Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Факультет комп’ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Метод стеганоаналізу на основі нейронних мереж

(тема)

Виконав:

студент II курсу, групи СПм-23-1

Чепурних М.А.
(прізвище, ініціали)

Спеціальність

123 «Комп’ютерна інженерія»
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма

Системне програмування
(повна назва освітньої програми)

Керівник: доц. Бологова Н.М.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ _____ Коваленко А.А.
(підпис) (прізвище, ініціали)

2025 р.

Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління

Кафедра електронних обчислювальних машин

Рівень вищої освіти другий (магістерський)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Системне програмування
(повна назва)

ЗАТВЕРДЖУЮ:
Зав.
кафедри _____
(підпис)

"____" _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Чепурних Максиму Андрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Метод стеганоаналізу на основі нейронних мереж

затверджена наказом по університету від " 22 " листопада 2024 р. № 1236 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 20 січня 2025 р.

3. Вхідні дані до роботи

4. Перелік питань, що потрібно опрацювати у роботі

1) огляд традиційних методів стеганоаналізу для виявлення стеганографічних схем;
2) вибір та обгрунтування методики та засобів дослідження;
3) програмна реалізація моделей протоколів;
4) проведення експериментальних досліджень;
5) висновки.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 14 слайдів

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Огляд традиційних методів стеганоаналізу	26.11.24-30.11.24	
2	Вибір та обгрунтування методики дослідження	02.12.24-05.12.24	
3	Вибір архітектури нейронної мережі	06.12.24-10.12.24	
4	Розробка концептуальної архітектури	11.12.24-21.12.24	
	глибокої нейронної мережі		
5	Проведення експериментів	23.12.24-03.01.25	
6	Оформлення матеріалів кваліфікаційної роботи	04.01.25-07.01.25	
7	Подання кваліфікаційної роботи керівникові та	08.01.25-11.01.25	
	її попередній захист		
8	Подання кваліфікаційної роботи на рецензу-	13.01.25-17.01.25	
	вання		

Дата видачі завдання 22 листопада 2024 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Бологова Н.М.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 55 с., 15 рис., 2 табл., 1 дод., 21 джерел.

ДКП, HUGO, JPEG, MIPOD, S-UNIWARD, СТЕГАНОАНАЛІЗ.

Метою кваліфікаційної роботи є розробка та аналіз ефективності нейронних мереж у контексті стеганоаналізу, а також порівняння їх результатів з традиційними методами виявлення прихованої інформації.

У ході виконання кваліфікаційної роботи розглядається метод стеганоаналізу, заснований на використанні нейронних мереж, зокрема на глибоких нейронних мережах для виявлення стеганографічних даних у цифрових зображеннях.

У цій кваліфікаційній роботі досліджено статистичний стеганоаналіз зображень. Спершу описано основи стеганографії та схеми вбудовування інформації, зокрема метод із використанням Судоку. Запропоновано схему, яка покращує якість стего-зображень за допомогою еталонної матриці.

Результатом роботи є розробка архітектуру глибокої нейронної мережі (ГНМ) із навчальним шаром високочастотної фільтрації, що покращує результати стеганоаналізу. Ця архітектура перевершує традиційні детектори (Xu-Net, Ye-Net, Yedroudj-Net) на великому наборі даних. У перспективі планується застосування ГНМ для аналізу в домені JPEG із можливістю використання трансферного навчання та оптимізації обчислювальних витрат.

ABSTRACT

Master’s thesis: 55 pages, 15 figures, 2 tables, 1 appendices, 21sources.

DCP, HUGO, JPEG, MIPOD, S-UNIWARD, STEGANOANALYSIS.

The major goal of this thesis is the qualification work is to develop and analyze the effectiveness of neural networks in the context of steganalysis, as well as to compare their results with traditional methods of detecting hidden information.

In order to the qualification work, the steganalysis method based on the use of neural networks, in particular, deep neural networks, for detecting steganographic data in digital images is considered.

In this qualification work, statistical image steganalysis is investigated. First, we describe the basics of steganography and information embedding schemes, including the Sudoku method. Then, a scheme is proposed that improves the quality of stego-images using a reference matrix.

The result of the work is the development of a deep neural network (DNN) architecture with a high-frequency filtering training layer that improves the results of steganalysis. This architecture outperforms traditional detectors (Xu-Net, Ye-Net, Yedroudj-Net) on a large dataset. In the future, it is planned to apply the GNM for analysis in the JPEG domain with the possibility of using transfer learning and optimizing computational costs.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ДКП – коефіцієнти дискретного косинусного перетворення

HUGO – стеганографічний метод, який оптимізований для того, щоб зміни у зображенні залишалися максимально непомітними для детекційних алгоритмів (англ., Highly Undetectable Steganography)

JPEG –формат стиснення зображень (англ., Joint Photographic Experts Group)

LSB – найменші значущі біти (англ., Least Significant Bits)

MiPOD – метод, спрямований на мінімізацію ймовірності виявлення прихованої інформації (англ., Minimizing the Power of Detection)

SVM – машина векторів підтримки (англ., Support Vector Machines

S-UNIWARD – універсальний адаптивний метод, що зменшує помітність (англ., Spatial-UNIversal WAvelet Relative Distortion)

WOW – метод, який використовує ваги, отримані за допомогою хвильового перетворення (англ., Wavelet Obtained Weights)

ВСТУП

У сучасному світі, де інформація є однією з найцінніших ресурсів, забезпечення її конфіденційності та захисту від несанкціонованого доступу стає надзвичайно важливим завданням. Окрім традиційних методів шифрування, все більшої популярності набуває стеганографія – технологія прихованого передавання інформації, яка дозволяє вмонтовувати секретні дані в різні носії, зокрема в зображення, аудіо чи відео. Однак, через поширення цих методів зловмисниками для приховування незаконної інформації, виникає потреба в ефективних способах виявлення стеганографічних даних, що призводить до розвитку науки стеганоаналізу.

Стеганоаналіз, як наука, має за мету розпізнавання прихованих даних у носіях інформації, зокрема через дослідження аномалій у цифрових зображеннях, аудіофайлах або відео. Однак традиційні методи стеганоаналізу, засновані на математичних та статистичних підходах, можуть бути недостатньо ефективними для виявлення прихованих даних у складних умовах, зокрема у випадках, коли використовуються сучасні стеганографічні техніки з адаптивними або обфускованими методами.

Нейронні мережі, зокрема глибоке навчання, набули значного розвитку за останні роки і продемонстрували високу ефективність у розв'язанні завдань, що пов'язані з аналізом складних структур даних. Використання нейронних мереж для стеганоаналізу дає змогу виявляти приховану інформацію навіть у складних умовах, коли традиційні методи не справляються. Моделі глибокого навчання можуть автоматично навчатися на великих обсягах даних і виявляти тонкі особливості, які є важливими для виявлення стеганографічних прихованих даних.

У межах цієї магістерської роботи розглядається метод стеганоаналізу, заснований на використанні нейронних мереж, зокрема на глибоких нейронних мережах для виявлення стеганографічних даних у цифрових

зображеннях. Метою дослідження є розробка та аналіз ефективності нейронних мереж у контексті стеганоаналізу, а також порівняння їх результатів з традиційними методами виявлення прихованої інформації.

1 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ

Стеганографія - це мистецтво і наука таємного зв'язку, метою якого є приховування секретних повідомлень у нешкідливих на вигляд об'єктах прикриття. В даний час більшість стеганографічних схем розробляються в області зображень, оскільки вони прості в реалізації і здатні приховувати секретну інформацію у всіх на виду, а головне - мають велику кількість можливостей для вбудовування в область зображень. З іншого боку, стеганоаналіз - це мистецтво і наука виявлення того, чи приховує даний об'єкт секретну інформацію. Стеганоаналіз зображень базується на припущенні, що достатньо малі зміни в зображеннях можуть призвести до помітних змін у деяких статистичних показниках зображення. Виявлення цих змін є основою стеганалізу зображень.

1.1 Приховування інформації

Почнемо з короткого обговорення різниці між стеганографією та криптографією. Взагалі кажучи, і стеганографія, і криптографія - це приховування інформації в сенсі таємного спілкування. Критична різниця між ними полягає в тому, що перша приховує існування секретної інформації, тоді як друга приховує зміст секретної інформації, що призводить до зовсім різних схем приховування інформації, а також схем її виявлення. У цьому розділі розглянемо стеганографію і зосередимося на її виявленні.

Стеганографія має довгу історію, і гарний вступ до неї можна знайти в [1]. Сучасна стеганографія базується на моделі з відомої задачі про ув'язнених [2]. Де, Аліса і Боб є ув'язненими, і їм дозволено спілкуватися, за чим спостерігає наглядач. Якщо наглядач виявить, що для передачі секретної інформації використовуються стеганографічні

схеми, то Аліса і Боб можуть бути покарані. У цьому випадку вважається, що наглядач успішно виявив стеганографію. Одне з припущень у цій моделі полягає у тому, що наглядач досконало знає схеми приховування інформації, які можуть застосувати Аліса та Боб. Це відомий принцип Керкхоффа. Варто зазначити, що цей принцип все ще широко використовується при вивченні стеганоаналізу, хоча він є менш практичним. Також широко досліджується сліпий стеганоаналіз, який послаблює принцип за рахунок ефективності виявлення.

У стеганографії зображень найменші значущі біти (LSB) значень пікселів зображення або коефіцієнтів дискретного косинусного перетворення (ДКП) (для зображень у форматі JPEG) злегка модифікуються стеганографічними методами з надією на те, що ці модифікації не будуть виявлені. По суті, ці модифікації дуже малі, і значення пікселів або коефіцієнтів ДКП змінюються максимум на одиницю. Тут слід звернути увагу на два моменти. По-перше, як змінюється значення даного пікселя або коефіцієнта ДКП, якщо ми знаємо, наприклад, що значення (яке є цілим числом) змінюється на одиницю. Підхід з використанням LSB-заміни полягає в тому, щоб просто перевернути LSB значення, а підхід з використанням LSB-відповідності - у випадковій зміні значення на одиницю. Перший підхід призводить до асиметрії, яку можна використати для побудови ефективних детекторів. Друга проблема полягає в тому, як визначити набір пікселів або коефіцієнтів, які використовуються для вбудовування інформації. До 2010 року найпопулярнішим підходом був випадковий вибір цих пікселів або коефіцієнтів за допомогою випадкових насінин. Такий вибір є простим і схема легко реалізується, що вперше було представлено в роботі [3]. Але схема вбудовування не залежить від вмісту зображення і тому є менш безпечною. Адаптивна до вмісту схема вбудовування стала популярною після того, як був запропонований інший підхід [4]. Така адаптивна до вмісту схема вбудовування полягає в мінімізації функцій спотворення таким чином, що гладкі області зображень

рідше обираються для вбудовування. Іншими словами, зашумлені області, такі як краї, мають високу ймовірність бути використаними для вбудовування. Було запропоновано багато стеганографічних методів, заснованих на адаптивній схемі вбудовування, таких як HUGO, WOW, S-UNIWARD і MiPOD [5-7].

