

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Факультет прикладної математики

Кафедра прикладної математики

«На правах рукопису»

УДК 519.688:004.855.5

«До захисту допущено»

Завідувач кафедри

_____ Олег ЧЕРТОВ

« ____ » _____ 2023 р.

Магістерська дисертація
на здобуття ступеня магістра
за освітньо-науковою програмою «Наука про дані та математичне
моделювання»
зі спеціальності 113 «Прикладна математика»
на тему: ««Математичне та програмне забезпечення системи
розпізнавання військової техніки за допомогою супутникових зображень»»

Виконав: студент 2 курсу, групи КМ-11мн

Давиденко Микола Андрійович

Керівник:

Доцент, канд. фіз.-мат. наук, доцент

Третиник Віолета Вікентіївна

Консультант з нормоконтролю:

старший викладач,

Мальчиков Володимир Вікторович

Рецензент:

професор кафедри математичних методів

системного аналізу ІІСА, д-р

техн. наук, доцент

О.Ю.Зайченко

Засвідчую, що в цій магістерській
дисертації немає запозичень із праць інших
авторів без відповідних посилань.

Студент _____

Київ — 2023

Національний технічний університет України

«Київський політехнічний інститут

імені Ігоря Сікорського»

Факультет прикладної математики

Кафедра прикладної математики

Рівень вищої освіти — другий (магістерський)

Спеціальність – 113 «Прикладна математика»

Освітньо-наукова програма «Наука про дані та математичне моделювання»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Олег ЧЕРТОВ

«__» _____ 2023 р.

ЗАВДАННЯ

на магістерську дисертацію студенту

Давиденко Миколі Андрійовичу

1. Тема дисертації: «Математичне та програмне забезпечення системи розпізнавання військової техніки за допомогою супутникових зображень», науковий керівник дисертації Третиник Віолета Вікентіївна, кандидат фіз.-мат. наук, доцент, затверджені наказом по університету від «30» березня 2023 р. № 1359-С.
2. Термін подання студентом дисертації: «15» травня 2023 р.
3. Об'єкт дослідження: Об'єктом дослідження методи та моделі розпізнавання об'єктів за супутниковими знімками на основі методів та моделей нейронних мереж. Моделі підвищення якості зображення, апскейлінг. Різні архітектури нейронних мереж: згорткові нейронні мережі, глибока згорткова нейронна мережа, U-Net, Залишкова нейронна мережа. Методи оптимізації та гіперпараметри нейронної мережі. Предмет дослідження: Математичне та програмне забезпечення системи розпізнавання морських об'єктів за супутниковими знімками на основі нейромережевої моделі. Застосування нейронних мереж на етапі попереднього тренування для розпізнавання військової техніки за допомогою супутникових

зображень, порівняльний аналіз методів оптимізації, що використовуються при навчанні нейронних мереж.

5. Перелік завдань, які потрібно розробити: знайти великий та змістовний датасет, проаналізувати на змістовність, виконати аналіз моделей, котрі використовуються для розпізнавання і вибрати одну. Застосувати алгоритми машинного навчання та нейроні мережі для знаходження закономірності, розробити інтерфейс.

6. Орієнтовний перелік ілюстративного матеріалу: приклади супутникових зображень, архітектура моделей нейронних мереж, архітектура розробленого програмного забезпечення, зображення з прикладами військової техніки.

7. Орієнтовний перелік публікацій: тези «Попереднє покращення якості зображення для підвищення точності розпізнавання об'єктів».

8. Дата видачі завдання: «1» лютого 2023 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Ґрунтовне ознайомлення з предметною областю	15.09.2022	
2	Визначення структури магістерської дисертації; вивчення літератури, пошук додаткової літератури	01.10.2022	
3	Робота над першим розділом магістерської дисертації	20.10.2022	
4	Проведення наукового дослідження; робота над другим розділом магістерської дисертації	15.01.2023	
5	Проведення наукового дослідження; робота над статтею за результатами наукового дослідження	15.02.2023	
6	Робота над третім розділом магістерської дисертації; підготовка статті за результатами наукового дослідження; розроблення програмного забезпечення	01.03.2023	
7	Оформлення текстової та графічної частин магістерської дисертації	25.05.2023	

Студент

Микола

ДАВИДЕНКО

Науковий керівник дисертації

Віолета ТРЕТИНИК

Реферат

Дисертацію виконано на 92 аркушах, вона містить 2 додатки та перелік посилань на використані джерела з 23 найменувань. У роботі наведено 30 рисунки та 3 таблиці.

Актуальність теми. Після 24 лютого 2022 року життя кожного українця назавжди змінилося від російського вторгнення та повномасштабної війни. Тож у сучасних життєвих реаліях війна зайняла окреме місце.

Щоб підвищити оборонну спроможність важливо мати танки, артилерію та літаки, але щоб підвищити якість всього потрібна допомога математики та машинного навчання. На заході вже давно розробляються автоматичні системи прийняття рішень на основі штучного інтелекту, а дрони навчають автоматично знаходити цілі, і це далеко не повний список. Тож мета цієї роботи розробити інструмент для підвищення оборонної спроможності нашої країни.

Щоб стежити та приймати відповідні рішення можна використовувати багато людських та фінансових ресурсів, що не є оптимально. Але є другий варіант, завдяки прогресу у супутниковій фотозйомці та нейронних мережах, це можна зробити повністю автоматизовано.

В наш час стало зрозуміло що незважаючи на 21 століття військові конфлікти нікуди не зникли, тому є сенс створювати оборонний потенціал. Важливу роль у сучасній війні є розвідка, бо якщо є інформація про розташування ворога, то можна зробити контрдії які можуть кардинально змінити хід війни.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконувалась згідно з планом науково-дослідних робіт кафедри прикладної математики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Мета і задачі дослідження. Розробка математичного та програмного забезпечення системи розпізнавання військової техніки за допомогою супутникових зображень на основі нейромережевої моделі для інтеграції з сайтами. Для розширення функціональних можливостей і адаптації системи до використання людьми без

профільної освіти. Програма має бути інтегрована з сайтом, додатком або впроваджена в робочий процес.

Об'єктом дослідження є методи та моделі розпізнавання об'єктів за супутниковими знімками на основі методів та моделей нейронних мереж. Моделі підвищення якості зображення, апскейлінг. Різні архітектури нейронних мереж: згорткові нейронні мережі, глибока згорткова нейронна мережа, U-Net, Залишкова нейронна мережа. Методи оптимізації та гіперпараметри нейронної мережі.

Предметом дослідження є математичне та програмне забезпечення системи розпізнавання військових об'єктів за супутниковими знімками на основі нейромережевої моделі. Застосування нейронних мереж на етапі попереднього тренування для розпізнавання морських об'єктів за супутниковими знімками, порівняльний аналіз методів оптимізації, що використовуються при навчанні нейронних мереж.

Методи дослідження. Для розв'язання поставленої задачі використовувались такі методи: методи машинного навчання (для розробки моделі нейронної мережі); методи оптимізації (для пошуку найкращого налаштування системи); методи обробки даних (для попередньої підготовки вхідних даних); методи теорії алгоритмів та програмування (для програмної реалізації розроблених алгоритмів).

Наукова новизна одержаних результатів включає в себе такі пункти:

- Комбінація використання навчених нейронних мереж з відкритим кодом для підвищення якості (апскейлінг) супутникових зображень, перед застосуванням основної згорткової нейронної мережі для розпізнавання об'єктів на фото.

- Запропоновано використання згорткових нейронних мереж для розпізнавання військової техніки в комбінації сучасних комерційних супутників з синтезованою апертурою.

Практичне значення одержаних результатів. Полягає в тому, що розроблена система для розпізнавання військової техніки за супутниковими знімками при подальшій модифікації може використовуватися військовими Збройними силами України.

Апробація результатів дисертації. Основні положення й результати роботи представлено на конференції ПМК 2022 (Прикладна математика та комп'ютинг).

Публікації. Результати дисертації викладено в одній науковій праці:

- Третиник В. В., Давиденко М.А. (2022) Попереднє покращення якості зображення для підвищення точності розпізнавання об'єктів. ПМК 2022. С. 169-173;

Ключові слова: супутникова зйомка , методи підвищення роздільної здатності, згорткова нейронна мережа, класифікація зображень, навчання нейронних мереж, розпізнавання об'єктів, комп'ютерне бачення, оптимізація параметрів.

ABSTRACT

The dissertation is made on 92 sheets, it contains 3 appendices and a list of references to the sources used from 23 titles. The paper contains 30 figures and 3 tables.

Relevance of the topic. After February 24, 2022, the life of every Ukrainian has changed forever from the Russian invasion and full-scale war. Therefore, in modern life realities, the war took a separate place.

It is important to have tanks, artillery and airplanes to increase the defensive capability, but to improve the quality of everything you need the help of mathematics and machine learning. In the West, automatic systems for bringing solutions based on artificial intelligence have long been developed, and drones are trained to automatically find targets, and this is far from a complete list. Therefore, the purpose of this work is to develop a tool to increase the defense capability of our country.

Many human and financial resources can be used to monitor and make appropriate decisions, which is not optimal. But there is a second option, thanks to advances in satellite photography and neural networks, that can be done in a fully automated way.

In our time, it became clear that despite the 21st century, military conflicts have not disappeared anywhere, so it makes sense to create a defensive potential. An important role in modern warfare is intelligence, because if there is information about the location of the enemy, then you can make counter-actions that can radically change the course of the war.

Connection of work with scientific programs, plans, topics. The dissertation work was carried out in accordance with the plan of research work of the Department of Applied Mathematics of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute."

Purpose and objectives of the study. Development of a mathematical and software system for recognizing military equipment using satellite images based on a neural network model for integration with sites. To expand functionality and adapt the system to the use of people without specialized education. The application must be integrated with the site, application, or implementation into the workflow.

The object of research is methods and models of object recognition by satellite snapshots based on methods and models of neural networks. Models for improving the quality of zoom image, upscaling. Different architectures of neural networks: convolutional neural networks, deep convolutional neural network, U-Net, Residual neural network. Optimization methods and hyperparameters of the neural network.

Subject of research: Mathematical and software systems for recognizing military objects from satellite images based on a neural network model. Application of neural networks at the stage of preliminary training for recognition of marine objects from satellite images, comparative analysis of optimization methods used in training neural networks.

Methods of research. To solve the problem, the following methods were used: machine learning methods (for developing a neural network model); Optimization methods (to find the best system configuration) data processing methods (for preliminary preparation of input data); methods of algorithm theory and programming (for software implementation of the developed algorithms).

Scientific contribution consists of the following::

- A combination of using trained open-source neural networks to improve the quality (upscaling) of satellite images, before using the main convolutional neural network to recognize objects in the photo.
- The use of convolutional neural networks for the recognition of military equipment in a combination of modern commercial satellites with a synthesized aperture is proposed.

The practical significance of the results obtained. It consists in the fact that the developed system for recognizing military equipment from satellite images during further modification can be used by the military of the Armed Forces of Ukraine.

Approbation of the results of the dissertation. The main provisions and results of the work are presented at the conference PMK 2022 (Applied Mathematics and Computing).

Publications. The results of the dissertation are presented in one scientific work:

- Tretinik V.V., Davydenko M.A. (2022) Preliminary improvement of image quality to improve the accuracy of object recognition. PMK 2022. PP. 169-173;

Key words: satellite shooting, resolution enhancement methods, convolutional neural network, image classification, neural network training, object recognition, computer vision, parameter optimization.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	14
ВСТУП.....	15
1 ОГЛЯД ІСНУЮЧИХ РІШЕНЬ	16
1.1 Бізнес цілі	16
1.2 Методи фіксації та види військової техніки.....	17
1.2.1 ENVI	20
1.2.2 Google Earth Pro	20
1.2.3 TerraSAR-X	22
1.2.4 WorldView-3.....	24
1.3 Огляд існуючих методів	25
1.3.1 Метод шаблонів.....	26
1.3.2 Метод каскаду Гаара.....	28
1.3.3 Метод розпізнавання за зовнішніми ознаками	29
1.3.4 Метод кластеризації.....	31
1.3.5 Метод глибокого навчання	33
1.4 Аналіз існуючих рішень	35
1.4.1 Google Lens	35
1.4.2 Amazon Rekognition	37
1.5.2 IBM Watson Visual Recognition.....	37
1.6 Висновки до розділу	39
2 ВІДОМОСТІ ТА РІЗНОВИД ВІЙСЬКОВОЇ ТЕХНІКИ	40
2.1 2С1 Гвоздика.....	41
2.2 BRDM-2	43
2.3 BTR-60	45
2.4 Т-62	46
2.5 Т-72	48
2.6 ZSU-23-4 Shilka.....	49

2.7 Zil-131	51
2.8 Макети	52
2.9 Висновки до розділу	53
3 ОГЛЯД ТА ОБҐРУНТУВАННЯ МАТЕМАТИЧНИХ МЕТОДІВ	55
3.1 CNN.....	55
3.2 Персептрон.....	58
3.3 Метрики.....	59
3.3.1 Categorical Accuracy	59
3.3.2 Weighted Average F2 Score	60
3.3.3 Matthews Correlation Coefficient.....	61
3.4 Підбір функції оптимізації	61
3.4.1 Adam	62
3.4.2 SGD.....	65
3.4.3 RMSprop	65
3.4.4 Adagrad	66
3.5 Upscaling.....	67
3.5.1 Бікубічна інтерполяція	68
3.5.2 Білінійна інтерполяція	69
3.5.3 EDSR.....	70
3.5.4 ESPCN	71
3.5.5 Вибір оптимального методу Upscaling	73
3.6 Висновки до розділу	76
4 КОМПОНЕНТНА МОДЕЛЬ СИСТЕМИ	79
4.1 Складові системи.....	79
4.2 Опис функціональних компонентів	80
5 ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ	81
5.1. Архітектура розроблених програмних засобів	81
5.2 Навчання моделі	81
5.3 Результат роботи	83

5.4 Висновки до розділу	87
ВИСНОВКИ.....	89
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	90
Додаток А ЛІСТИНГ ПРОГРАМИ.....	93
Додаток Б ІЛЮСТРАТИВНИЙ МАТЕРІАЛ.....	97

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ML – machine learning;

БД – база даних;

ЗНМ – згорткова нейронна мережа;

НМ – нейронна мережа;

ОС – операційна система;

СУБД – система управління базами даних;

ПЗ – програмне забезпечення;

ШІ – штучний інтелект.

