. (2016). *Deep Learning*. MIT Press.
3. Zhou C. & Paffenroth R. (2017). *Anomaly detection with robust deep autoencoders.* Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. P. 665–674. http://doi.org/10.1145/3097983.3098052.
4. Hayes J. & Danezis G. (2017). *Generating steganographic images via adversarial training.* Advances in Neural Information Processing Systems. Vol. 30.
5. Montgomery D. C. (2017). *Design and Analysis of Experiments*. John Wiley & Sons.
6. Gupta S., Joshi R. C. & Misra M. A. (2019). *SeGAN: Segment-based image steganography using generative adversarial network.* IET Image Processing. Vol. 13, No. 10. P. 1706–1713. http://doi.org/10.1049/iet-ipr.2018.6233.
7. Lu X., Li B. & Huang J. (2019). *GAN-based image steganography: review and research trends.* IEEE Access. Vol. 7. P. 179097–179110. http://10.1109/ACCESS.2019.2958574.
8. Zhang R., Wang S. & Wang L. (2021). *Robust deep steganography with pixelwise adversarial training.* Neural Computing and Applications. Vol. 33. P. 2357–2369. http://doi.org/10.1007/s00521-020-05047-2

9. Nissar A. U. & Mir A. H. (2010). *Classification of steganalysis techniques: A study*. Digital Signal Processing. Vol. 20, No. 6. P. 1758–1770. http://doi.org/10.1016/j.dsp.2010.01.017.
10. Li B., Luo X., Liu T. & Huang J. (2011). *A survey on image steganography and steganalysis*. Journal of Information Hiding and Multimedia Signal Processing. Vol. 2, No. 2. P. 142–172.
11. Ker A. D. (2005). *Steganalysis of LSB matching in grayscale images*. IEEE Signal Processing Letters. Vol. 12, No. 6. P. 441–444. http://doi.org/10.1109/LSP.2005.847889.
12. Petitcolas F. A. P., Anderson R. J. & Kuhn M. G. (1999). *Information hiding - a survey*. Proceedings of the IEEE. Vol. 87, No. 7. P. 1062–1078. http://doi.org/10.1109/5.771065.
13. Reddy A., Acharya N. & Mandal J. K. (2018). *A new approach to transform domain-based robust steganography using wavelet families*. Arabian Journal for Science and Engineering. Vol. 43. P. 5079–5090. http://doi.org/10.1007/s13369-018-3205-0.
14. Liu Z., Su Z. & Hou D., Li H. (2018). *A robust CNN-based method for image steganography and steganalysis*. Multimedia Tools and Applications. Vol. 77. P. 21769–21785. http://doi.org/10.1007/s11042-018-6089-1.
15. Tang S. & Wu X. (2020). *Adaptive steganography based on deep reinforcement learning*. Neurocomputing. Vol. 370. P. 35–46. http://doi.org/10.1016/j.neucom.2019.08.090.
16. Yu W. & Chen S. (2021). *Improved autoencoder-based image steganography*. IEEE Access. Vol. 9. P. 41395–41406. http://doi.org/10.1109/ACCESS.2021.3064091
17. Duan Y., Yang B., & Gao H. (2021). *Deep hiding in video frames with convolutional neural networks* // Information Sciences. Vol. 551. P. 27–43. http://doi.org/10.1016/j.ins.2020.11.038.
18. Melnyk, K.V., Melnyk, V.M., & Koptyuk, Y.Y. (2019). Research of image recognition methods based on neural networks. *Scientific journal «Computer-integrated technologies: education, science, production»*. Lutsk. Issue No. 35. C. 161-165 [in Ukrainian].
19. Yarovyi, A.A., Kashubin, S.G., & Kulyk, O.O. (2015). Recognition of mimic microexpressions of the human face. *Systems of technical vision and artificial intelligence with image processing and stratification*. Vinnytsia: VNTU. C.76-83 [in Ukrainian].

Dmytro Prokopovych-Tkachenko[1], Assoc. Prof., PhD tech. sci.,
Liudmyla Rybalchenko[1], Assoc. Prof., PhD econ. sci., **Yaroslav Derkach**[2]
[1]*University of Customs and Finance, Dnipro, Ukraine*
[2]*Kharkiv National University of Radio Electronics, Kharkiv, Ukraine*

Steganographic Methods in Information Security

The article discusses the possibilities of using deep autoencoders in the field of information hiding (steganography). It is shown that the combination of steganography methods with deep learning makes it possible to improve system reliability and increase the bandwidth of the hidden data transmission channel.

The purpose of this paper is to analyse modern approaches to the use of deep autoencoders in steganography, to determine their main advantages and disadvantages, and to formulate promising directions for further research. The development of deep learning opens up broad prospects for enhancing classical steganographic approaches. The use of deep neural networks allows increasing the bandwidth of the steganographic concealment channel, increase resistance to attacks, in particular, to compression, noise and other distortions and adapt to different types of media and specific security requirements will allow identifying new tools for optimising the process of information concealment. The scientific novelty lies in the combination of the ideas of steganography and deep auto-encoders, which makes it possible to obtain high quality recovery of hidden information with minimal visual, acoustic or other artefacts.

A comparative review of modern autoencoder architectures is carried out, the principles of encoding and decoding are analysed, and the results of experimental studies demonstrating the effectiveness of the proposed approaches are summarised. The prospects for the development of this area in terms of security, efficiency and resistance to attacks are assessed through a detailed analysis of potential vulnerabilities and practical implementation scenarios.

The results of the study indicate the significant potential of deep autoencoders in the field of information security, in particular for integration with steganographic methods. A number of recommendations for further improvement of the technology are proposed, including optimisation of the neural network architecture, expansion of the scope of applications, and consideration of ethical and legal aspects.

The results of the study can be used in various areas of information security, including: digital watermarks for copyright protection; covert transmission of messages in communication channels resistant to interception; security of IoT devices, where it is important to minimise the amount of data and at the same time maintain secrecy. The methodology itself can be improved by integrating additional encryption at the level of the hidden message, which will increase the overall level of security.

deep autoencoders, steganography, information security, deep learning, neural networks

Одержано (Received) 21.03.2025 *Прорецензовано (Reviewed) 31.03.2025*
Прийнято до друку (Approved) 22.04.2025