1.2 Виявлення інформації

Статистичний стеганоаналіз - це дослідження виявлення інформації за допомогою статистичних інструментів. Спотворення зображень, спричинені ретельно розробленими стеганографічними методами, майже не помітні людським оком. Для їх виявлення необхідно використовувати статистичні методи. Як було показано вище, ця задача є задачею бінарної класифікації. А саме, треба вирішити, чи є дані зображення обкладинкою (чистими) або стего (брудними) зображеннями. В аналізі зображень розглядаються два підходи: перевірка статистичних гіпотез і машинне навчання/глибоке навчання. Перший базується на моделях зображень обкладинки, а другий - на вилученні ознак.

У рамках перевірки гіпотез припускається, що значення пікселів або коефіцієнтів ДКП мають розподіл, подібний до узагальненого гаусівського. Вбудовування, проведене стеганографічними методами, може бути змодельоване як зміна параметрів моделі/розподілу. Слід зазначити, що в моделі розподілу використовується припущення i.i.d. або незалежне припущення. Це непрактично, але модель, здається, добре працює для певних стеганографічних методів, як показано експериментами в наведених вище посиланнях. Однак, з моєї точки зору, цей фреймворк має труднощі з адаптивними до вмісту стеганографічними методами. Більш поширеним є використання методів машинного навчання/глибокого навчання для складних стеганографічних методів.

Машинне навчання природно застосовується в стеганоаналізі,

оскільки виявлення секретної інформації розглядається як завдання класифікації. Лю та ін. [8] вперше використали вейвлет-статистику (як ознаки) та машину векторів супортів (SVM) для стеганалізу зображень. Після цієї новаторської роботи стеганоаналіз на основі машинного навчання привернув до себе велику увагу. З того часу було виконано багато робіт зі стеганоаналізу, зосереджених на вилученні ознак. Для стеганоаналізу з використанням фреймворку машинного навчання основна увага приділяється вилученню ознак. Незважаючи на те, що методи класифікації є дуже важливими у стеганоаналізі зазвичай використовується добре розроблені методи класифікації (наприклад, SVM, ансамблеві методи) для вирішення поставленої задачі. Крім того, для стеганоаналізу широко використовуються методи зменшення розмірності, що працюють з високорозмірним простором ознак, такі як PCA, випадкова проекція. До 2015 року стеганоаналіз зображень був значною мірою пов'язаний з пошуком ручної роботи, яка значною мірою покладалася на знання предметної області. Багато дослідників докладали багато зусиль протягом багатьох років, і врешті-решт була зібрана колекція з більш ніж 30 000 ознак. Модель з такими характеристиками називається повною моделлю [9]. Зауважте, що повна модель має високорозмірний простір ознак, що може вимагати багато пам'яті та значних обчислень під час навчання.

Глибоке навчання - це підгалузь машинного навчання, яка використовує глибокі нейронні мережі для автоматичного вилучення ієрархічних ознак у процесі навчання. Завдяки успіху глибоких нейронних мереж у багатьох завданнях, таких як комп'ютерний зір, глибоке навчання було впроваджено в стеганоаналіз зображень у 2015 році Qian та ін. [10]. Результати, отримані ними в статті, були багатообіцяючими, хоча метод з використанням глибоких нейронних робіт не дотягував за продуктивністю до ансамблевого методу з багатою моделлю.

2 СТЕГАНОГРАФІЯ ЗОБРАЖЕНЬ

Стеганографія - це мистецтво і наука приховування існування секретної інформації, яка вбудовується в об'єкти прикриття таким чином, щоб отримані стегооб'єкти не викликали підозр. Зображення широко використовуються як об'єкти прикриття в стеганографічних методах, хоча інші типи об'єктів, що містять надлишковість, також можуть бути застосовані. Для зображень найпоширеніші стеганографічні методи ґрунтуються на зміні БГП. Більш конкретно, стеганографічні методи змінюють LSB значень пікселів для просторових зображень, і змінюють LSB коефіцієнтів перетворення (наприклад, коефіцієнтів DCT) для зображень частотної області (наприклад, зображень JPEG). Тобто, секретна інформація, представлена рядком двійкових цифр, ховається в зображенні шляхом модифікації LSB пікселів або коефіцієнтів зображення.

2.1 Заміна LSB проти узгодження

Розглянемо простий приклад, який описує заміну та зіставлення LSB. Шість значень пікселів, або коефіцієнтів вибираються для вбудовування секретного повідомлення, закодованого у вигляді двійкового рядка 011010. Вважається, що ці шість значень дорівнюють 7, 7, 8, 6, 5, 5. Аліса (відправник), можливо, змінить значення таким чином, що LSB модифікованих шести значень будуть відповідати секретному двійковому рядку 011010. LSB початкових шести значень дорівнюють 110011. Легко помітити, що перше значення 7, третє значення 8 і останнє значення 5 потрібно змінити. Для заміни LSB просто поміняємо місцями LSB трьох значень, і вони стануть 6, 9 і 4. Шість значень, які отримає Боб (одержувач), будуть 6, 7, 9, 6, 5, 4, і він зможе легко витягти секретний двійковий рядок

011010 з LSB. Слід зазначити, що Боб знає, звідки він може отримати ці шість значень за принципом Керкгоффа. Зверніть увагу, що існує закономірність для заміни LSB. Тобто, непарне значення завжди зменшується і парне значення завжди збільшується, коли застосовується заміна LSB. На основі цієї закономірності було запропоновано багато детекторів, таких як аналіз пар зразків, структурний стеганоаналіз та метод зваженого стего

Вирівнювання LSB спрямоване на усунення асиметрії шляхом випадкового зменшення або збільшення значень на одиницю. Знову ж таки, розглянемо приклад вище. Перше значення 7 можна було випадковим чином змінити на 8 або 6. Той самий підхід застосовується до двох інших значень. Зауважте, що значення змінюється на одиницю за допомогою LSB-порівняння, але у змінах може бути задіяно більше LSB.

2.2 Випадкові та контентно-адаптивні схеми вбудовування

Як згадувалося раніше, вибір набору пікселів або коефіцієнтів, які використовуються для вбудовування інформації (наприклад, заміна або співпадіння LSB), є важливою темою в стеганографії зображень. Для схеми випадкового вбудовування Аліса і Боб можуть обмінюватися випадковими пікселями (за принципом Керкгоффа), так що вони обидва знають набір пікселів або коефіцієнтів. Бобу легко витягти секретну інформацію, закодовану двійковим рядком. Як показано у прикладі вище, Боб просто переглядає LSB пікселів з набору у певному порядку, використовуючи, наприклад, випадкові пікселі. Недоліком випадкового вбудовування є те, що воно не залежить від вмісту зображення. Це може призвести до значних спотворень на гладких ділянках зображення. Щоб вирішити цю проблему, було запроваджено адаптивну до вмісту схему вбудовування шляхом мінімізації функції вартості спотворення.

Розглянемо зображення обкладинки у відтінках сірого, позначене

$X = \{x_1, x_2, ..., x_N\}$. Тут ігноруємо кореляцію між пікселями для простоти і використовуємо послідовність для двовимірного зображення. Позначимо її стего-зображення через $Y = \{y_1, y_2, ..., y_N\}$, де x_i , $y_i \in \{0, 1, ..., 255\}$, i = 1, 2, ..., N. Для оцінки спотворень вбудовування вводиться функція вартості спотворень $\rho_i(X, y_i)$. Задаємо Y так, щоб мінімізувати наступні спотворення

$$D(X,Y) = \sum_{i=1}^{N} \rho_i \mid x_i - y_i \mid, \tag{2.1}$$

де ρі - витрати на зміну хі на уі . Різні варіанти функції витрат ρ призводять до різних схем адаптивного вбудовування контенту, таких як HUGO, WOW, S-UNIWARD та MiPOD.

На рисунку 2.1 показано зображення обкладинки, її стего-зображення та різницю між ними. Тут застосовано LSB зіставлення з випадковою схемою вбудовування. Для порівняння, для цього ж зображення обкладинки використано адаптивну до контенту схему вбудовування, а саме S-UNIWARD, рисунок 2.3. З порівняння видно, що на рисунку 2.1 показано випадковість, а на рисунку 2.3 - адаптивність з точки зору місця вбудовування. Рисунок 2.2 демонструє, що схема вбудовування WOW також є адаптивною до контенту.

2.3 Стеганографія на основі судоку

Тут наведено конкретний приклад з використанням стеганографії. Хочу зазначити, що цей приклад не є адаптивним до контенту. Він показує, що стеганографія зображень може бути проведена за допомогою інших підходів.

Почнемо з опису стеганографічних методів приховування секретної інформації в зображеннях на основі Судоку. Рішення Судоку відіграють важливу роль у побудові еталонної матриці, на основі якої значення пікселів зображення обкладинки модифікуються для приховування секретної

інформації. Більш конкретно, припустимо, що маємо розв'язок Судоку, який можна записати у вигляді матриці 9 × 9 S. Нехай B(i, j) позначає запис в i-му рядку та j-му стовпчику матриці B. А еталонна матриця M, пов'язана з розв'язком Судоку S, будується за допомогою задання

$$M(i, j) = S(i', j') - 1, \;\; i' = i \bmod 9, \; j' = j \bmod 9. \qquad (2.2)$$

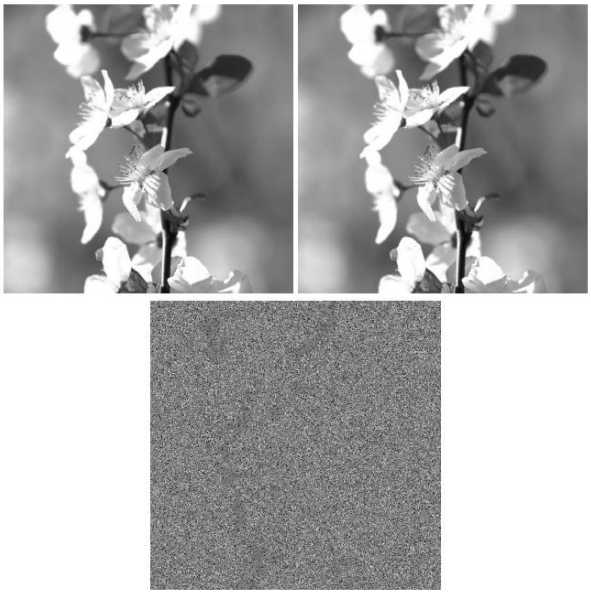

Рисунок 2.1 – Вбудовування: LSB Matching 0.4bpp; вгорі ліворуч: зображення обкладинки; вгорі праворуч: зображення стего; внизу: змінені пікселі: білий → 1, чорний → -1

Тут індекси рядків і стовпців матриці починаються з 0. Для того, щоб вбудувати секретну інформацію в задане зображення обкладинки у відтінках сірого, ми спочатку генеруємо послідовність пікселів, випадковим чином перебираючи всі пікселі обкладинки, а потім створюємо список L пар пікселів, що не перетинаються, просто з'єднуючи сусідні пікселі в послідовності. Оскільки значення пікселів зображення у відтінках сірого є цілими числами від 0 до 255, розмір опорної матриці M фіксований і становить 256 × 256. Нижче наведено розв'язок судоку.