ВСТУП

Після 24 лютого 2022 року життя кожного українця назавжди змінилося від російського вторгнення та повномасштабної війни. Тож у сучасних життєвих реаліях війна зайняла окреме місце.

Щоб підвищити оборонну спроможність важливо мати танки, артилерію та літаки, але щоб підвищити якість всього потрібна допомога математики та машинного навчання. На заході вже давно розробляються автоматичні системи прийняття рішень на основі штучного інтелекту, а дрони навчають автоматично знаходити цілі, і це далеко не повний список. Тож мета цієї роботи розробити інструмент для підвищення оборонної спроможності нашої країни.

Щоб стежити та приймати відповідні рішення можна використовувати багато людських та фінансових ресурсів, що не є оптимально. Але є другий варіант, завдяки прогресу у супутниковій фотозйомці та нейронних мережах, це можна зробити повністю автоматизовано.

За останні 10 років було проведено багато роботи для автоматичного вилучення об'єктів із супутникових зображень зі значними результатами, але без ефективного оперативного ефекту. Але тепер сучасні нейронні мережі та швидкі процесори можуть підвищити точність і швидкість для автоматичного виявлення військової техніки.

1 ОГЛЯД ІСНУЮЧИХ РІШЕНЬ

1.1 Бізнес цілі

Війни були та залишаються дуже дорогими та кровопролитними протягом усієї історії людства. Є багато прикладів коли після невдалої війни держави не витримували післявоєнну кризу, а саме великий дефіцит грошей та велику втрату населення, і зазнавали краху, дуже гарним прикладом цього є Перша світова війна. Тому у 21 столітті найвпливовіші держави передивившись цю концепцію зрозуміли, що життя кожного солдата коштує дуже дорого, тому краще заплатити значну суму на екіпіровку та додаткове оснащення, щоб підвищити шанси зберегти життя людині ніж дати йому дешеву екіпіровку, яка йому ніяк не допоможе та приведе до 0 результату.

Зрозумівши це, з'явився пріоритет на фінансування безліч технологій, які напряду або дотичною можуть врятувати життя солдатів та людей. До них відноситься як звичайні бронешилети, так і системи раннього виявлення ракетних обстрілів.

Тобто чим технологічна армія, тим більш результативно вона буде працювати.

Звичайно технології це не 100 відсотків перемоги, але без них у сучасній війні не здобудеш перемогу. Яскравим прикладом сучасної армії є Ізраїль, невелика країна, але яка має великий оборонний комплекс з найрізноманітнішими технологіями. Саме ця країна показала, що її армія є найякісніша.

Тож повертаючись до теми дисертації, можна сказати що більшість військових технологій, великих чи малих за вартістю, стоїть за тим, щоб зберегти життя кожному солдату. Тоді можна чітко зрозуміти цілі які поставлені у даній дисертаційній роботі: по перше, зробити систему для розпізнавання ворожої військової техніки за супутниковими фотографіями, підвищуючи якість нашої армії, та, по-друге тим самим врятувавши життя звичайним солдатам, які не мають ціни взагалі.

1.2 Методи фіксації та види військової техніки

Ще з початку минулого століття основним методом розвідки були групи піхоти розвідників які своїми очима фіксували ворожі сили і потім поверталися з добутою інформацією та доповідали капітанам. Це було дуже повільно та з великим ризиком для самих розвідників. Тому не дивно що їх хотіли замінити на щось ефективніше. Це стало абсолютно новою промисловістю – аеророзвідкою.

Спочатку це були дирижаблі та аеростати, але вони були дуже повільними та помітними, тому вони надовго не затримались. Саме в літаках військові побачили великий потенціал для розвідки. Перші військові літаки були зроблені без будь-якого озброєння, але вони були двомісні та другий пілот брав з собою камеру для повітряної зйомки. Дана технологія відкрила величезний потенціал для військових, тому що вони могли наперед знати де сили ворога більші, а де їх дуже мало і саме туди спрямовувати удар.

З часом літаки та обладнання для повітряної розвідки ставали все краще і краще, тим самим дозволяючи робити якісніші знімки та залишатися непоміченим. Вершиною інженерної думки став американський літак розвідник U-2, його особливістю була дуже висока висота польоту - майже 22 000 метрів. На такій висоті він був невразливий для більшості ППО і мав літати та збирати дані в будь-який час. Згодом все ж він став вразливим до ППО, але все одно такий літак був дуже корисним у часи холодної війни та зіграв важливу роль.

Навпаки, від аеророзвідки фото зроблені з супутника були дуже поганої якості та це було притаманно аж до 90-х років минулого століття. Головною перевагою супутників є те, що вони знаходяться на орбіті і їх неможливо простежити. З розвитком технічної складової камери отримали здібність фотографувати з більш-менш нормальної якості, що дозволило майже своєчасно отримувати дані про будь-яку частину землі. У 2023 році якість фотографій з супутників вражає, можливо

отримати фото навіть однієї людини в високій роздільній якості, що робить супутники ідеальним інструментом для розвідки.

Головною проблемою для звичайних фотокамер є природні явища, такі як хмари та дощ. Такі явища можуть поставити під сумнів всю суть космічних супутників з оптичними камерами. Але на більшості військових та в меншості на комерційних встановлюють радар із синтезованою апертурою.

Радар із синтезованою діафрагмою - це бічна когерентна радіолокаційна система бічного огляду, як правило встановлюють на літаки чи супутники, в якій рух платформи використовується для електронного моделювання надзвичайно великої діафрагми. Використовуючи такий радіолокатор, під час дистанційного зондування отримують радіолокаційні зображення високої роздільної здатності. Суть методу синтезу діафрагми полягає в тому, що радіолокаційні дані, отримані в кожному окремому періоді зонда, залишаються спільно з даними про розташування платформи, що відповідають цим періодам, а усі ці дані зберігаються протягом певного інтервалу часу[1]. Відстань, на яку буде рухатися платформа протягом цього інтервалу, визначає розмір синтезованої діафрагми. В кінці інтервалу отримані дані обробляються спільно, як ніби вони були отримані за допомогою антенного ґриля, елементи, які відповідають положенням платформи з радіолокатором протягом кожного періоду зондування. При обробці враховуються зміни частоти Доплера, які будуть різними для різних взаємних положень радарів та цілей у кожному періоді звучання. Таким чином, досягається значно більша роздільна здатність радіолокаційного зображення, ніж у радіолокації з тією ж антеною при використанні реальної діафрагми.

Принцип функціонування радіолокатора із синтезованою діафрагмою схожий на принцип функціонування радіолокатора з поетапною антенною решіткою. Різниця полягає в тому, що замість індивідуальних антенних елементів поетапної решітки в цьому випадку використовується лише одна антена, але з поєднанням інформації з її виходу в різні моменти часу і через рух платформи, в різних точках простору. У процесорі радіолокатора з синтезованою діафрагмою всі ехо-сигнали (у вигляді

амплітуд і фаз), прийняті протягом часу T -інтервалу, що відповідає зміщенню платформи з точки A до точки D завдяки цьому, можна сформувати сигнал, який може бути сформований, за допомогою антени з довжиною діафрагми $v \cdot t$, де v - швидкість платформи. Таким чином, синтезується діафрагма значної довжини, що не існує в реальності. Звідси назва цього типу радіолокатора. Зі збільшенням часу T синтезована діафрагма також збільшується в довжині, і, отже, досягається колись висока кутова (поперечна) роздільна здатність. Як тільки ціль (наприклад, корабель) потрапляє в промінь радіолокатора, прийняті відбиті сигнали для кожного імпульсу зонда починають записувати. Платформа з радіолокатором продовжує рухатися, а ехо-сигнали записуються, поки ціль знаходиться в проміні антени. Точка, коли платформа знаходиться, в якій ціль через деякий час залишає поле зору радіолокатора, визначає довжину синтезованої (модельованої) діафрагми антени. Розширюється синтезована ширина променями в поєднанні зі збільшенням часу, щоб знайти ціль у промені зі збільшенням діапазону треків, врівноважує один одного, щоб роздільна здатність залишалася постійною по всій смугі. Потенційна роздільна здатність у тангенціальному напрямку не залежить від діапазону до цілі та довжини хвилі та визначається лише половиною значень розкриття реальної антени[1]. На рисунку 1.1 показано схематичне зображення роботи технології SAR.

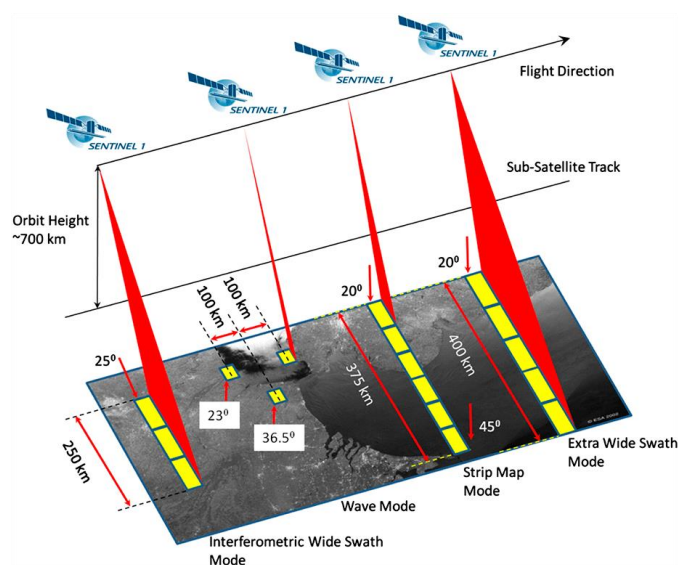


Рисунок 1.1 - Приклад роботи SAR [2]

1.2.1 ENVI

ENVI (Environment for Visualizing Images) - це програмне забезпечення для обробки супутникових фотографій та іншої геопросторової інформації, що розроблене компанією Harris Geospatial[3].

ENVI дозволяє користувачам виконувати наступні операції з супутниковими зображеннями:

- обрізання та ресемплінг зображень;
- підготовка даних для класифікації та векторизації;
- класифікація зображень на основі піксельних значень та векторних даних ;
- створення та редагування векторних шарів;
- розрахунок NDVI (Normalized Difference Vegetation Index) та інших геоіндексів;
- аналіз геоінформаційних даних, включаючи статистику та зв'язок між різними шарами даних;

ENVI також має інтерфейс з іншими програмами геоінформатики та статистичними пакетами, що дозволяє інтегрувати різноманітні геопросторові дані та забезпечити аналітику на їх основі.

ENVI може бути використане в таких галузях як екологія, агрономія, геологія, міське планування та інші, де геопросторова інформація є важливим елементом дослідження.

1.2.2 Google Earth Pro

Google Earth Pro - це програмне забезпечення для обробки супутникових фотографій, яке надає користувачам можливість перегляду, аналізу та маніпулювання

географічними даними на глобальному рівні. Програма має багато корисних функцій, що дозволяють виконувати широкий спектр завдань, пов'язаних з обробкою супутникових знімків[4]. На рисунку 1.2 показано приклад роботи Google Earth Pro.