Рисунок 2.2 – Вбудовування: WOW 0.4bpp; вгорі ліворуч: зображення обкладинки; вгорі праворуч: зображення стего; внизу: змінені пікселі: білий → 1, чорний → -1

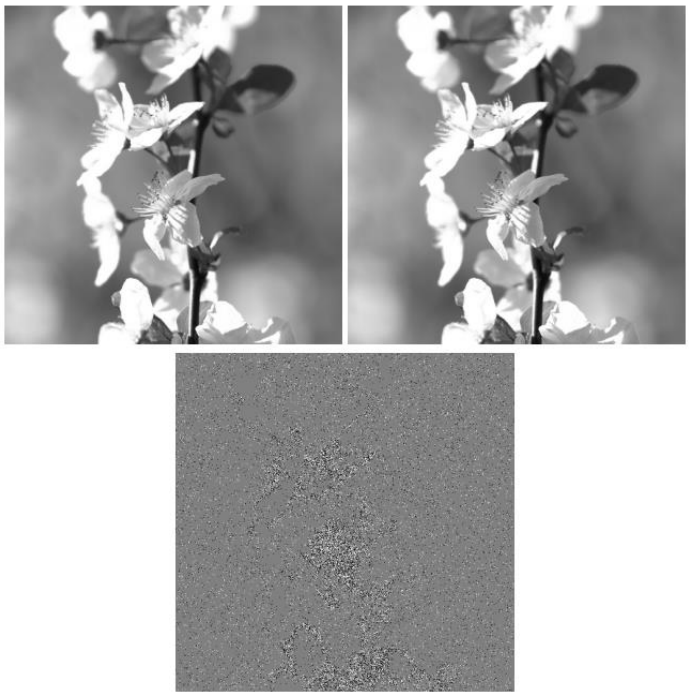

Рисунок 2. 3 – Вбудовування: S-UNIWARD; 0.4bpp; Вгорі ліворуч: зображення обкладинки; Вгорі праворуч: зображення стего; Внизу: змінені пікселі: білий → 1, чорний → -1

$$S = \begin{pmatrix} 7\,8\,2 & 4\,9\,1 & 5\,3\,6 \\ 1\,4\,6 & 5\,7\,3 & 9\,2\,8 \\ 5\,3\,9 & 6\,2\,8 & 7\,4\,1 \\ 3\,5\,8 & 1\,6\,4 & 2\,9\,7 \\ 4\,9\,1 & 7\,5\,2 & 8\,6\,3 \\ 6\,2\,7 & 3\,8\,9 & 4\,1\,5 \\ 2\,7\,5 & 9\,3\,6 & 1\,8\,4 \\ 8\,1\,3 & 2\,4\,5 & 6\,7\,9 \\ 9\,6\,4 & 8\,1\,7 & 3\,5\,2 \end{pmatrix} \quad (2.3)$$

Зверніть увагу, що кожен рядок або стовпчик у будь-якому розв'язку Судоку містить рівно дев'ять різних цифр від 1 до 9. Також існує дев'ять блоків 3×3, що не перетинаються, кожен з яких містить різні цифри від 1 до 9. Ця властивість розв'язків Судоку використовується для розробки алгоритмів пошуку. Секретний ключ, спільний для відправника та одержувача, складається з пікселів, еталонної матриці та можливого символу, що вказує на кінець секретної інформації.

Припустимо, що секретна інформація, включаючи символ закінчення, може бути представлена послідовністю R секретних цифр у 9-ричній системі числення. Нехай r_i та (x_i , y_i) позначають i-ту секретну цифру у послідовності R та i-ту пару у списку L відповідно. Зауважимо, що кожна пара пікселів у списку L відповідає певному місцю в опорній матриці M. Процедура вбудовування виконується наступним чином:

- спочатку i ← 1;
- знайдіть розташування (u, v) в M таке, що M(u, v) = r_i і (u, v) найближче до (x_i , y_i) в нормі L1 (або L2);
- змініть пару пікселів (x_i , y_i) на (x_i , y_i) ← (u, v);
- i ← i + 1;
- повторити (i)-(iv), поки i > length(R) або i > length(L).

Використовуючи властивість розв'язку Судоку, пошук найближчого розташування (u, v) до (xi , yi) у нормі L1 (тобто відстані Манхеттена) у

наведеній вище процедурі вбудовування здійснюється над трьома блоками, що містять (x_i , y_i): горизонтальним блоком 1 × 9, вертикальним блоком 9 × 1 і блоком 3 × 3. Зокрема, у випадку $3 < x_i < 252$ і $3 < y_i < 252$, три вищевказані блоки

$$\begin{matrix} & & & M(x_i-4,y_i) & \\ (M(x_i,y_i-4)\ldots M(x_i,y_i+4)), & (& \ldots &), & \\ & & & M(x_i+4,y_i) & \end{matrix}$$
$$\begin{matrix} M(s,\ t) & M(s,\ t\ +\ 1) & M(s,\ t\ +\ 2) \\ (M(s\ +\ 1,\ t) & M(s\ +\ 1,\ t\ +\ 1) & M(s\ +\ 1,\ t\ +\ 2)) \\ M(s\ +\ 2,\ t) & M(s\ +\ 2,\ t\ +\ 1) & M(s\ +\ 2,\ t\ +\ 2) \end{matrix} \quad (2.3)$$

де s та t задовольняють $x_i = 3 - s + z$, $0 \le z < 3$, та $y_i = 3 - t + w$, $0 \le w < 3$, відповідно. В інших випадках нам може знадобитися зсунути горизонтальний або вертикальний блок.

Використовуючи секретний ключ, приймач може згенерувати список L пар пікселів з отриманого стего-зображення. Далі процедура вилучення виконується наступним чином:

- спочатку i ← 1;
- витягти i-ту пару пікселів (xi , yi) з L і отримайте i-ту секретну цифру M(xi , yi) з матриці M;
- i ← i + 1;
- повторити (i)-(iv) до тих пір, поки i > length(L) або не буде отримано символ завершення.

Далі покажемо, що спотворення, спричинені процедурою вбудовування, можна зменшити, замінивши опорну матрицю, побудовану з розв'язку Судоку, на нову, яку буде описано нижче. Розглянемо пару пікселів (u, v) у відтінках сірого на головному зображенні, де і u, і v є ненасиченими, тобто $0 < u, v < 255$. Зауважимо, що u та v можна збільшити або зменшити на два за допомогою алгоритму вбудовування, описаного раніше, якщо використовується опорна матриця, пов'язана з розв'язком Судоку. Наприклад,

припустимо, що пара пікселів і секретна цифра дорівнюють (3, 4) і 2 відповідно. Використовуючи покращений алгоритм пошуку, застосований у [11], та еталонну матрицю M, побудовану з (2.1) та (2.2), найближчими кваліфікованими розташуваннями до (3, 4) є (1, 5) та (5, 3) у нормі L2. Отже, v або u потрібно зменшити або збільшити на два, щоб приховати секретну цифру. Щоб ще більше зменшити спотворення, ми представляємо нову конструкцію еталонної матриці M, яка використовується в процедурі вбудовування. Стеганографічний метод, заснований на новій конструкції еталонної матриці, має таку саму здатність вбудовування, як і в [11], але досягає вищих значень PSNR, що буде продемонстровано за допомогою експериментальних результатів у наступному розділі. Нова конструкція еталонної матриці описується наступним чином. Спочатку ми створюємо матрицю T розміром 3 × 3, яка містить рівно дев'ять цілих чисел від 0 до 8. Еталонна матриця M, пов'язана з T, будується за допомогою задання

$$M(i, j) = T(i', j'), i' = i \bmod 3, j' = j \bmod 3, i, j = 0,1,...,255 \qquad (2.4)$$

Легко бачити, що існує 9! = 362880 різних варіантів для матриці T. Можемо довільно вибрати один з них. Для прикладу, нехай

$$T = \begin{pmatrix} 012 \\ 345 \\ 678 \end{pmatrix}. \qquad (2.5)$$

Таким чином, опорна матриця M, побудована на основі (2.4) та (2.5), має вигляд

$$M = \begin{pmatrix} 0 & 1 & 2 & 0 & 1 & 2 & 0 & \cdots & 0 \\ 3 & 4 & 5 & 3 & 4 & 5 & 3 & \cdots & 3 \\ 6 & 7 & 8 & 6 & 7 & 8 & 6 & \cdots & 6 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & \cdots & 0 \\ 3 & 4 & 5 & 3 & 4 & 5 & 3 & \cdots & 3 \\ 6 & 7 & 8 & 6 & 7 & 8 & 6 & \cdots & 6 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & \cdots & 0 \end{pmatrix} \tag{2.6}$$

Зауважте, що будь-який блок 3 × 3 у матриці M, побудований з (2.4), містить рівно дев'ять цілих чисел від 0 до 8. Це можна довести наступним чином. Припустимо, що у матриці M існує блок 3 × 3, який містить одне і те ж число у двох різних місцях. Позначимо через (p_1, q_1) і (p_2, q_2) два різних місця в матриці M відповідно. Отже, $M(p_1, q_1) = M(p_2, q_2)$. З конструкції (2.5) випливає, що M(p1, q1) = M(p2, q2) тоді і тільки тоді, коли

$$(p_1 - p_2) = 0 \; mod \; 3 \; and \; (q_1 - q_2) = 0 \; mod \; 3. \tag{2.7}$$

Це суперечить тому, що (p1, q1) і (p2, q2) знаходяться у блоці 3 × 3. Отже, всі блоки 3 × 3 у матриці M, побудованій з (2.3), мають різні цілі числа від 0 до 8.

Для ненасиченої пари пікселів (u, v) на обкладинці у відтінках сірого вищезгадана властивість матриці M гарантує, що u або v буде збільшено або зменшено не більше ніж на одиницю. Щоб переконатися у цьому, розглянемо блок 3 × 3 у матриці M у наступному прикладі

$$\begin{pmatrix} M(u-1, v-1) & M(u-1, v) & M(u-1, v+1) \\ M(u, v-1) & M(u, v) & M(u, v+1) \\ M(u+1, v-1) & M(u+1, v) & M(u+1, v+1) \end{pmatrix} \tag{2.8}$$

Оскільки блок вище містить всі дев'ять цифр від 0 до 8, то найближче кваліфіковане місце до (u, v) у нормі L2 повинно знаходитись у цьому блоці. Таким чином, для того, щоб приховати будь-яку з дев'яти цифр, потрібно змінити u або v максимум на одиницю. Як згадувалося раніше у прикладі, u або v у ненасиченій парі пікселів можуть бути збільшені або зменшені на два, якщо опорна матриця побудована з розв'язку Судоку за допомогою (2.1). Таким чином, нова конструкція (2.3) еталонної матриці M надає можливості для ненасичених пар пікселів зменшити спотворення, спричинені процедурою вбудовування. У запропонованому нами стеганографічному методі значення пікселів у насиченій парі пікселів можуть бути змінені на два. Але зауважте, що кількість насичених пікселів становить невелику частину від загальної кількості пікселів на більшості зображень природи. Покращення якості стегозображень можна очікувати за рахунок використання еталонної матриці M, побудованої з (2.4), як буде показано далі.

Процедура вбудовування та вилучення у запропонованому нами стеганографічному методі є такою ж, як і у попередньому підрозділі, за винятком того, що еталонна матриця M, побудована за формулою (2.1), замінюється на матрицю, побудовану за формулою (2.4).

Далі розглядаєтьсґ ефективність стеганографічного методу на основі Судоку, використаного в [11], та запропонованого нами стеганографічного методу. Для оцінки якості стего-зображень широко використовується показник PSNR (Peak Signal-to-Noise Ratio), який для стего-зображення у відтінках сірого N’ розміром r × c визначається як

$$PSNR = 10 \cdot log_{10} \frac{255^2}{MSE} \tag{2.9}$$

де MSE (середньоквадратична похибка) між стегозображенням N` та

відповідним зображенням в якому приховується інформація N визначається як

$$MSE = \frac{1}{rc} \sum_{i=0}^{r-1} \sum_{j=0}^{c-1} \left[N(i, j) - N'(i, j) \right]^2 \quad (2.10)$$

Дев'ять зображень розміром 512×512 вибрано з бази даних зображень USC-SIPI [12], серед яких Baboon, House, Lena, Peppers і Splash були перетворені у відтінки сірого рисунок 2.4.

Рисунок 2.4 – Дев'ять зображень у відтінках сірого з роздільною здатністю 512 × 512

Дев'ять зображень були використані як вер-зображення і протестовані стеганографічним методом на основі Судоку [13,14] та запропонованим нами

методом з використанням нових еталонних матриць. Всі три стеганографічні методи мають однакову здатність вбудовування, а саме $(\log_2 9)/2$ біт на піксель. При максимальній швидкості вбудовування послідовність секретних цифр, отриманих за допомогою генератора псевдовипадкових чисел, вбудовується в дев'ять зображень обкладинки відповідно трьома методами. Еталонна матриця M, що використовується в цьому експерименті, побудована на основі (2.1) і (2.2) для стеганографічного методу на основі Судоку в [13,14], а також для запропонованого методу. Порівняння ефективності методів наведено в таблиці 2.1 за показником PSNR.

Таблиця 2.1 – Порівняння ефективності трьох стеганографічних методів

Зображення	Метод в роботі [13]	Метод в роботі [14]	Запропонований метод
Aerial	47.4951	48.2037	49.7787
Baboon	47.4603	48.1418	49.7931
Boat	47.5021	48.2108	49.7286
Elaine	47.4814	48.2108	49.7198
House	47.4591	48.1058	49.7269
Lena	47.4950	48.1372	49.7055
Peppers	47.4758	48.1531	49.7434
Splash	47.4553	48.1166	49.6863
Truck	47.6078	48.2269	49.7171

Як видно з таблиці, стеганографічний метод на основі Судоку в [14] дає середнє значення PSNR на 0,64 дБ вище, ніж метод, отриманий в [13]. У той же час, PSNR запропонованого методу в середньому на 1.5 вищий, ніж PSNR, отриманий за допомогою методу в [15]. Експериментальні результати показують, що запропонований метод зменшує спотворення, спричинені вбудовуванням, порівняно зі стеганографічним методом на основі Судоку в [13, 14].

3 НЕЙРОННІ МЕРЕЖІ ДЛЯ СТЕГАНОАНАЛІЗУ

3.1 Загальні відомості про нейронні мережі

Стеганаліз зображень за допомогою системи перевірки гіпотез було обговорено в попередных розділах. Ця система перевірки гіпотез для стеганоаналізу значною мірою покладається на моделі зображення обкладинки (наприклад, неоднорідний гаусівський розподіл) і схему вбудовування (наприклад, LSB зіставлення з випадковим вбудовуванням). Кореляція або залежність пікселів в основному ігнорується для спрощення моделей. Крім того, адаптивну до вмісту схему вбудовування набагато важче інтегрувати з моделями обкладинки, ніж схему випадкового вбудовування. Для розробки ефективних детекторів для складних стеганографічних методів, таких як HUGO, WOW і S-UIWARD, люди вдаються до методів машинного навчання. Методи машинного навчання для стеганоаналізу показали дуже хороші результати [15, 16]. Оскільки велика модель, яка включає понад 30 000 ознак, була запропонована [17], підхід, який поєднує велику модель з ансамблевим навчанням, став гарним детектором для стеганоаналізу. Проте, ефективність детектора на основі машинного навчання сильно залежить від якості ручних ознак. Пошук корисних ручних ознак є складним завданням і потребує знань предметної області та великої кількості експериментів з використанням усіх видів методів машинного навчання. Після завершення роботи над великою моделлю, здається, потрібні нові підходи, які виходять за рамки вилучення «ручних» ознак.

Успішне застосування глибоких нейронних мереж у комп'ютерному зорі, обробці природної мови та інших завданнях змусило людей звернути увагу та розглянути можливість застосування нейронних мереж у своїх власних завданнях. Однією з головних переваг використання глибоких нейронних мереж є здатність до самонавчання/видобування ознак з даних під

час процесу навчання. Це зовсім інший підхід для вилучення ознак. Крім того, глибокі нейронні мережі мають дуже велику ємність, оскільки вони можуть легко мати тисячі і мільйони параметрів. Навіть поверхнева нейронна мережа може мати сотні тисяч параметрів. Така величезна ємність дозволяє нейронним мережам самонавчатися складним структурам/особливостям з даних і досягати видатної продуктивності в багатьох завданнях. Як правило, нейронні мережі поєднують в собі вилучення ознак і класифікацію та забезпечують комплексне рішення для стеганоаналізу. Більшість шарів в архітектурі нейронних мереж зазвичай використовуються для вилучення ознак, а останній шар - для класифікації. Проблеми використання глибоких нейронних мереж включають збір даних, розробку архітектури нейронних мереж, обчислення тощо. Дослідники з різних дисциплін досягли значного прогресу у вирішенні цих проблем.

Для стеганалізу зображень фреймворк глибокого навчання пропонує новий підхід, що виходить за рамки ручної роботи. Що ще важливіше, нас може мотивувати прогрес глибинного навчання для інших завдань, особливо для комп'ютерного зору. У 2014 році Тан та ін. [18] зробили спробу застосувати стекові автокодери для стеганоаналізу, хоча результати виглядали не дуже добре. У 2015 році Цянь та ін. [18] запропонували архітектуру згорткової нейронної мережі (CNN), яка показала багатообіцяючі результати. Це вважається ранньою спробою використання глибокого навчання для стагнаналізу зображень. Їхній дизайн архітектури CNN був заснований на тих принципах, що застосовуються в комп'ютерному зорі. Єдиним винятком є те, що Qian та ін. [18] на початку розробки архітектури CNN використовували фіксований фільтр високих частот. Фільтр високих частот 5×5 визначається наступним чином.

$$F^{(0)} = \frac{1}{12}\begin{pmatrix} -1 & 2 & -2 & 2 & -1 \\ 2 & -6 & 8 & -6 & 2 \\ -2 & 8 & -12 & 8 & -2 \\ 2 & -6 & 8 & -6 & 2 \\ -1 & 2 & -2 & 2 & -1 \end{pmatrix}. \quad (3.1)$$

Це KV-фільтр високих частот, який використовується в [18] під назвою квадратний фільтр S5a. В основному, цей фільтр високих частот розглядається як такий, що імітує процес виділення ознак. В [18] стверджується, що цей фільтр високих частот допомагає покращити продуктивність нейронної мережі. Варто зазначити, що цей фільтр є симетричним і сума всіх елементів дорівнює нулю. Цей модуль використовується для вилучення стегоінформації, прихованої в зображеннях. Зауважте, що секретна інформація вбудовується в LSB значень пікселів або коефіцієнтів DCT, щоб уникнути значних спотворень. Крім того, як згадувалося раніше, адаптивні до вмісту схеми вбудовування з великою ймовірністю приховують секретну інформацію в областях зображень, які не є гладкими, наприклад, лініях. З цих причин вважається, що стегосигнали ховаються в рівні шуму зображень.

Для того, щоб краще зрозуміти архітектуру CNN, повинні ввести деякі термінологію.