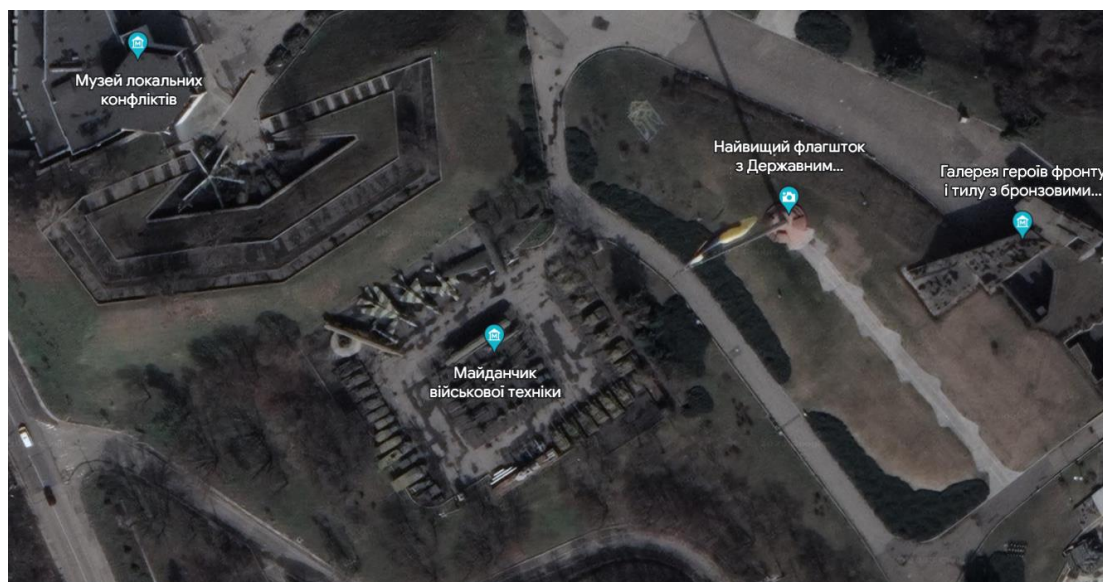


Рисунок 1.2 – Приклад роботи Google Earth Pro

Основні функції Google Earth Pro:

- Перегляд супутникових знімків: програма має доступ до найбільш точних супутникових знімків з усього світу, що дозволяє користувачам переглядати будь-яку точку планети з різних ракурсів;
- 3D-моделювання: програма дозволяє користувачам створювати 3D-моделі будь-яких географічних об'єктів, використовуючи знімки та векторні дані;
- Мітки та шари: Google Earth Pro дозволяє користувачам додавати мітки та шари до супутникових знімків, що дозволяє організовувати та категоризувати географічні дані ;
- Аналіз даних: програма дозволяє користувачам аналізувати різні географічні дані, такі як висота, нахил території, розташування об'єктів тощо;
- Розміщення на картах: Google Earth Pro дозволяє користувачам розміщувати на картах свої об'єкти та маршрути, що дозволяє ділитися даними з іншими користувачами;

- Інтеграція з іншими програмами: програма підтримує інтеграцію з іншими програмами, такими як AutoCAD, що дозволяє обмінюватися географічними даними між різними програмами;
- Перегляд зображень Землі в режимі реального часу, зміну масштабу та поворот зображення;
- Можливість вимірювати відстані, площі та об'єми;
- Можливість перегляду зображень, знятих з супутника, включаючи зображення з різних часових періодів;

З недоліків можна додати що поки що ця система не має радара під синтетичну апаратуру, під котру буде розроблятися система, але незважаючи на це її можна використовувати в тандемі з майбутньою моделлю для досягнення кращих результатів.

1.2.3 TerraSAR-X

TerraSAR-X - це німецький космічний радар, який був запущений на орбіту у червні 2007 року. Це супутник зі штучним діелектричним відмінностям, що дозволяє йому досліджувати Землю, використовуючи мікрохвильові радарні сигнали. TerraSAR-X був розроблений з метою отримання високоточних даних для застосування в таких галузях, як картографія, геологія, метеорологія, а також військові дослідження[5]. На рисунку 1.3 показано приклад роботи TerraSAR-X.

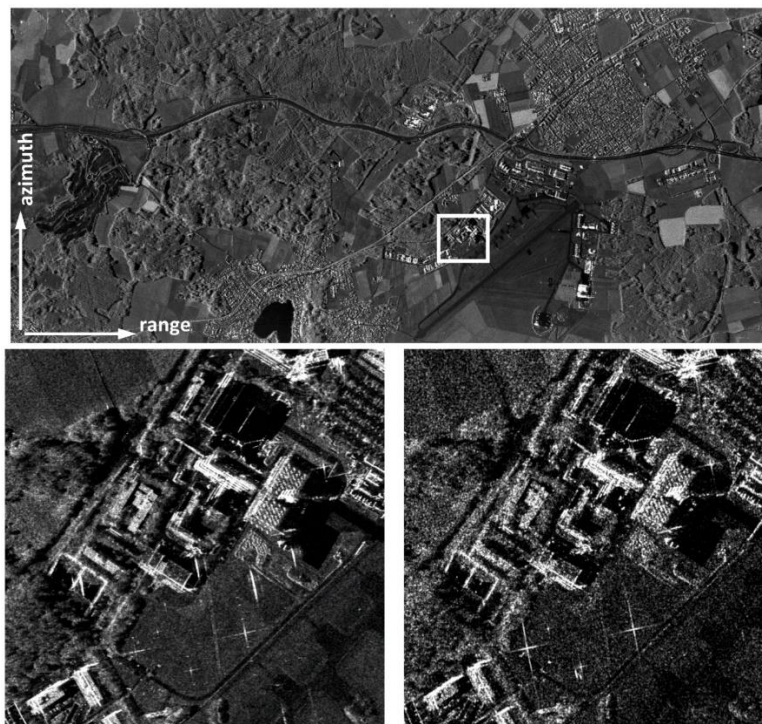


Рисунок 1.3 - Приклад роботи TerraSAR-X [5]

Основні характеристики TerraSAR-X:

- Робоча частота: X-діапазон (9,6-9,9 ГГц);
- Роздільна здатність: від 1 м до 40 м;
- Площа зображення: до 1 000 000 км² на день ;
- Детектування руху з відхиленням від місцевого гравітаційного поля Землі;
- Можливість отримувати зображення в будь-який час доби та при будь-якій погоді;
- Для забезпечення високої якості даних, TerraSAR-X має систему стабілізації та корекції помилок;

TerraSAR-X також може використовуватися для розробки цифрових моделей рельєфу, моніторингу льодовиків та вулканів, виявлення нафти, газу та мінеральних родовищ, а також для запобігати природним катастрофам.

Дані, отримані від TerraSAR-X, доступні для наукових досліджень, а також для комерційних застосувань через спеціалізовані послуги, такі як Airbus Defence and Space. Тобто жоден охочий може за відведену суму грошей орендувати цей супутник

та використовувати його у своїх потребах, що дозволяє реалізувати систему для розпізнавання військової техніки за допомогою супутникових фотографій без додаткових проблем.

1.2.4 WorldView-3

WorldView-3 - це комерційний супутник високої роздільної здатності, запущений компанією DigitalGlobe в серпні 2014 року. Він є одним з найбільш потужних супутників, які в цей час перебувають у космосі, і має здатність робити знімки з висоти 617 км з роздільною здатністю до 31 см на піксель[6]. На рисунку 1.4 показано приклад роботи WorldView-3.



Рисунок 1.4 - Приклад роботи WorldView-3 [6]

Основні характеристики WorldView-3:

- Розширення: 31 см на піксель в режимі панхроматичного зображення і 1,24 м на піксель в режимі мультиспектрального зображення;

- Ширина смуги зображення: до 13,1 км в панхроматичному режимі та до 52,4 км в мультиспектральному режимі;
- Спектральні канали: 8 каналів, що покривають ультрафіолетовий, видимий та інфрачервоний діапазони, включаючи панхроматичний канал;
- Точність позиціонування: 3 м горизонтально та 4,5 м вертикально з 90% інтервалом довіри;
- Час знімання: можливість здійснювати знімки в будь-який час доби;

WorldView-3 використовується для різних цілей, таких як картографування, геологічні дослідження, моніторинг змін клімату, контроль за дотриманням договорів, охорона довкілля, а також для допомоги у врегулюванні різних конфліктів. Його дані використовуються урядовими органами, дослідницькими інститутами, комерційними компаніями та іншими зацікавленими сторонами з усього світу.

1.3 Огляд існуючих методів

Маючи інструмент для отримання зразків фото з технікою треба описати наступну важливу деталь системи, як саме розпізнати щось на фото, а саме техніку. Існує багато методів розпізнавання образів, що базуються на аналізі зображень. Ось декілька методів розпізнавання по фото:

- Метод шаблонів (Template Matching) - порівнює зображення з попередньо заданим шаблоном;
- Метод каскаду Гаара (Haar Cascade) - заснований на використанні класифікатора з групою властивостей зображення (фільтрів);
- Метод глибокого навчання (Deep Learning) - використовується нейронна мережа, щоб автоматично визначати характеристики зображення та розпізнавати об'єкти;

- Метод кластеризації (Clustering) - групує пікселі зображення в класи за допомогою алгоритмів кластеризації, що допомагає виділити об'єкти на зображенні;
- Метод розпізнавання за зовнішніми ознаками (Feature-based recognition) - використовується для розпізнавання об'єктів, що мають визначені зовнішні ознаки (форма, розмір, текстура тощо).

Ці методи можуть використовуватися як окремо, так і в комбінації для досягнення більш точних результатів розпізнавання по фото. Тепер розглянемо окремо кожен метод більш якісніше.

1.3.1 Метод шаблонів

Метод шаблонів (англ. template matching) - це один з найпростіших та найбільш поширених методів розпізнавання образів в комп'ютерному зорі. Цей метод використовується для пошуку та визначення позиції об'єкта на зображенні.

Ідея методу полягає в тому, щоб зіставити вихідне зображення з набором шаблонів, які відповідають пошуку об'єкта. Шаблон може бути простим рисунком або складнішою формою, яка відображає контур об'єкта. Якщо шаблон відповідає ділянці зображення, то можна вважати, що об'єкт знайдено[7].

Наприклад якщо взяти за шаблон зображення монети то на рис. 1.5 алгоритм знайде всі монетки на екрані:

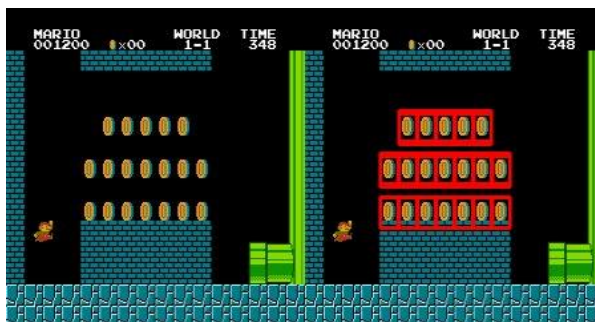


Рисунок 1.5 - Приклад роботи методів шаблону[8]

Для реалізації методу шаблонів використовуються алгоритми, що складаються з наступних етапів:

- Підготовка шаблону та зображення до порівняння (перетворення, зменшення шуму, підсвічування контурів);
- Розрахунок міри подібності (кореляції) між шаблоном та областю зображення. Чим вище значення міри подібності, тим більша ймовірність того, що об'єкт знайдено;
- Визначення точки збігу міри подібності та шаблону, яка відповідає лівому верхньому куту об'єкта;

Однією з основних переваг методу шаблонів є його простота та швидкість виконання. Цей метод вже реалізовано у бібліотеці з відкритим кодом OpenCV. Це робить цей інструмент дуже простим у використанні. Однак, він має деякі недоліки:

- Залежність від шаблону: метод шаблонів дуже залежить від вибраного шаблону. Якщо шаблон не точно відповідає реальному об'єкту, то розпізнавання може бути неточним;
- Велика чутливість до шуму: метод шаблонів не дуже ефективний у випадку, коли зображення містить багато шуму або спотворень;
- Високі вимоги до потужності обчислювальної системи: для того, щоб виконати пошук шаблону великого зображення, необхідно виконати значну кількість операцій, що може займати багато часу та ресурсів;
- Обмежена відображеність об'єктів: метод шаблонів може бути обмеженим у відображенні об'єктів, які не відповідають жодному з вибраних шаблонів;
- Відсутність адаптивності: метод шаблонів не може адаптуватися до нових об'єктів або сценаріїв. Для кожного нового об'єкту необхідно створювати нові шаблони, що може бути дуже час витратно і неефективним.

Тому, хоча метод шаблонів має свої переваги у простоті та зрозумілості, він не завжди є ефективним для розпізнавання образів у різних умовах.

1.3.2 Метод каскаду Гаара

Метод каскаду Гаара (або метод обличчя Хаара) - це алгоритм машинного навчання для детекції облич у зображеннях. Він базується на використанні особливих фільтрів, що називаються фільтрами Хаара, для виділення характеристик обличчя на зображеннях.

Основна ідея методу каскаду Гаара полягає у використанні багатоетапного процесу виявлення облич, де кожен наступний етап уточнює результат попереднього. Спочатку зображення поділяється на блоки, на яких застосовуються фільтри Хаара для визначення наявності характеристик обличчя. Потім, за допомогою класифікатора (наприклад, машини опорних векторів), визначається, чи містить блок обличчя чи ні. Цей процес повторюється кілька разів, де кожен наступний етап зосереджується на визначенні більш точних характеристик обличчя [9]. На рисунку 1.6 показано приклад роботи алгоритму каскаду Гаара.

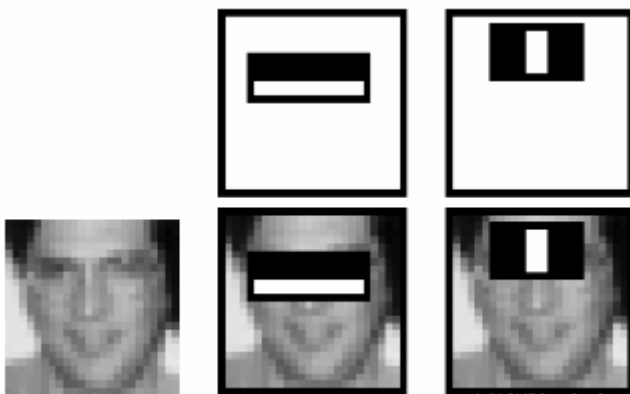


Рисунок 1.6 - Приклад роботи каскаду Гаара[8]

Одним з найбільш важливих переваг методу каскаду Гаара є його швидкість, яка дозволяє обробляти зображення в режимі реального часу. Однак, цей метод може бути досить чутливим до змін освітлення, орієнтації та масштабування зображень.

Крім того, він може пропустити обличчя, які мають нетипові риси або знаходяться під кутом.

Хоч цей метод і є доволі вдалим, але він не є гнучким і працює тільки з зображеннями облич людей. Тож інтегрувати до майбутньої системи навряд чи є можливим.

1.3.3 Метод розпізнавання за зовнішніми ознаками

Метод розпізнавання за зовнішніми ознаками, також відомий як метод розпізнавання образів, є одним з методів машинного навчання, що використовується для класифікації об'єктів за їх зовнішніми ознаками або властивостями. Цей метод базується на тому, що різні об'єкти можуть мати різні ознаки або характеристики, які можуть бути використані для їх розпізнавання [10].

Один з найпоширеніших підходів до методу розпізнавання за зовнішніми ознаками - це використання нейронних мереж. Вони можуть бути навчені розпізнавати різні об'єкти, використовуючи набір даних з відомими зовнішніми ознаками для кожного об'єкта. На рисунку 1.7 показано приклад роботи алгоритму Appearance-Based Recognition.

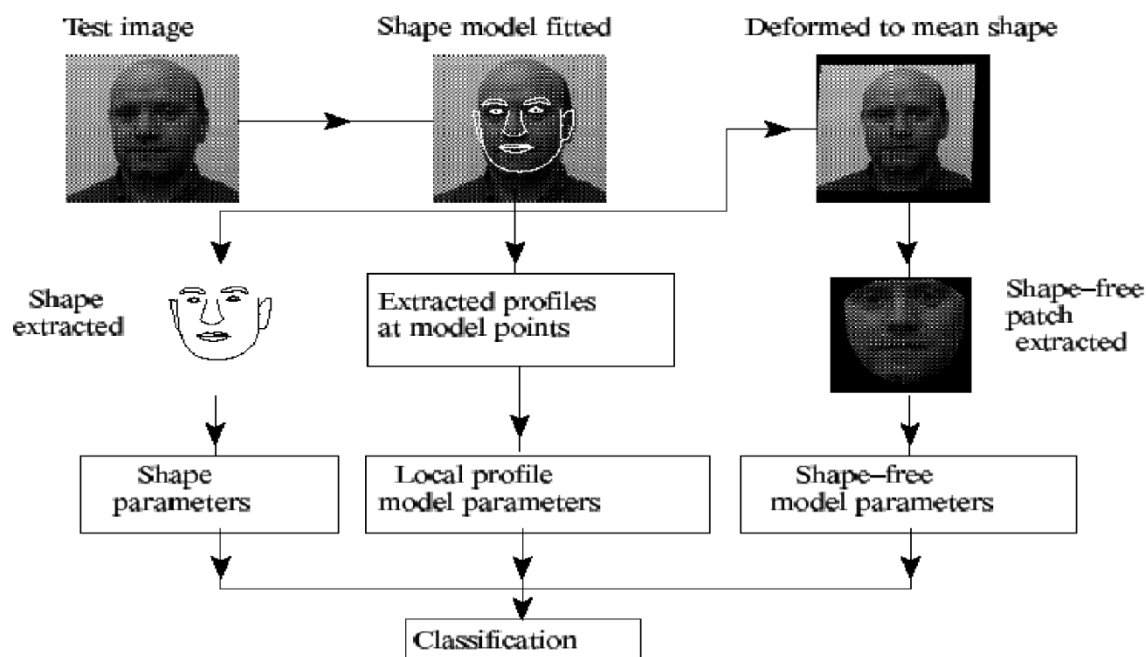


Рисунок 1.7 – приклад роботи Appearance-Based Recognition [8]

Наприклад, для розпізнавання зображень облич людей можна використовувати метод розпізнавання за зовнішніми ознаками, навчивши нейронну мережу розпізнавати особливі ознаки, такі як форма обличчя, розмір очей, ніс та рота.

Інші приклади методу розпізнавання за зовнішніми ознаками включають класифікацію рукописних цифр, розпізнавання мовлення та автоматичне розпізнавання тексту на зображеннях. Метод розпізнавання за зовнішніми ознаками має деякої недолі, зокрема:

- Він може бути не дуже точним, оскільки зовнішні ознаки можуть бути досить суб'єктивними, а також можуть залежати від контексту та інших факторів;
- Цей метод може бути обмеженим в тих випадках, коли зовнішні ознаки не забезпечують достатньої інформації для точного розпізнавання. Наприклад, коли зображення занадто мале або низької якості;
- Метод може бути досить чутливим до змін в зовнішніх ознаках, таких як зміна освітлення, волосся або віддалення від камери. Це може призвести до помилок в розпізнаванні;

- Важко створити набір даних для тренування, оскільки зовнішні ознаки можуть бути дуже різноманітними, а також можуть залежати від культурних, етнічних та інших факторів.

1.3.4 Метод кластеризації

Метод кластеризації - це метод машинного навчання, який використовується для сегментації зображень на декілька кластерів. Основна мета полягає у групуванні пікселів або областей зображення зі схожими характеристиками в окремі класи або кластери, залежно від певних параметрів. Цей метод може бути використаний для багатьох завдань обробки зображень, таких як сегментація зображень, компресія зображень, визначення ознак та інших[11].

Одним з основних підходів до кластеризації є k -середніх алгоритмів, який розділяє зображення на підмножини або кластери, використовуючи евклідову відстань між пікселями. У цьому алгоритмі спочатку випадковим чином вибираються k центрів кластерів, а потім кожен піксель призначається до ближчого центру. Потім позиції центрів перераховуються на основі середньої відстані пікселів до кожного центру, і процес повторюється до тих пір, поки позиції центрів не залишаються стабільними. На рисунку 1.8 показано приклад роботи алгоритму кластеризації.



Рисунок 1.8- Приклад роботи кластеризації для фото будівлі [8]

Метод кластеризації для зображень має кілька недоліків, серед яких:

- Специфікація кількості кластерів: у методі кластеризації потрібно заздалегідь вказати кількість кластерів, що потрібно створити. Якщо кількість кластерів заздалегідь не відома, визначення оптимальної кількості кластерів може бути складним завданням;
- Залежність від початкового вибору центроїдів: визначення початкових центроїдів кластерів може вплинути на кінцевий результат. Якщо початкові центроїди обрані недостатньо точно, можуть виникнути неправильні кластери або кластери можуть бути зміщені;
- Підходить тільки для статичних зображень: метод кластеризації підходить тільки для статичних зображень, які не залежать від часу. Для аналізу динамічних зображень, які залежать від часу, потрібні більш складні методи;
- Не враховує семантичний зміст: метод кластеризації аналізує зображення лише за пікселями і не враховує семантичний зміст зображення. Якщо необхідно здійснити аналіз зображення на основі його семантичного змісту, можуть знадобитися більш складні методи, такі як нейронні мережі.

1.3.5 Метод глибокого навчання

Метод глибокого навчання (Deep Learning) - це один із найефективніших методів розпізнавання об'єктів на фото, що базується на нейронних мережах.

Глибокі нейронні мережі складаються з багатьох шарів, які навчаються розпізнавати різноманітні ознаки на зображенні на різних рівнях абстракції. Наприклад, на першому рівні мережа може виявляти різні кольори та геометричні форми, на другому - текстури та складніше складені об'єкти, а на останньому рівні - цілі об'єкти, такі як автомобілі, люди, тварини та інші[12].

Для навчання глибоких нейронних мереж використовуються великі набори зображень з позначками, що містять інформацію про те, який об'єкт зображений на фото. Після навчання мережа може застосовуватись для розпізнавання об'єктів на нових фото.

Навчання проходить у декілька етапів:

- Збір і підготовка даних, на початку необхідно зібрати достатню кількість даних, які включають зображення з об'єктами, що потрібно розпізнати, та зображення без об'єктів. Ці дані потрібно підготувати до навчання шляхом нормалізації та обрізки;
- Побудова моделі, далі необхідно побудувати модель глибокого навчання. Це зазвичай здійснюється за допомогою нейронних мереж, які вивчають зображення з об'єктами та без них і встановлюють залежності між зображенням та його вмістом;
- Навчання моделі, після побудови моделі проводиться навчання, яке передбачає подачу зображень на вхід мережі та оцінювання результату, який порівнюється з очікуваним вихідним значенням. Навчання повторюється доти, поки модель не досягне задовільної точності;

- Оцінка та тестування моделі, після навчання модель оцінюється та тестується за допомогою даних, які не використовувалися для навчання. Це дозволяє перевірити, наскільки добре модель може розпізнавати об'єкти на нових зображеннях.

- Використання моделі, нарешті, після успішного навчання та оцінки модель може бути використана для розпізнавання об'єктів на фото у реальному часі.

Одним із головних переваг методу глибокого навчання є його здатність до автоматичного виявлення та вивчення корисних ознак для розпізнавання об'єктів, що дозволяє отримати дуже точні результати на великих наборах даних. Крім того, глибокі нейронні мережі можуть використовуватись для розпізнавання об'єктів у реальному часі та для вирішення задач обробки зображень, таких як сегментація та генерація зображень.

Щодо відомих прикладів для застосування Deep Learning у задачах розпізнавання об'єктів на фото треба згадати:

- Розпізнавання облич, DeepFace від Facebook та FaceNet від Google - це два з найбільш відомих методів розпізнавання облич. Вони використовують глибокі нейронні мережі для визначення особливостей облич та встановлення зв'язку між різними фотографіями тієї ж самої людини;

- Розпізнавання об'єктів у відео: YOLO (You Only Look Once) - це алгоритм, що використовує глибокі нейронні мережі для розпізнавання об'єктів у відео. Він здатний розпізнавати об'єкти в реальному часі та з високою точністю;

- Детекція та класифікація медичних зображень: DeepLesion - це система, що використовує глибокі нейронні мережі для розпізнавання та класифікації ракових ознак на медичних зображеннях. Ця система допомагає лікарям швидше та точніше діагностувати рак та вирішувати питання щодо лікування;

- Розпізнавання зображень з дронів: Dronet - це система, що використовує глибокі нейронні мережі для розпізнавання об'єктів на зображеннях, знятих з дронів. Вона може допомогти у військовій розвідці, забезпеченні безпеки на різних об'єктах та ділянках, а також в агрокультурному виробництві.

Наразі цей метод є дуже точним та сучасним, і для його реалізації є дуже багато готових інструментів, за допомогою яких можливо побудувати майбутню систему.

1.4 Аналіз існуючих рішень

Оскільки ціллю даної роботи є побудова застосунку з можливістю розпізнавати військової техніки за супутниковими знімками, необхідно розглянути існуючі рішення, виокремити їхні переваги та недоліки. На сьогодні існує багато рішень для розпізнавання об'єктів на фото, які є комерційно вдалими та більшість мають хмарну систему, тобто працюють на хмарі. Але майже в більшості є внутрішні обмеження для використання у військових цілях.

1.4.1 Google Lens

Google Lens - це технологія компанії Google для розпізнавання об'єктів на фотографіях та відео. Її можна використовувати на смартфонах з операційною системою Android та iOS, а також в деяких інших програмах та пристроях, таких як камери на смартфонах Pixel та Google Nest Hub Max.

Google Lens використовує технологію глибокого навчання для розпізнавання об'єктів, тексту та інформації на фотографіях. Завдяки цьому вона може розпізнавати різні типи об'єктів, такі як рослини, тварини, логотипи, книги, пам'ятники та інші[13].

Однією з основних функцій Google Lens є розпізнавання тексту, що може бути корисним в різних ситуаціях. Також Google Lens може розпізнавати текст на фото візиток, що дозволяє швидко додати новий контакт до свого телефону. На рисунку 1.9 показано приклад роботи Google Lens.

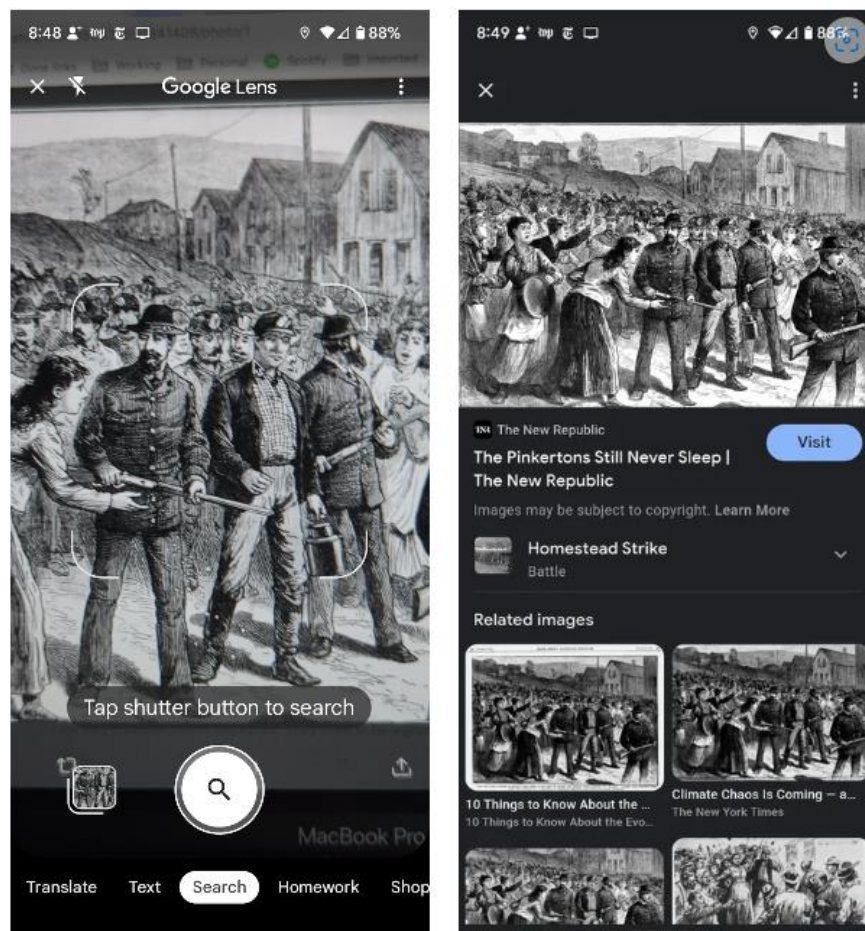


Рисунок 1.9 – Приклад роботи Google Lens [13]

До недоліків можна віднести:

- Неспроможність використання даного застосунку в системах пошуку військової техніки;
- погане розпізнавання об'єктів, хоча Google Lens має велику базу даних об'єктів, які може розпізнати, він не завжди може правильно ідентифікувати об'єкти на фото, особливо якщо фото не чітке або має складний фон;
- Проблеми з конфіденційністю, Google Lens може збирати інформацію про користувачів та їх фотографії, що може бути проблемою для людей, які стежать за своєю приватністю.