На рисунку 3.1 показано базову архітектуру нейронної мережі, яка складається з вхідного шару, двох прихованих шарів та вихідного шару. Кожен шар має групу нейронів, за винятком вихідного шару (в цій архітектурі). Нейрони в сусідніх шарах повністю з'єднані, як показано на 3.1. Це дійсно повністю з'єднані шари. У нашому випадку вхід, який є 16-вимірним вектором, рухається вперед, щоб пройти перший прихований шар. Обчислюється добуток матриці та вектора, і на виході отримуємо 12-вимірний вектор. Після проходження другого прихованого шару обчислювальний процес повторюється, і на виході отримуємо 10-мірний

вектор. Цю концепцію можна узагальнити на архітектуру CNN. З точки зору CNN, кожен нейрон у прихованих шарах вважається фільтром з певними розмірами (наприклад, 3 × 3, 5 × 5). Приховані шари в цьому випадку фактично називаються згортковими шарами. Кожна вхідна одиниця розглядається як канал зображення. Наприклад, RGB-зображення має три канали, а зображення у відтінках сірого - лише один. Коли вхідне зображення проходить через перший згортковий шар, відбувається згортка, і результат називається картою ознак, яка в нашому випадку є набором з 12 зображень. Цей процес повторюється, коли вхідне зображення рухається вперед шар за шаром.

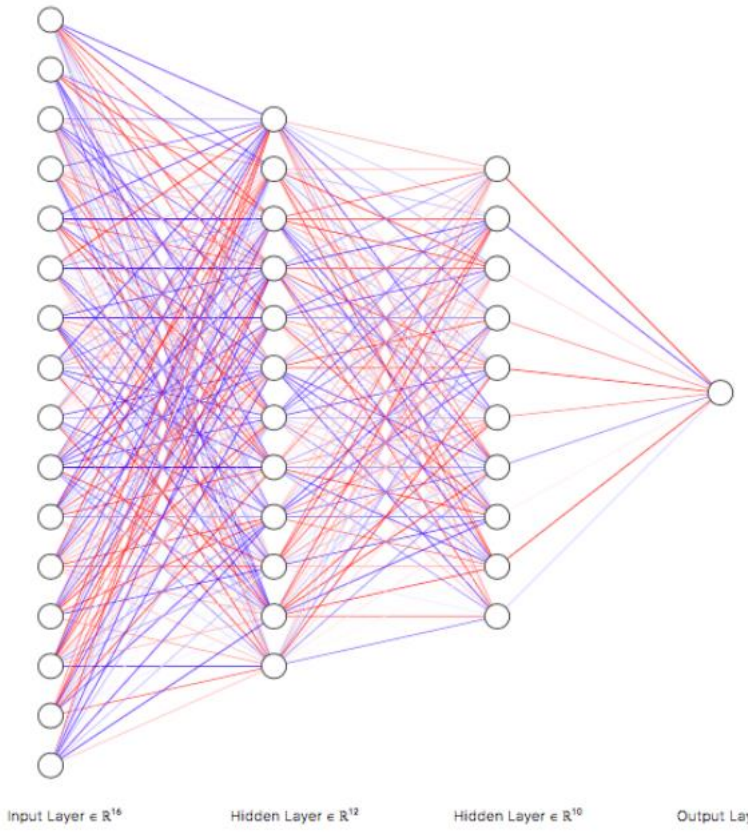

Рисунок 3.1 – Проста нейронна мережа

Далі використовуємо математичні позначення для опису згортки. Позначення запозичено з [18]. Нехай $I^{(0)}$ - відфільтроване зображення, яке є

результатом проходження через фіксований фільтр високих частот (3.1). Нехай $F_k^{(l)}$ позначає k-й фільтр з шару l = {1, ..., L}, де L - кількість шарів згортки, і k ∈ {1, ..., K(l)}, де K(l) - кількість фільтрів l-го шару. Згортка з першого шару з k-м фільтром дає відфільтроване зображення, позначене через $\overline{I}_k^{(l)}$, таке, що

$$\overline{I}_k^{(l)} = \mathrm{I}^{(0)} * F_k^{(l)} \tag{3.2}$$

Починаючи з другого і до останнього шару згортки, згортка є менш традиційною, оскільки на вхід подається K(l-1) карт ознак (тобто K(l-1) зображень), позначених через $I_k^{(l)}$ з k = 1, ..., K(l-1). Зауважте, що згортка, яка призведе до k-го зображення фільтра $\overline{I}_k^{(l)}$ в результаті згорточного шару l, насправді є сумою K(l-1) операцій згортки. Тобто

$$\overline{I}_k^{(l)} = \sum_{i=1}^{K^{(i-1)}} I_i^{(l-1)} * F_{k,i}^{(l)} \tag{3.3}$$

де $F_{k,i}^{(l)}$ і K(l-1) фільтри для заданого k.

Функція активації - це нелінійна функція, яка вносить нелінійність у мережі. Сигмоїдна функція була популярною функцією активації. З точки зору CNN, найбільш часто використовуваною функцією активації після операцій згортки є випрямлена лінійна одиниця (ReLU), яка визначається як

$$\mathrm{Re}\,LU(x) = \max(0, x) \tag{3.4}$$

Інша активаційна функція TLU, запропонована Ye та ін. [20], визначається як

$$TLU(x) = \begin{cases} -T, & x < -T \\ x, & -T \le x \le T \\ T, & x > T \end{cases} \tag{3.4}$$

де T - параметр. Крім того, абсолютна функція активації gросто визначається як

$$f(x) = |x| \tag{3.5}$$

Пакетна нормалізація - це тип шару, метою якого є адаптивна нормалізація даних. Він нормалізує розподіл кожної ознаки до нульового середнього значення та одиничної дисперсії, а потім масштабує та транслює розподіл під час навчання. Основна перевага використання пакетної нормалізації полягає в тому, що вона допомагає підтримувати поширення градієнта. Це може прискорити процес навчання за рахунок використання більшої швидкості навчання. Зауважте, що два параметри: параметр зсуву і параметр масштабу навчаються на основі навчальних даних.

Операція об'єднання поділяється на дві категорії: середнє об'єднання та максимальне об'єднання. Для розпізнавання зображень кращим є максимальне об'єднання, яке має локальну інваріантність при перекладі, коли ознаки перераховуються. Однак середнє об'єднання найчастіше використовується в стеганоаналізі, оскільки вважається, що інформація, яка вбудовується, ховається в шумі зображення. Використання максимального об'єднання може призвести до отримання корисної інформації, яка допоможе класифікувати звичайні зображення та стегозображення. Часто об'єднання використовується для зменшення розміру вихідних карт ознак шляхом вибору відповідного кроку. Наприклад, середній або максимальний крок об'єднання 2 × 2 на зображенні зменшить вибірку вдвічі.

3.2 Запропоновані нейронні мережі

На основі сучасних досліджень, згаданих вище, пропонується створити архітектуру CNN, яка використовує 30 фільтрів високих частот, що навчаються, у першому згортковому шарі на початку архітектури. Ідея цієї конструкції з високою фільтрацією полягає в тому, що хочемо розмістити самонавчальний згортковий шар з високою фільтрацією на початку. Термін "самонавчання" означає, що ядра цього шару ШНМ можуть навчатися, а саме, ваги ядер коригуються в процесі навчання. Як згадувалося раніше, Цянь та ін. [18] використовували один фіксований фільтр високих частот (4.1), який не піддається навчанню. Ye-Net [20] використовували шар ШНМ, що навчається, з ядрами 30 фільтрів високих частот як початковими ознаками. Але ваги ядер змінюються під час навчання, і малоймовірно, що вони продовжуватимуть виконувати фільтрацію високих частот.

Для того, щоб 30 фільтрів працювали як фільтри високих частот під час навчання, додамо деякі обмеження на ці фільтри. Зробимо їх симетричними і тримаємо суму ваг ядер (для кожного ядра) рівною нулю, що є властивістю фільтрів високих частот. Можна припустити, що ядра, які навчаються, зможуть самонавчатися на основі даних під час процесу навчання. Знову ж таки, суть полягає в тому, щоб накласти деякі обмеження на ядра, що навчаються. Yedroudj-Net зафіксував 30 фільтрів високих частот на початку для проведення фільтрації високих частот

Варто зазначити, що в Xu-Net, YeNet тощо існують успішні шари, такі як шар усічення, середній шар об'єднання. Ці шари, можуть бути корисними, тому і вводяться в запропоновані архітектури CNN. Крім того, залишкові зв'язки широко використовуються в глибоких нейронних мережах для навчання моделей. Вони також розглядаються в запропонованій архітектурі.

На рисунках 3.3 та 3.4 представлено запропоновану архітектуру CNN, яка складається з двох блоків, показаних на рисунку 3.2. Архітектурна схема Xu-Net представлена на рисунку 3.5, архітектурна схема Ye-Net - на рисунку

3.6, а архітектурна схема Yedroudj-Net - на рисунку 3.7.