1.4.2 Amazon Rekognition

Amazon Rekognition - це сервіс розпізнавання облич, об'єктів та сцен, який розробляється Amazon Web Services. Він використовує глибоке навчання для аналізу зображень та відео з метою ідентифікації та аналізування вмісту.

Завдяки Amazon Rekognition можна здійснювати розпізнавання облич, виявляти інформацію про людей, виявляти об'єкти на зображеннях та відео, знаходити подібні зображення в бібліотеці, а також здійснювати аналіз настрою та емоцій на зображеннях.

Цей сервіс використовується в багатьох галузях, таких як безпека, реклама, медіа, маркетинг та багато інших. Він може бути корисним для підприємств, які потребують автоматизації процесів ідентифікації та аналізу зображень та відео.

До недоліків можна віднести:

- Неспроможність використання даного застосунку в системах пошуку військової техніки;
- Вартість, Використання Amazon Rekognition може бути досить дорогим для певних бізнесів або організацій, особливо якщо вони мають потребу в більшій кількості операцій з розпізнавання зображень.

1.5.2 IBM Watson Visual Recognition

IBM Watson Visual Recognition - це хмарна платформа для розпізнавання об'єктів на зображеннях. Вона використовується для створення рішень у таких галузях як медицина, автомобільна промисловість, роздрібна торгівля та інше.

IBM Watson Visual Recognition може розпізнавати різноманітні об'єкти на зображеннях, такі як тварини, люди, автомобілі, будівлі, пейзажі та інше. Він також

може розпізнавати обличчя та аналізувати вікові ознаки, стать та емоційний стан людей на зображеннях[14].

IBM Watson Visual Recognition працює на основі глибоких нейронних мереж, які навчаються на великій кількості зображень. Платформа може бути легко інтегрована в існуючі програми та вебсайт за допомогою API.

Одним з головних переваг IBM Watson Visual Recognition є висока точність розпізнавання об'єктів та швидкість обробки зображень. Також платформа має велику базу зображень для навчання своїх моделей розпізнавання, що дозволяє їй використовувати глибоке навчання для отримання кращих результатів

До недоліків можна віднести:

- Питання конфіденційності даних, оскільки дані, надіслані до сервісу IBM Watson Visual Recognition, можуть містити конфіденційну інформацію, користувачі можуть мати сумніви щодо того, як ці дані будуть використовуватися та зберігатися;
- Високі витрати, IBM Watson Visual Recognition може бути досить дорогим для користувачів з обмеженим бюджетом;
- Неспроможність використання даного застосунку в системах пошуку військової техніки.

Результат аналізу існуючих рішень з функціоналом розпізнавання об'єктів на фото наведений у таблиці 1.1.

Таблиця 1.1 – Порівняння існуючих рішень

Критерій порівняння	Google Lens	IBM Watson Visual Recognition	Amazon Rekognition
Хмарний доступ	+	+	+
Можливість розширення функціоналу	-	-	-
Конфіденційність	+	-	-
Відкрите API	-	-	-

Підсумовуючи результати аналізу декількох існуючих рішень, жоден з них не підходить для інтеграції в системи розпізнавання військової техніки за супутниковими знімками. Жоден додаток неможливо розширити додатковим функціоналом. Для майбутнього додатку важливими критеріями будуть:

- простота використання;
- можливість інтеграції в системи розпізнавання військової техніки за супутниковими знімками.

1.6 Висновки до розділу

У даному розділі було поставлено бізнес цілі, описанні методи фіксацій техніки за допомогою супутника з SAR апаратурою. Розглянуто діючі супутники які можливі у використанні в системі. Детально описано про існуючі методи розпізнавання об'єктів по фото, та проаналізовано існуючі рішення, підкреслені їх переваги та недоліки. З'ясовано? що майбутня модель буде використовувати нейронну мережу з використанням глибокого навчання та буде мати свою систему конфіденційності.

2 ВІДОМОСТІ ТА РІЗНОВИД ВІЙСЬКОВОЇ ТЕХНІКИ

Різнovid військової техніки є дуже величезний і зробити уніфіковану систему неможливо, але нещастя це можливо зробити для конкретного автопарку окремої країни. Очевидно що в даній роботі це буде автопарк срср/рф. Бо в даних реалі не має сенсу робити таку систему для техніки НАТО. Завдяки американським військові розвідці в 1990-х років було зібрано фотографії за допомогою технології SAR більшості розповсюдженої техніки радянського союзу. В знайденому датасеті є наступні класи:

- 2S1
- BRDM-2
- BTR-60
- T-62
- ZIL-131
- ZSU-23-4 Shilka

Також в цьому наборі даних є фотографії звичайного трактору та макету техніки як варіант для перевірки, щоб система мала змогу розуміти різницю між військовою технікою та господарсько.

Ще важливою особливістю є те, що фотографії зроблені за допомогою технології SAR є не дуже зрозумілими для звичайного ока, приклад цих фотографій можна побачити нижче:

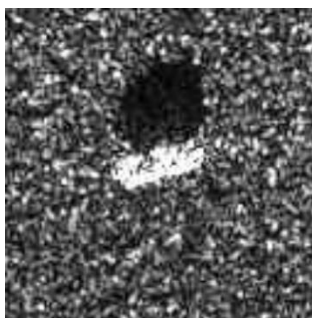


Рисунок 2.1 - Приклад фото з набору даних

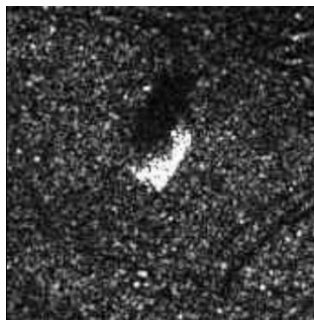


Рисунок 2.2 - Приклад другого фото з набору даних

Але для нейронної мережі це зовсім не є проблемою. Отже, подивимось які саме види техніки має датасет та зробимо окремо опис кожного типу техніки.

2.1 2С1 Гвоздика

2С1 «Гвоздика» — 122-мм самохідна гаубиця російського виробництва на базі гусеничної броньованої машини. Перший прототип був готовий у 1969 році. 2С1 надійшов на озброєння радянської армії на початку 1970-х років і вперше був показаний на параді польської армії в 1974 році. 2С1 Гвоздика має екіпаж із чотирьох солдатів, включаючи командира, навідника, вантажника, водія. 2С1 завжди використовується багатьма збройними силами по всьому світу, як і ВСУ, так в армії рф. Виготовлена кількість машин складає понад 10 000 одиниць. Приклад даної техніки зображено на рис. 2.3:



Рисунок 2.3 - Приклад фото 2С1 [15]

2S1 озброєний модифікованою версією гармати 2А31, яка також встановлена до буксирної гаубиці D-30. Зовнішній сталевий корпус 2S1 розділений на три відсіки з водієм спереду зліва, двигуном за водієм, і баштою ззаду. Система підвіски 2S1 Гвоздики схожа на багатоцільовий відстежений транспортний засіб МТ-ЛВ, і будь-яка сторона складається з семи дорожніх коліс із зірочкою приводу в передній частині, холостого ходу ззаду і без переробки доріжок. Незвичайна особливість М-1974 полягає в тому, що підвіска може бути відрегульована, щоб надати різній висоті, що особливо використовується, коли транспортний засіб транспортується тактичним транспортним літальним апаратом.

Приклад фото у датасеті на рис. 2.4 та рис. 2.5:

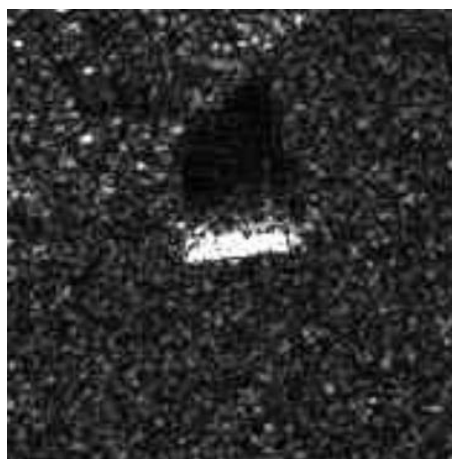


Рисунок 2.4 - Приклад фото 2С1 у датасеті

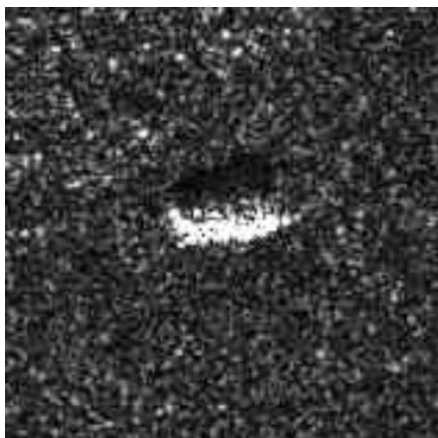


Рисунок 2.5 - Приклад другого фото 2С1 у датасеті

2.2 BRDM-2

BRDM-2-це легкий броньований автомобіль 4х4, розроблений В.К. Рубцовим з ОКБ Дедкова на початку 1960-х. Ця машина надійшла на озброєння радянської армії в 1962 році з першим виробництвом в 1963 році і була вперше показана публіці в 1966 році. Всього виготовлено 9400 одиниць. Приклад даної техніки зображено на рис. 2.6:



Рисунок 2.6 - Приклад фото БРДМ-2 [15]

БРДМ-2 в основному використовується як розвідувальна броньована машина і все ще знаходиться на озброєнні Росії збройними силами та багатьма іншими країнами по всьому світу. Тисячі транспортних засобів будуть виготовлені та використані на полі під час сучасних конфліктів з 1960 року до сьогодні. Машина може вмістити до п'яти військовослужбовців. За останні десять років він зазнав кілька модернізацій, головним чином завдяки новому двигуну та покращеній броні. Також можна звернути увагу, що є багато модифікації, на яких БРДМ-2 має замість кулемета протитанкові або протиповітряні ракети.

Приклад фото у датасеті на рис. 2.7 та рис. 2.8:



Рисунок 2.7 - Приклад BRDM-2 у датасеті

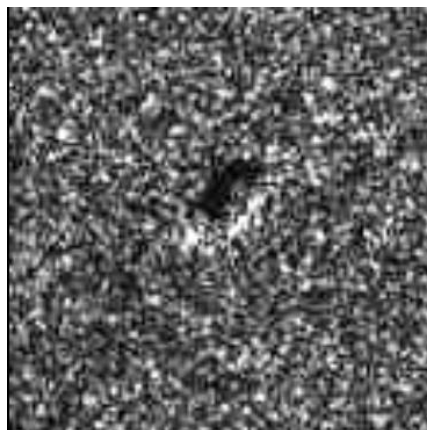


Рисунок 2.8 – Приклад BRDM-2 у датасеті №2

2.3 BTR-60

Бронетранспортер БТР-60 є подальшим розвитком попереднього БТР-152. Виготовлявся великими тиражами для Радянської Армії та радянських союзників. Виробництво почалося в 1966 році і припинилося в 1976 році. Довгий час БТР-60 використовувався для транспортування радянських мотопіхотних військ. Його також експортували до ряду країн. Незважаючи на свій вік, БТР-60 все ще перебуває на озброєнні щонайменше 30 країн світу. Виготовлено більш ніж 25 000 одиниць. Приклад даної техніки зображено на рис. 2.9:



Рисунок 2.9 – БТР-60 [15]

Цей бронетранспортер має сталевий зварний корпус. У порівнянні з попередньою моделлю БТР-60ПА покращено броньовий захист. БТР-60ПБ забезпечує захист від вогню стрілецької зброї та осколків артилерійських снарядів. Повний захист від снарядів калібру 7,62 мм, випущених з дистанції 100 метрів. У бойовому порядку передні лобові скла закриті броньованими жалюзі. БТР-60ПБ оснащений системою РХБ захисту.

Приклад фото з датасета на рис. 2.10 та рис 2.11:

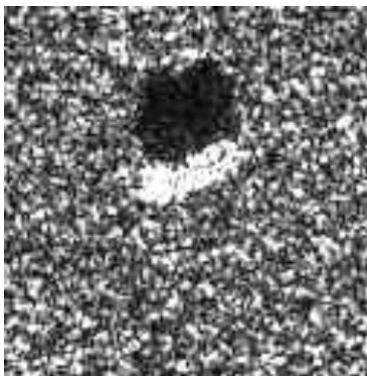


Рисунок 2.10 - Приклад БТР-60 у датасеті



Рисунок 2.11 –Приклад 2 БТР-60 у датасеті

2.4 T-62

T-62 — це основний бойовий танк (ОБТ) російського виробництва, який є подальшим розвитком основного бойового танка Т-55. Т-62 випускався між 1961 і 1975 роками. Він став стандартним основним бойовим танком у російських збройних силах і Варшавському договорі під час холодної війни, частково замінивши Т-55, хоча цей танк продовжував виготовлятися в росії та в інші країни світу. Основний бойовий танк Т-62 був прийнятий на озброєння російської армії в 1961 році. Т-62 завжди стоїть на озброєнні російської армії, але в оновленому або модернізованому варіанті. Багато збройних сил світу використовують Т-62 як стандартний основний бойовий танк. Виготовлено 20 000 шт. Приклад даної техніки зображено на рис. 2.12:



Рисунок 2.12 - Т-62 [15]

Основний бойовий танк Т-62 має класичне компонування. Корпус Т-62 розділений на три відділення: переднє відділення механіка-водія, центральне — бойове, а двигун і трансмісія — ззаду. Лобові бронеплити мають значний кут нахилу: нахил glacis становить 60° , а нахил нижнього листа — 54° . Обидва мають товщину 100 мм.

Приклад фото з датасету на рис. 2.13 та рис. 2.14:

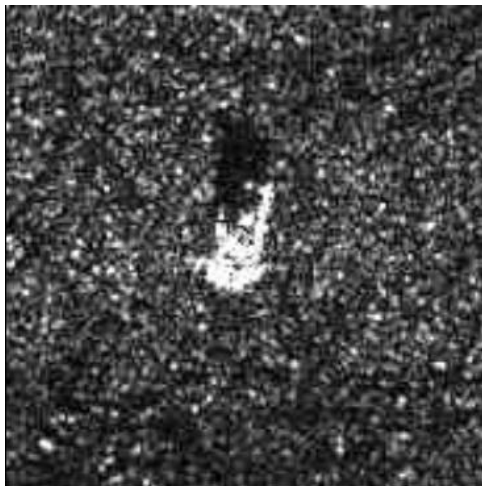


Рисунок 2.13 - Приклад Т-62 у датасеті

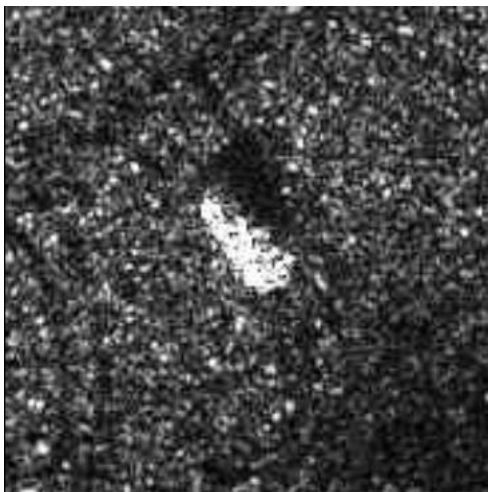


Рисунок 2.14 – Приклад 2 Т-62 у датасеті

2.5 Т-72

Т-72 — сімейство радянських основних бойових танків, які почали виготовляти в 1969 році. Т-72 був розвитком Т-64, який був стурбований високою вартістю та залежністю від незрілої технології розробки. Було побудовано близько 25 000 танків Т-72, а модернізація дозволила багатьом з них залишатися на озброєнні протягом десятиліть. Він широко експортувався і використовувався в 40 країнах і в численних конфліктах. Російський Т-90, представлений у 1992 році, і китайський Тип 99 є подальшим розвитком Т-72. Виробництво і розробка різних модернізованих моделей Т-72 триває і сьогодні. Приклад даної техніки зображено на рис. 2.15:



Рисунок 2.15 – Т-72 [15]

Загальна компоновка Т-72А схожа на всі радянські серійні танки: відділення механіка-водія розташоване спереду, бойова вежа — по центру, двигун і трансмісія — ззаду. Т-72А має екіпаж з трьох осіб, причому водій сидить спереду, а інші два члени екіпажу сидять у вежі, навідник ліворуч і командир праворуч.

Приклад фото з датасету на рис. 2.16 та рис. 2.17:

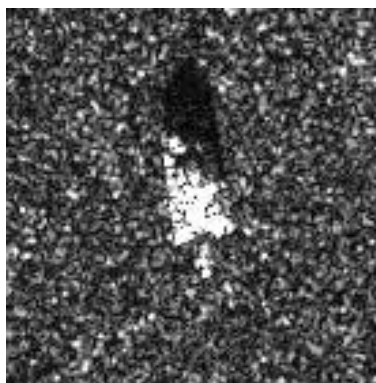


Рисунок 2.16 - Приклад Т-72 у датасеті

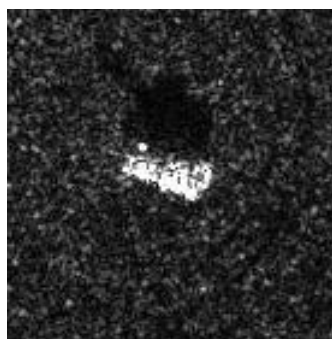


Рисунок 2.17 - Приклад Т-72 у датасеті

2.6 ZSU-23-4 Shilka

ЗСУ-23-4 «Шилка» — російський легко броньований самохідний зенітний комплекс з радіолокаційним наведенням. ЗСУ-23-4 «Шилка» була розроблена на початку 1960-х років і вперше була помічена на публіці під час параду в Москві в

листопаді 1965 року. ЗСУ-23-4 виявилася високоефективною під час війни на Близькому Сході 1973 року і нарахувала велику кількість озброєнь. Втрати ізраїльських літаків. ЗСУ-23-4 також використовувалися як у Північному, так і в Південному В'єтнамі, а також у війні в Перській Затоці іранцями та іракцями. Виготовлено 6 500 шт. Приклад даної техніки зображено на рис. 2.18:



Рисунок 2.18 – ЗСУ-23-4 [15]

ЗСУ-23-4 Шилка створена на базі шасі гусеничної машини ГМ-575, на якій використані компоненти легкого танка-амфібії ПТ-76. Двигун і трансмісія розташовані в задній частині корпусу, як і газова турбіна DG4M-1 у поєднанні з механічною коробкою передач з 5 передачами вперед і 1 передачею назад. Торсіонна система підвіски складається з шести одинарних опорних коліс з гумовими шинами з натяжним колесом спереду та ведучою зірочкою ззаду.

Приклад фото у датасеті на рис. 2.19 та рис. 2.20:

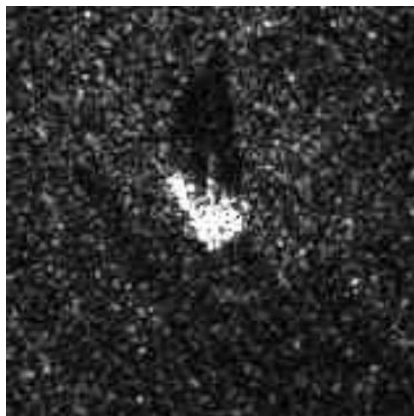


Рисунок 2.19 - Приклад ЗСУ-23-4 у датасеті

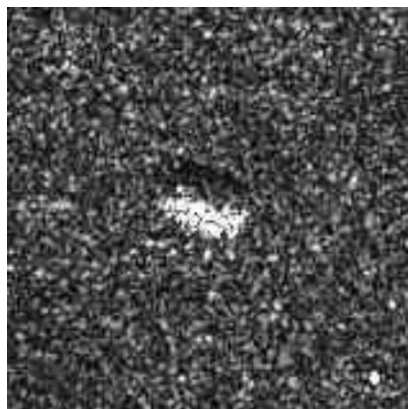


Рисунок 2.20 - Приклад ЗСУ-23-4 у датасеті

2.7 Zil-131

ЗІЛ-131 — військова вантажівка загального призначення. Він був розроблений на основі комерційної вантажівки ЗІЛ-130 з повною колісною формулою, який випускався масово. Виробництво ЗІЛ-131 почалося в 1966 році. Він прийшов на зміну попередньому ЗІЛ-157. Виробництво було припинено в 1994 році, випущено майже 1 000 000 таких вантажівок. Його широко використовували радянські військові. Незважаючи на свій вік, ЗІЛ-131 все ще використовується російськими військовими, а також ряду інших країн.

Автомобіль має звичайне компонування. Базова модель ЗІЛ-131 - вантажний автомобіль. Він має вантажність 3 500 кг на бездоріжжі та 5 000 кг на дорогах з твердим покриттям. Використовується для перевезення військ, вантажів та іншої військової техніки. Вантажний відсік прикритий дугами і брезентовим покриттям. Автомобіль також може буксирувати причепи або артилерійські знаряддя максимальною вагою 4 000 кг на бездоріжжі та 6 500 кг на дорогах з твердим покриттям. Приклад даної техніки зображено на рис. 2.21:



Рисунок 2.21 – Зіл-131 [15]

2.8 Макети

Для того, щоб навчити нейронну мережу відрізнити військову техніку від звичайних тракторів та просто макетів, в даному датасеті присутні і їх фотографії. А саме трактора американського виробництва D7 та мішень «Slicy».

Мішень «Slicy» — це точно розроблена та оброблена інженерна тестова мішень, що містить стандартні примітивні форми радіолокаційного відбивача, такі як плоскі пластини, двогранники, тригранники та циліндри. Вона зображена на рис. 2.22:



Рисунок 2.22 – Фальшива ціль [16]

2.9 Висновки до розділу

В цьому розділі були розглянута військова техніка, що буде програмно розпізнаватися, їх зовнішній вигляд, клас техніки та їх розповсюдженість. Це набір з 9 об'єктів: 2С1 Гвоздика, Т-72, БРДМ-2, БТР-60, Т-62, ZSU-23-4 Shilka, Zil-131, D7, Slicy. У таблиці 2.1 описані характеристики кожного виду техніки та ознаки кожної з них.

Також, було описано будову та для чого саме призначена техніка. До кожного опису додавались зображення з візуальним представленням кожної техніки.

Таблиця 2.1 – Збірка головних характеристик зазначеної техніки

Назва техніки	Характеристики
2С1 Гвоздика	Самохідна артилерійська установка, виготовлена у розмірі понад 10 000 одиниць.
БРДМ-2	Бойова розвідувальна-дозорна машина, виготовлена у розмірі 9 400 одиниць.

Назва техніки	Характеристики
БТР-60	Найпоширеніший колісний бронетранспортер, виготовлено у розмірі 25 000 одиниць.
T-62	Основний бойовий танк другого покоління, виготовлено у розмірі 22 700 одиниць.
T-72	Основний бойовий танк другого покоління, заміна T-62, виготовлено у розмірі понад 30 000 одиниць у різних модифікаціях.
ZSU-23-4 Shilka	Зенітна самохідна установка, що призначена для безпосереднього протиповітряного прикриття наземних військ, виготовлено у розмірі 6 500 одиниць.
Zil-131	Радянський вантажний автомобіль, створений для потреб логістики армії, також може використовуватись як транспорт для піхоти.
D7	Середній гусеничний трактор, поширений у всьому світі.
Slicy	Інженерна тестова мішень, що містить стандартні примітивні форми радіолокаційного відбивача.

3 ОГЛЯД ТА ОБҐРУНТУВАННЯ МАТЕМАТИЧНИХ МЕТОДІВ

3.1 CNN

Згорткові нейронні мережі (Convolutional Neural Networks або CNNs) - це один з типів глибоких нейронних мереж, які використовуються для обробки зображень і відео. Вони мають здатність розпізнавати зображення та розуміти його зміст.

Основна ідея застосування згорткових нейронних мереж полягає в тому, щоб використовувати згортки, які знаходять локальні особливості зображення, такі як лінії, кути, форми тощо. Згорткові шари в згорткових нейронних мережах здійснюють операцію згортки з фільтрами (ядрами), що дає можливість витягнути різні властивості зображення[17]. Структура роботи CNN зображена на рис. 3.1:

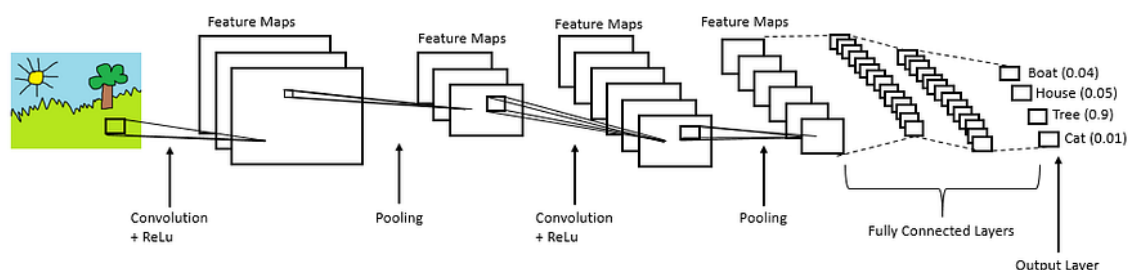


Рисунок 3.1- Структура роботи CNN [17]

Крім згорткових шарів, згорткові нейронні мережі також мають пулінгові шари, що допомагають зменшити кількість параметрів мережі та зробити її більш робастною до невідомих даних. В пулінговому шарі використовується операція пулінгу для зменшення розмірності зображення. Зазвичай, найпоширеніша операція пулінгу - максимальний пулінг, що зводить відстань між відповідями фільтра на мінімальний рівень.