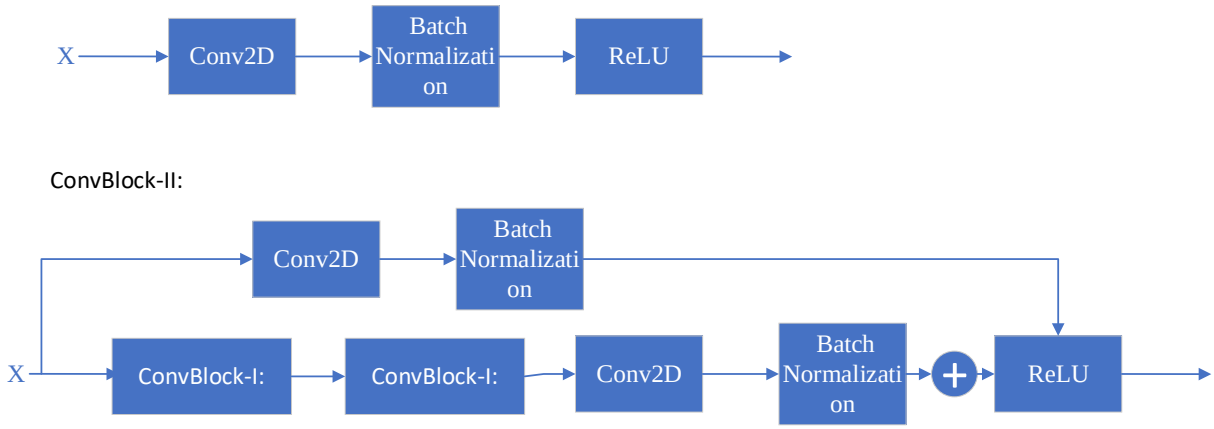

Рисунок 3.2 – Два типи блоків: Згортковий блок-I та згортковий блок-II

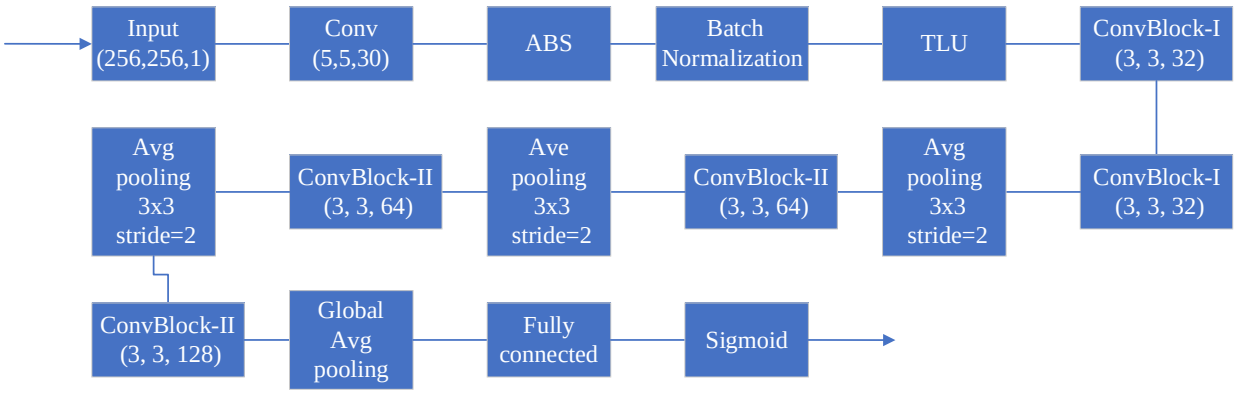

Рисунок 3.3 – Схема запропонованої архітектури CNN-I

Розглянемо три схеми адаптивного вбудовування контенту: WOW, MiPOD та SUIWARD. Порівнюємо продуктивність великої моделі з ансамблевим навчанням, Xu-Net, Ye-Net, Yedroudj-Net та запропонованої нами нейронної мережі на наборі зображень, що складається з 40 000 зображень розміром 256 × 256. Цей набір зображень [21] був створений Пібре та ін. на основі популярного набору даних BOSSbase. Набір даних BOSSbase містить 10 000 напівтонових зображень розміром 512×512. Кожне зображення в наборі даних BOSS розділене на чотири зображення розміром

256. Таким чином, вони отримали 40 000 зображень з набору даних BOSSbass і розмістили їх на своєму сайті.

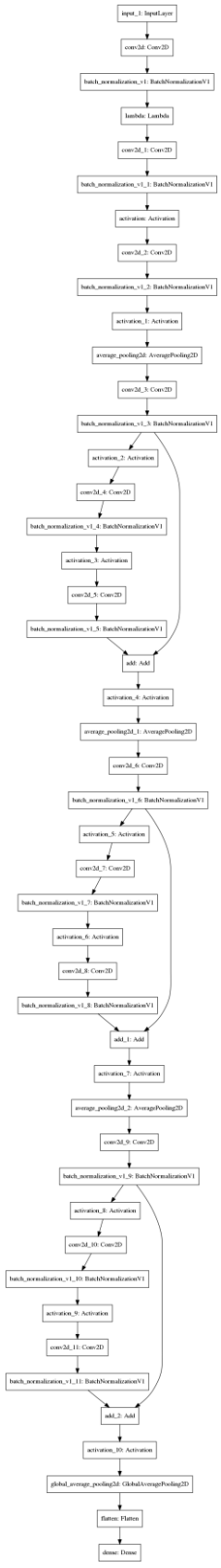

Рисунок 3.4 – Схема запропонованої архітектури CNN- II

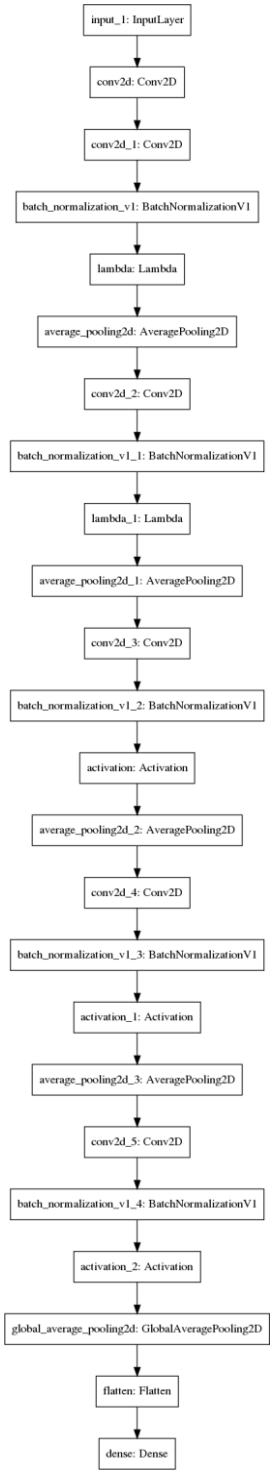

Рисунок 3.5 – Схема архітектури Xu-Net

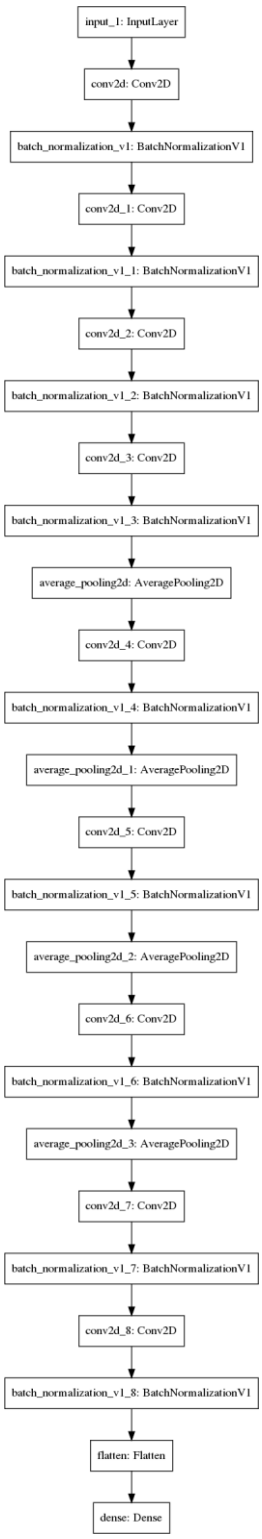

Рисунок 3.6 – Схема архітектури Ye-Net

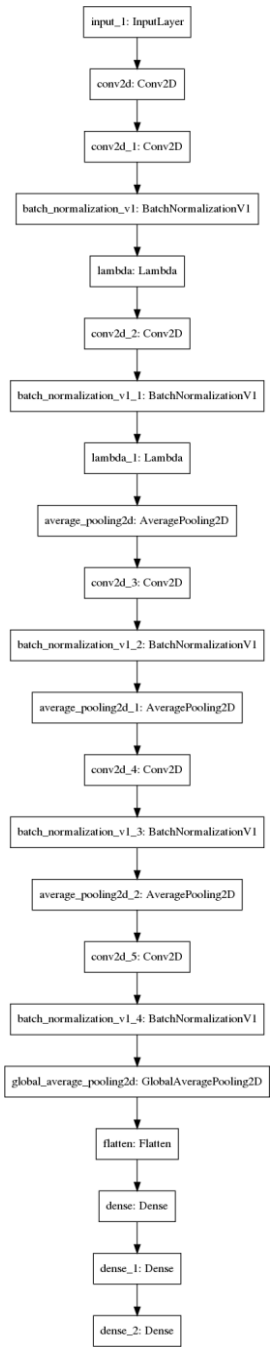

Рисунок 3.7 – Схема архітектури Yedroudj-Net

Для кожної схеми вбудовування було створено 40 000 стегозображень і, відповідно, маємо 40 000 пар зображення обкладинка/стегозображення (тобто, 40 000 оригінальних зображень + 40 000 стегозображень). Далі генеруємо навчальний, валідаційний та тестовий набори даних наступним чином. Випадковим чином вибираємо 4 000 пар обкладинка/стего для формування тестового набору даних і 8 000 пар обкладинка/стего для

формування валідаційного набору даних. Навчальний набір даних складається з решти 28 000 пар зображень обкладинка/стего. Всі архітектури CNN працюють на одному графічному процесорі з 24 ГБ пам'яті. Під час навчання CNN фіксуємо максимум 450 епох. Також було використано популярний API для глибокого навчання Keras з Tensorflow в якості бекенду для реалізації всіх трьох архітектур CNN, що використовуються в експерименті.

3.3 Результати експерименту

Далі покажемо кілька результатів навчання для Ye-Net, Yedroudj-Net та запропонованої нейронної мережі. На рисунку 3.8 показано точність навчання та валідації, а також похибки для Ye-Net. Схема вбудовування - WOW з корисним навантаженням 0.4. На рисунках 3.9 та 3.10 показано точність навчання та валідації, а також втрати для Yedroudj-Net. Схема вбудовування - S-UNIWARD з корисним навантаженням 0.4 та 0.2 відповідно. На рис. 3.11 показано точність навчання та валідації, а також втрати для запропонованої нейронної мережі. Схема вбудовування - WOW з корисним навантаженням 0,4.

Таблиця 3.1 показує частоту помилок виявлення для трьох різних схем вбудовування з використанням SRM («модель з великим обсягом вхідних даних з ансамблевим навчанням»), Xu-Net, Ye-Net, YedroudjNet та запропонованої архітектури CNN. Як видно, запропонована нейронна мережа перевершує інші для схем вбудовування WOW і S-UNIWARD з корисним навантаженням 0.2, 0.3 і 0.4, а також для схеми вбудовування MiPOD з корисним навантаженням 0.2. Yedroudj-Net показує найкращі результати для схеми впровадження MiPOD з корисним навантаженням 0.3 і 0.4.

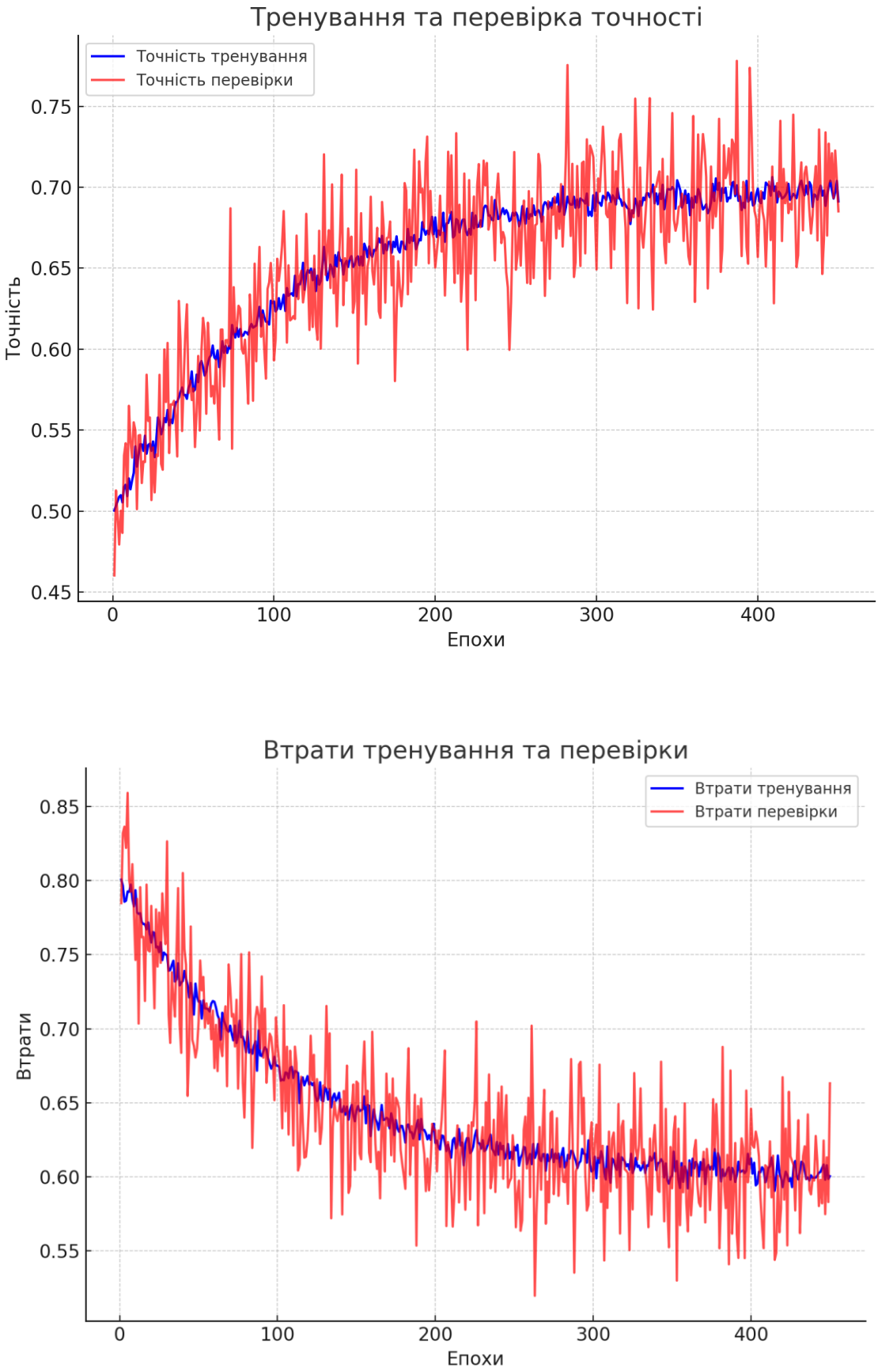

Рисунок 3.8 – Ye-Net для WOW з корисним навантаженням 0.4pp

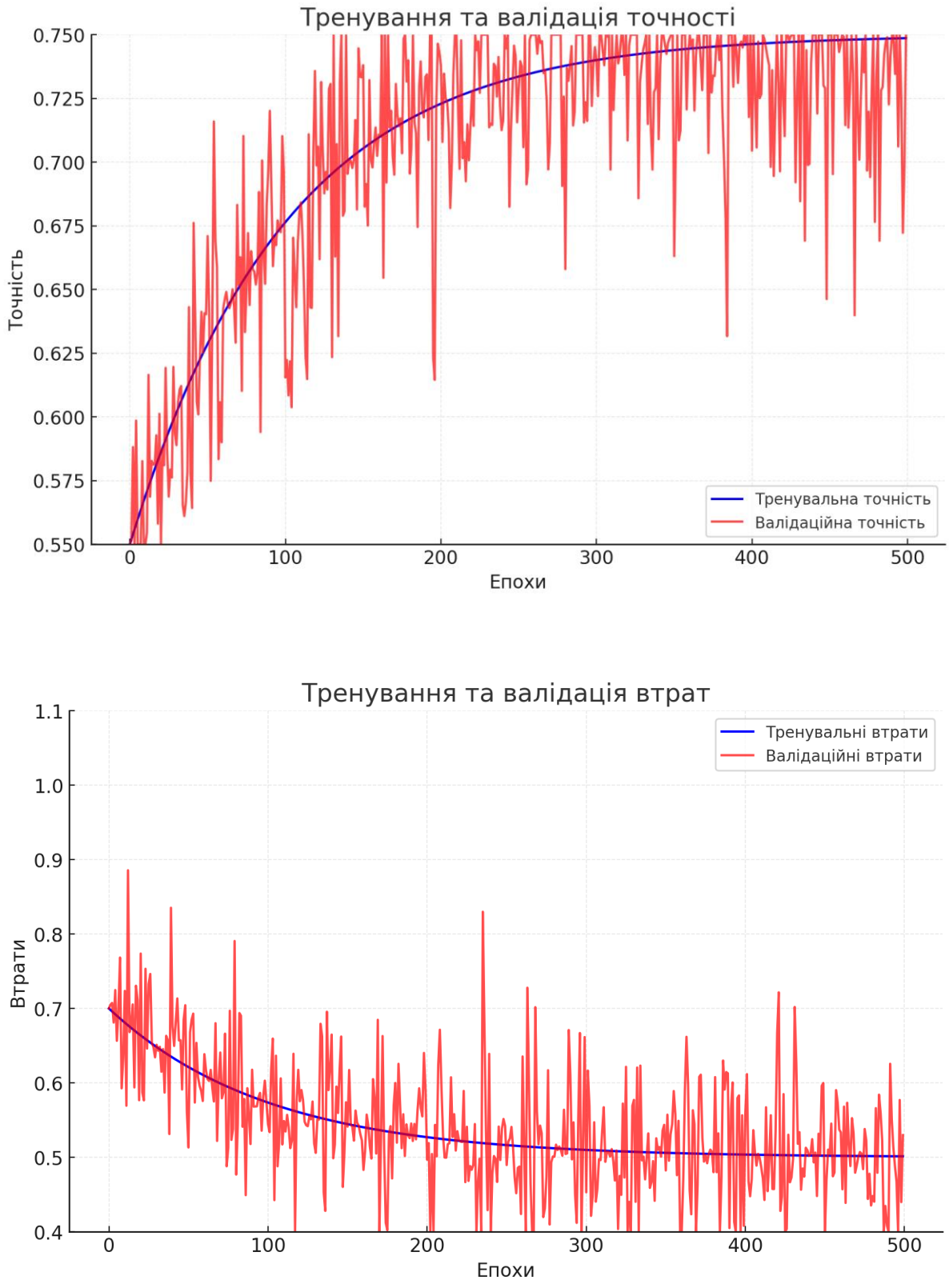

Рисунок 3.9 – Yedroudj-Net для S-UIWARD з корисним навантаженням 0.4pp

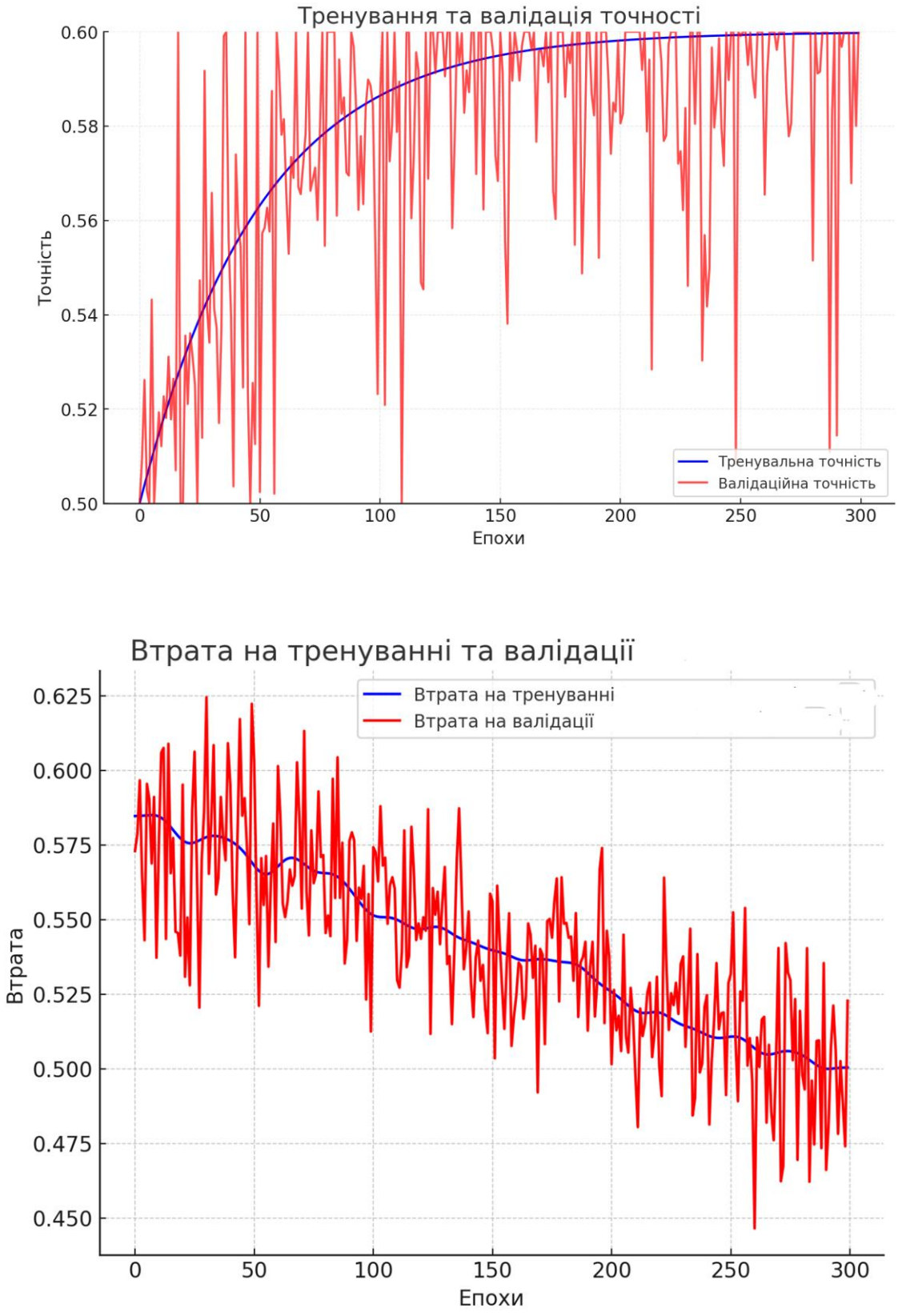

Рисунок 3.10 – Yedroudj-Net для S-UIWARD з корисним навантаженням 0.2pp

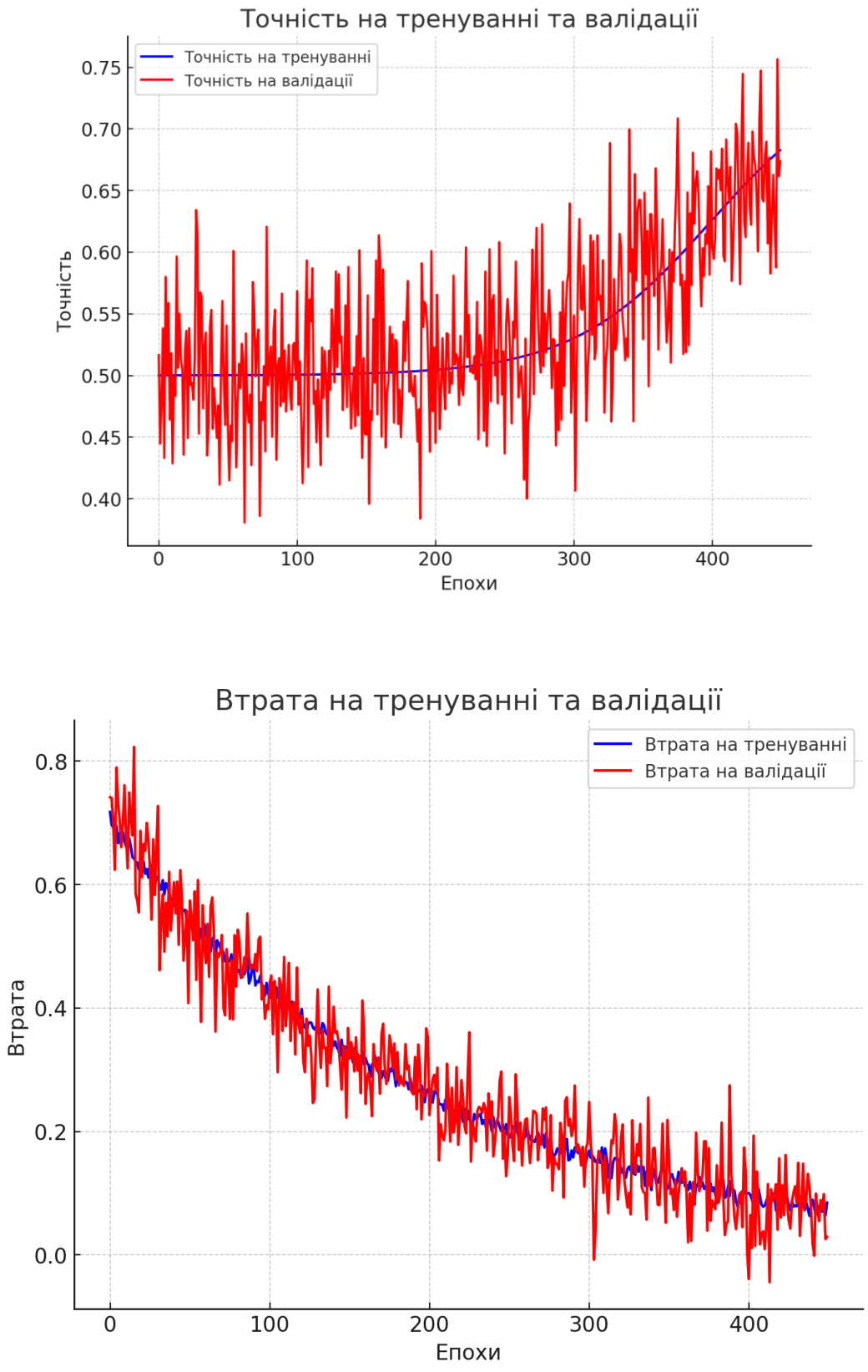

Рисунок 3.11–Запропонована мережа для WOW з корисним навантаженням 0.4pp

Таблиця 3.1 – Рівень помилок виявлення для WOW, MiPOD та S-UNIWARD.

Метод стеганографії	Корисне навантаження (bpp)	SRM	Xu-Net	Ye-Net	Yedroudj-Net	Пропонована мережа
WOW.	0.2	0.391	0.409	0.397	0.392	0.388
	0.3	0.345	0.348	0.343	0.340	0.337
	0.4	0.286	0.302	0.288	0.281	0.274
MiPOD	0.2	0.396	0.386	0.380	0.379	0.377
	0.3	0.322	0.318	0.317	0.306	0.311
	0.4	0.284	0.262	0.255	0.246	0.249
	0.2	0.398	0.391	0.387	0.385	0.379
S-UNIWARD	0.3	0.353	0.349	0.341	0.315	0.302
	0.4	0.299	0.287	0.276	0.271	0.267

ВИСНОВКИ

У цій кваліфікаційній роботі досліджено статистичний стеганоаналіз зображень. Спершу описано основи стеганографії та схеми вбудовування інформації, зокрема метод із використанням Судоку. Запропоновано схему, яка покращує якість стего-зображень за допомогою еталонної матриці.

Далі розглянуто стегоаналіз із застосуванням фреймворку перевірки гіпотез, що дозволяє оцінювати ефективність аналізу. Показано переваги запропонованого тесту над LR-тестом, особливо за низьких швидкостей вбудовування. Водночас, виклики, пов'язані з адаптивними схемами вбудовування, підштовхнули до використання підходів глибокого навчання.

Розроблено архітектуру глибокої нейронної мережі (ГНМ) із навчальним шаром високочастотної фільтрації, що покращує результати стеганоаналізу. Ця архітектура перевершує традиційні детектори (Xu-Net, Ye-Net, Yedroudj-Net) на великому наборі даних. У перспективі планується застосування ГНМ для аналізу в домені JPEG із можливістю використання трансферного навчання та оптимізації обчислювальних витрат.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Fridrich, Jessica. *Steganography in digital media: principles, algorithms, and applications*. No. 190288. Cambridge University, 2009.

2. Simmons, G. J. (1984, August). The prisoners' problem and the subliminal channel. In *Advances in Cryptology: Proceedings of Crypto 83* (pp. 51-67). Boston, MA: Springer US.

3. J. Mielikainen, "LSB matching revisited," in *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 285-287, May 2006, doi: 10.1109/LSP.2006.870357.

4. Bas, P., Filler, T., Pevný, T. (2011). "Break Our Steganographic System": The Ins and Outs of Organizing BOSS. In: Filler, T., Pevný, T., Craver, S., Ker, A. (eds) Information Hiding. IH 2011. Lecture Notes in Computer Science, vol 6958. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-24178-9_5

5. Прогонов, Д. О. "Теоретико-інформаційні оцінки спотворень контейнерів при формуванні стеганограм." (2018).

6. Прогонов, Дмитро Олександрович. "Структурний синтез і параметрична оптимізація методів побудови стегодетекторів для цифрових зображень." (2024).