Після декількох згорткових та пулінгових шарів зазвичай використовуються повнозв'язні шари (FC), щоб згенерувати фінальний вектор ознак. Повнозв'язні шари приймають вектор ознак і виконують класифікацію на основі нього.

Кількість ядер у згортковому шарі залежить від розміру вхідного зображення та кількості ознак, які ми хочемо виділити. Наприклад, якщо ми працюємо з кольоровим зображенням розміром 224x224 пікселів, то ми можемо застосувати до нього кілька згорткових шарів з декількома ядрами кожен, які будуть виділяти різні ознаки зображення, такі як контури, текстури та форми об'єктів. Приклад роботи фільтра з зображенням на рис. 3.3:

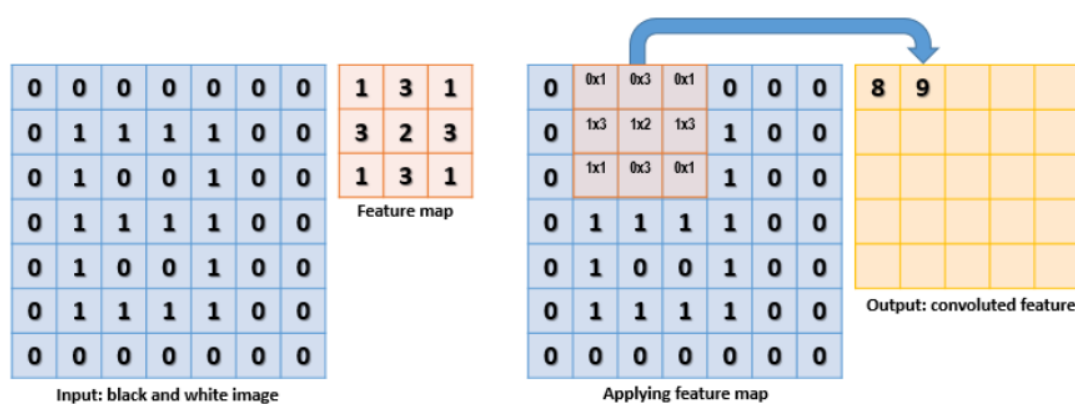


Рисунок 3.3 – Приклад роботи фільтра з зображенням [17]

Для уникання перенавчання, покращення роботи мережі, та підвищення стійкості використовується додатковий шар або його ще називають метод *dropout*, який виключає з себе деякі нейронні зв'язки[18].

Dropout - це метод регуляризації в згорткових нейронних мережах та інших глибоких нейронних мережах, який допомагає уникнути перенавчання (*overfitting*).

Основна ідея *Dropout* полягає в тому, що в процесі навчання нейронна мережа випадковим чином викидає з певною ймовірністю (наприклад, 0,5) деякі нейрони та їх зв'язки, які не беруть участь у обчисленні ваги та зсуву на наступному шарі. Це допомагає уникнути залежності між певними нейронами, тому що мережа повинна вчитися уважно на корисну інформацію, яка знаходиться в багатьох різних шляхах крізь мережу, а не тільки через один конкретний шлях[19]. Схематичне зображення *Dropout* на рис. 3.4:

Neural network dropout: Before and after

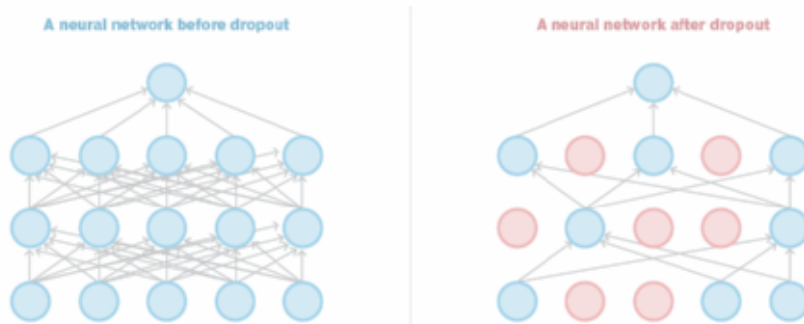


Рисунок 3.4 – Схематичне зображення Dropout [19]

Dropout можна застосовувати на різних шарах нейронної мережі, а не тільки на повнозв'язних шарах, і це може допомогти зменшити кількість параметрів, що потрібні для навчання мережі, зменшити час навчання та запобігти перенавчанню.

3.2 Персептрон

Персептрон - це проста форма штучної нейронної мережі, яка є основним будівельним блоком глибокого навчання. Він складається з вхідного шару, вихідного шару та шару прихованих нейронів, які зв'язуються між собою за допомогою зважених з'єднань. Кожен нейрон в прихованому та вихідному шарах має активаційну функцію, яка залежить від зваженого додавання вхідних сигналів. Причому ваги з'єднань між нейронами змінюються під час навчання для досягнення оптимальної функціональності мережі[20]. Візуалізація будови моделі персептрона на рис 3.5:

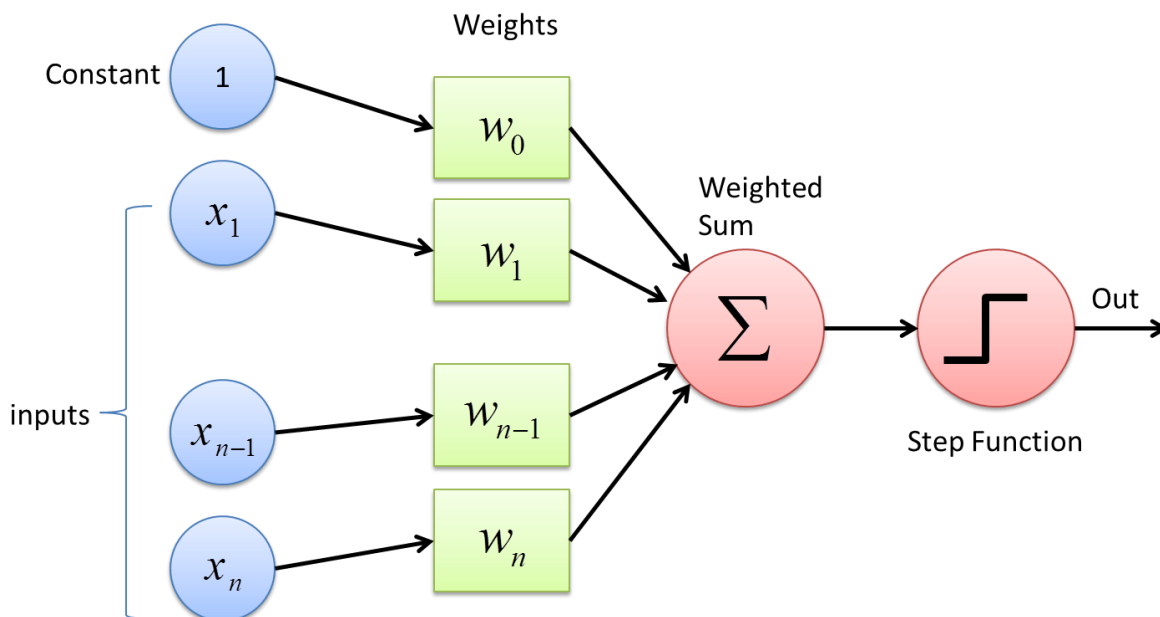


Рисунок 3.5 – Візуалізація будови моделі перцептрона [20]

Перцептрон може розв'язувати прості задачі класифікації і регресії. Наприклад, для класифікації зображень перцептрон приймає на вхід пікселі зображення та виводить клас, який він вважає, що зображення належить.

Однак, перцептрони не завжди здатні вирішувати складні завдання, такі як обробка природних мов або розпізнавання зображень, через їх обмежену архітектуру. Тому для вирішення більш складних завдань використовують глибокі нейронні мережі, що складаються з багатьох перцептронів, з'єднаних в складну топологію.

3.3 Метрики

3.3.1 Categorical Accuracy

CategoricalAccuracy - це метрика, яка використовується для оцінки якості роботи нейронної мережі у задачах класифікації з багатьма класами. Вона вимірює точність прогнозів моделі, порівнюючи передбачення з відповідними мітками класів.

Якщо передбачення моделі співпадає з міткою класу, то метрика збільшується на одиницю, інакше - залишається незмінною. Наприклад, якщо модель правильно передбачає класи для 8 з 10 зразків, то метрика CategoricalAccuracy буде дорівнювати 0.8.

3.3.2 Weighted Average F2 Score

Weighted Average F2 Score - це метрика, яка використовується для оцінки якості роботи моделі у задачах багатокласової класифікації з незбалансованими даними. Вона обчислюється як середнє гармонійне між точністю і повнотою, де більший ваговий коефіцієнт надається меншій класі.

$$F = \frac{1 + \beta^2}{\frac{\beta^2}{Recall} + \frac{1}{Precision}} \quad (3.1)$$

Наприклад, якщо у задачі є 3 класи, і кількість зразків в кожному класі різна (200, 100, 50), то для обчислення Weighted Average F2 Score будуть використовуватись вагові коефіцієнти (0.5, 0.25, 0.125), які відображають співвідношення кількості зразків у кожному класі. Таким чином, меншому класу буде присвоєно більший ваговий коефіцієнт, щоб він мав більший вплив на значення метрики.

3.3.3 Matthews Correlation Coefficient

Matthews Correlation Coefficient (MCC) - це метрика, яка використовується для оцінки якості бінарної класифікації. Вона базується на матриці помилок, яка показує, скільки зразків було правильно/неправильно класифіковано моделлю (рис. 3.6).

		Передбачені	
Дійсні	Істинно негативні	Хибно позитивні	
	Хибно негативні	Істино позитивні	

Рисунок 3.6 – Матриця помилок

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} [21] \quad (3.2)$$

MCC враховує не тільки точність передбачення моделі, а й дисбаланс між класами. Вона приймає значення в діапазоні від -1 до +1, де +1 означає ідеальну кореляцію між спостережуваними та передбаченими класами, 0 - випадкову кореляцію, а -1 - повну протилежність між класами.

3.4 Підбір функції оптимізації

Оптимізація в нейронних мережах - це процес налаштування параметрів моделі, щоб зменшити помилку передбачення і покращити її ефективність. Цей процес

зазвичай містити вибір оптимального алгоритму оптимізації та налаштування його параметрів, таких як швидкість навчання, коефіцієнти регуляризації та інші.

Навчання мережі проходить за допомогою алгоритму зворотного розповсюдження помилки. Він використовується для обчислення градієнта функції втрат відносно кожного вагового коефіцієнта в мережі. Цей градієнт потім використовується для оновлення вагових коефіцієнтів за допомогою алгоритму градієнтного спуску, що допомагає зменшити значення функції втрат і збільшити точність передбачень моделі [20]. Але цей процес може бути дуже різноманітним тому існує багато алгоритмів оптимізації, обравши котрий можна суттєво змінити властивості цілої мережі.

3.4.1 Adam

Adam - це адаптивний оптимізатор градієнтного спуску для нейронних мереж, який зазвичай використовується в глибокому навчанні. Він був запропонований в 2015 році.

Adam об'єднує інтегрування та адаптивний підбір норми кроку, тобто він підлаштовує швидкість навчання для кожного параметру окремо на основі статистики другого моменту градієнтів. Це дозволяє ефективно налаштовувати параметри моделі, що може зменшити час навчання та покращити точність.

Його назва походить від оцінки адаптивного моменту , і причина, чому вона так називається, полягає в тому, що Адам використовує оцінки першого та другого моментів градієнта, щоб адаптувати швидкість навчання для кожної ваги нейронної мережі[22]. Основне поняття моменту випадкової величини визначається як очікуване значення цієї змінної в степені n :

$$m_e = E[X^n] \quad (3.3)$$

Де m – це момент, а X випадкова величина.

Щоб оцінити моменти, Адам використовує експоненціально ковзні середні, обчислені на основі градієнта, оціненого на поточній міні-серії:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \beta g_1 \quad [22] \quad (3.4)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad [22] \quad (3.5)$$

Де m і v — ковзні середні, g — градієнт на поточній міні-пакеті, а β — нові введені гіперпараметри алгоритму. Вони мають дійсно хороші значення за замовчуванням 0,9 і 0,999 відповідно. Ці значення майже ніхто ніколи не змінює. Вектори ковзних середніх ініціалізуються нулями на першій ітерації [23].

Щоб побачити, як ці значення корелюють з моментом, визначеним як у першому рівнянні, подивимось на очікувані значення наших ковзних середніх. Оскільки m і v є оцінками першого і другого моментів, ми хочемо мати таку властивість:

$$E[m_t] = E[g_t] \quad (3.6)$$

$$E[v_t] = E[g_t^2] \quad (3.7)$$

Зафіксувавши цей шаблон, ми можемо переписати формулу для ковзного середнього:

$$m_t = (1 - \beta_t) \sum_{i=0}^t \beta_t^{t-i} g_i \quad (3.8)$$

З чого випливає наступне:

$$E[g_i](1 - \beta_t^t) + \varepsilon \quad (3.9)$$

Тепер нам потрібно виправити оцінювач, щоб очікуване значення було тим, що ми хочемо. Цей крок зазвичай називають корекцією зсуву. Остаточні формули для нашого оцінювача будуть такими[23]:

$$\widehat{m}_t = \frac{m_t}{1-\beta_1^t} \quad (3.10)$$

$$\widehat{v}_t = \frac{v_t}{1-\beta_2^t} \quad (3.11)$$

Єдине, що залишилося зробити, це використати ці ковзні середні для індивідуального масштабування швидкості навчання для кожного параметра. Спосіб, як це робиться в Adam, дуже простий, щоб виконати оновлення ваги, ми робимо наступне:

$$w_t = w_{t-1} - \eta \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t + \varepsilon}} \quad (3.12)$$

Де w — це ваги моделі, η — розмір кроку, це правило оновлення для Адама.