7. Прогонов, Д. О. "Ефективність стегоаналізу цифрових зображень у випадку попередньої фільтрації стеганограм, сформованих згідно адаптивних методів MG та MIPOD." (2020).

8. Siwei Lyu and Hany Farid. Steganalysis using color wavelet statistics and one-class support vector machines. In Security, steganography, and watermarking of multimedia contents VI, volume 5306, pages 35– 45. International Society for Optics and Photonics, 2004.

9. J. Kodovsky, J. Fridrich and V. Holub, "Ensemble Classifiers for Steganalysis of Digital Media," in *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 432-444, April 2012, doi:

10.1109/TIFS.2011.2175919

10. Tan, Tieniu, et al. "Image Steganalysis Based on Deep Learning." U.S. Patent Application No. 15/557,080.

11. Li, Xuejing, Yonglong Luo, and Weixin Bian. "Retracing extended sudoku matrix for high-capacity image steganography." *Multimedia Tools and Applications* 80.12 (2021): 18627-18651.

12. Allan Weber Technical questions about the database images, the image format, or problems obtaining the images should be directed to the database editor [Електронний ресурс] : The USC-SIPI Image Database Режим доступу: https://sipi.usc.edu/database/.

13. Chin-Chen Chang, Yung-Chen Chou, and The Duc Kieu. An information hiding scheme using sudoku. In 2008 3rd international conference on innovative computing information and control, pages 17–17. IEEE, 2008.

14. Wien Hong, Tung-Shou Chen, and Chih-Wei Shiu. Steganography using sudoku revisited. In 2008 Second International Symposium on Intelligent Information Technology Application, volume 2, pages 935– 939. IEEE, 2008.

15. Laishram, Debina, and Themrichon Tuithung. "A secure adaptive Hidden Markov Model-based JPEG steganography method: (J-HMMSteg)." *Multimedia Tools and Applications* 83.13 (2024): 38883-38908.

16. Jia, Ju, et al. "Transferable heterogeneous feature subspace learning for JPEG mismatched steganalysis." Pattern Recognition 100 (2020): 107105.

17. Jessica Fridrich and Jan Kodovsky. Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 7(3):868–882, 2012.

18. S. Tan and B. Li, "Stacked convolutional auto-encoders for steganalysis of digital images," Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific, Siem Reap, Cambodia, 2014, pp. 1-4, doi: 10.1109/APSIPA.2014.7041565

19. Yinlong Qian, Jing Dong, Wei Wang, and Tieniu Tan. Deep learning for steganalysis via convolutional neural networks. In Media Watermarking, Security,

and Forensics 2015, volume 9409, page 94090J. International Society for Optics and Photonics, 2015.

20. Jian Ye, Jiangqun Ni, and Yang Yi. Deep learning hierarchical representations for image steganalysis. IEEE Transactions on Information Forensics and Security, 12(11):2545–2557, 2017.

21. The cropped BOSSBase stego images with S-UNIWARD at 0.4 bpp [Електронний ресурс] CroppedBossBase-1.0-256x256_stego_SUniward0.4bpp.rar. Режим доступу: https://www.lirmm.fr/~chaumont/SteganalysisWithDeepLearning/CroppedBossBase-1.0-256x256 _stego _SUniward0.4bpp.rar