Одним з головних переваг Adam є те, що він здатний автоматично налаштовувати швидкість навчання на кожному кроці оптимізації. Це означає, що швидкість навчання може змінюватись на основі інформації, яку модель отримує під час навчання, що може забезпечити швидку збіжність моделі.

Adam також має вбудований механізм для регуляризації моделі, що може допомогти зменшити перенавчання та покращити загальну точність моделі. Однак, як і у будь-якого оптимізатора, є певні недоліки, такі як складність підбору гіперпараметрів та можливість зациклення на локальному мінімумі.

3.4.2 SGD

SGD (Stochastic Gradient Descent) - це один з базових оптимізаторів, який використовується в нейронних мережах. Оптимізатор SGD використовує градієнтний спуск для оновлення вагів мережі[24].

Зазвичай SGD використовується для навчання мережі на великих даних, так як він розрахований на обробку окремих зразків даних випадковим чином. Це допомагає запобігти застряганню в локальному мінімумі, оскільки параметри мережі будуть оновлюватися за допомогою випадково обраних зразків.

Хоча SGD є простим та ефективним методом оптимізації, він може страждати від проблеми застряганню в локальному мінімумі. Тому були розроблені різні модифікації SGD, такі як SGD з моментом, Nesterov Accelerated Gradient та інші, які допомагають покращити його ефективність.

3.4.3 RMSprop

Оптимізатор RMSprop (Root Mean Square Propagation) є адаптивним методом градієнтного спуску для навчання нейронних мереж. Він був запропонований в 2012 році Геффом Хінтоном для покращення методу градієнтного спуску.

Оптимізатор RMSprop використовує експоненційно зважене середніх квадратів градієнтів для кожного параметра моделі. Це дозволяє різним параметрам моделі адаптивно налаштовувати швидкість навчання в залежності від історії градієнтів. Зокрема, оптимізатор RMSprop використовує швидкість навчання, яка залежить від середнього квадрата останніх градієнтів для кожного параметра.

Оптимізатор RMSprop допомагає уникнути проблеми "затухання градієнтів", що може статися при використанні методу градієнтного спуску з фіксованою швидкістю навчання. Він також може допомогти прискорити збіжність моделі.

Проте, як і в інших адаптивних методах оптимізації, може виникати проблема "перестрілювання", коли модель перестає генералізувати на нові дані через занадто агресивні налаштування швидкості навчання для кожного параметра. Також, оптимізатор RMSprop може бути вимогливим до великих обсягів пам'яті і обчислень.

3.4.4 Adagrad

Оптимізатор Adagrad (Adaptive Gradient Algorithm) - це адаптивний метод оптимізації градієнтного спуску для нейронних мереж.

На відміну від стандартного градієнтного спуску, який використовує сталу швидкість навчання для всіх параметрів моделі, Adagrad змінює швидкість навчання для кожного параметра окремо, що дозволяє краще керувати процесом навчання. Швидкість навчання для кожного параметра залежить від історії градієнтів, що використовувалися для оновлення цього параметра раніше.

Ідея полягає в тому, щоб зберігати квадрат кожного градієнту та додавати його до суми квадратів градієнтів попередніх ітерацій. Потім швидкість навчання для кожного параметра обчислюється як ділення поелементного добутку градієнта на інвертоване значення квадратного кореня зі суми квадратів градієнтів попередніх ітерацій. Це дозволяє збільшувати швидкість навчання для параметрів, які мають рідкі градієнти, та зменшувати швидкість навчання для параметрів, які мають часті градієнти.

Одним з головних переваг Adagrad є те, що він дозволяє ефективно навчати моделі з великою кількістю параметрів, таких як згорткові нейронні мережі. Однак, недоліком може бути те, що з часом сума квадратів градієнтів може стати дуже

великою, що призводить до зменшення швидкості навчання, що може призвести до повільного навчання моделі.

3.5 Upscaling

Upscaling - це процес збільшення розміру зображення без втрати якості. Це може бути корисним, якщо ми маємо мало розширення зображення, або якщо ми хочемо роздрукувати або відобразити зображення на більшому екрані. Приклад Апскейлінгу зображено на рис. 3.7:

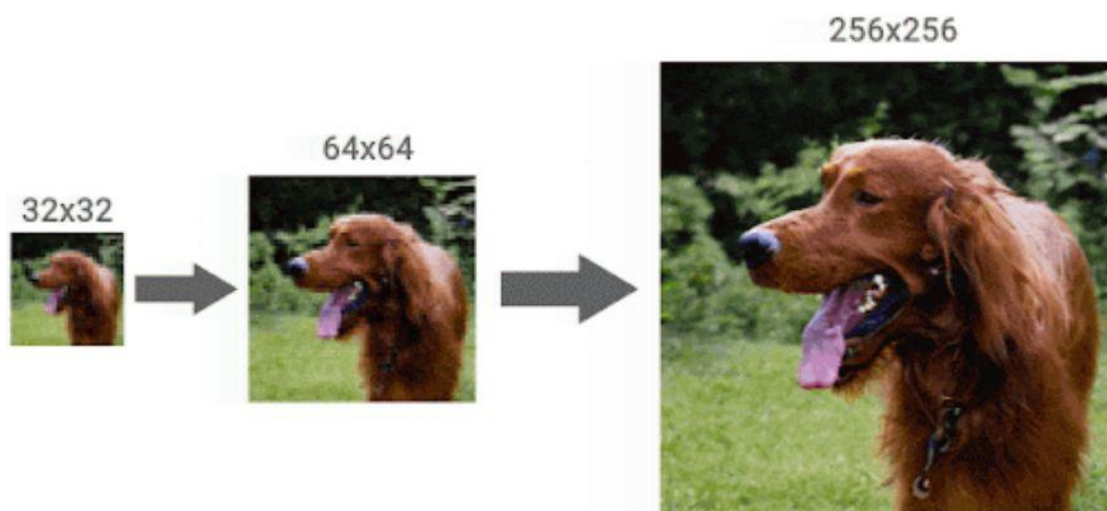


Рисунок 3.7 - Приклад Апскейлінгу [25]

Для upscaling фотографії існує кілька методів. Одним з найбільш поширених є інтерполяція, де нові пікселі заповнюються на основі інформації з навколишніх пікселів. Однак, цей метод може призвести до втрати якості, особливо якщо масштаб збільшення великий.

Інші методи upscaling включають в себе використання нейронних мереж, що можуть навчитися відновлювати втрачені деталі у зображенні при збільшенні розміру. Ці методи називаються super-resolution (від англ. "надзвичайне розширення") і використовуються в багатьох програмах та сервісах для upscaling фотографій. Зі збільшенням розміру зображення надлишки пікселів повинні якось інтерполювати. Основні методи обробки зображень не дають гарних результатів, оскільки вони не враховують контекст масштабування. Глибоке навчання та зокрема генеративні змагальні мережі, приходять на допомогу і дають набагато кращі результати.

Найпопулярніша бібліотека комп'ютерного зору з відкритим кодом, яка містить велику колекцію гарних алгоритмів це OPENCV. Вона містить простий інтерфейс для використання та реалізації Upscaling на основі методів глибокого навчання. Інтерфейс містить попередньо натреновані моделі, які є легкими у використанні. Наразі є багато реалізацій на різних мовах програмування зокрема і на Python[25].

3.5.1 Бікубічна інтерполяція

Бікубічна інтерполяція є одним з методів розмиття зображень під час їх масштабування. Цей метод зазвичай використовується для збільшення розміру цифрового зображення без втрати якості.

В основі цього методу лежить ідея, що кожен новий піксель на збільшеному зображенні має бути отриманий на основі околу змінних пікселів на вихідному зображенні. Для цього використовується функція, яка підібрана для роботи з великими даними - бікубічна функція.

Бікубічна інтерполяція використовує множину з чотирьох змінних пікселів, які оточують новий піксель на збільшеному зображенні, для визначення його значення. Ця функція використовується для підрахунку вагових коефіцієнтів, які використовуються для обчислення нового значення пікселя.

Однією з переваг бікубічної інтерполяції є те, що вона забезпечує високу якість зображення при збільшенні його розміру, зберігаючи деталі зображення та гладкість. Однак, наявність високої якості зображення за рахунок збільшення його розміру може призвести до збільшення розміру файлу, що може бути недоцільно в деяких випадках.

Узагалі, бікубічна інтерполяція є популярним методом збільшення розміру зображень, який використовують у багатьох галузях, включаючи фотографію, медичну техніку, обробку відео та комп'ютерну графіку.

3.5.2 Білінійна інтерполяція

Білінійна інтерполяція є одним із методів збільшення розмірів зображення відносно його вихідного розміру. Вона використовується для створення більш реалістичного і гладкого зображення під час збільшення масштабу зображення.

Основна ідея полягає у тому, що нові пікселі заповнюються на основі значень пікселів, що оточують даний піксель. Кожен новий піксель розміщується на середині між чотирма найближчими пікселями. Для кожного нового пікселя використовуються значення яскравості сусідніх пікселів, що знаходяться на тій же лінії або колонці з новим пікселем.

Білінійна інтерполяція дозволяє покращити якість зображення після збільшення масштабу, проте порівняно з іншими методами, такими як бікубічна інтерполяція або методи машинного навчання, вона може викликати певні артефакти. Однак, вона є досить ефективним і швидким методом для збільшення масштабу зображення з невеликими змінами на оригінальному зображенні.

3.5.3 EDSR

EDSR (Enhanced Deep Super-Resolution) - це метод глибокого навчання, який використовується для удосконалення якості зображень в машинному зорі. Він використовує глибоку згорткову нейронну мережу (CNN) для досягнення високої якості удосконалення зображень[26]. Архітектура EDSR зображена на рис. 3.8:

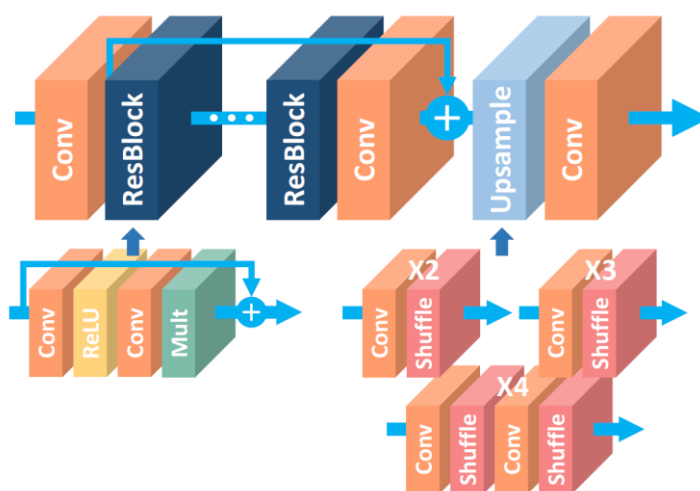


Рисунок 3.8 - Архітектура EDSR [26]

Основним етапом методу EDSR є побудова глибокої згорткової нейронної мережі з багатьма шарами. Кожен шар мережі складається зі згортки та ReLU-функції активації, яка допомагає зменшити кількість нейронів, що передаються на наступний шар, тим самим зменшуючи розмір моделі та підвищуючи швидкість обчислень.

Для удосконалення зображення метод EDSR використовує резидуальний блок, який допомагає підвищити якість удосконалення зображення. У кожному резидуальному блоку використовується згортка для екстракції особливостей та зворотний прохід для реконструкції зображення.

Його дизайн залишкових блоків відрізняється від дизайну ResNet. Шари пакетної нормалізації було видалено разом з остаточною активацією ReLU, як

показано на правій стороні наступного малюнку. На рис.3.9 показано дизайн залишкового блоку в ResNet і в EDSR відповідно

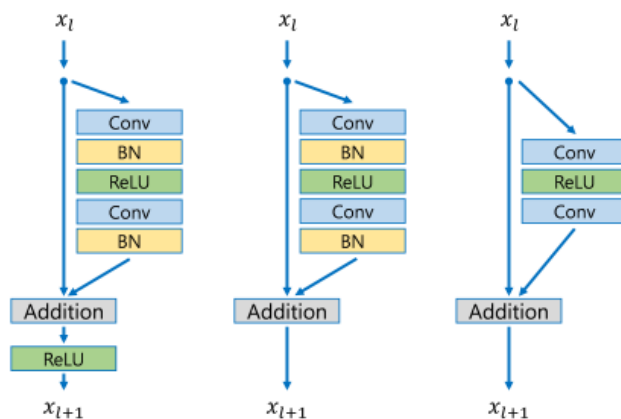


Рисунок 3.9 - Дизайн залишкового блоку в ResNet і в EDSR відповідно [26]

Окрім цього, метод EDSR використовує техніку "upsampling" для підвищення розміру зображення, що збільшує його роздільну здатність. У процесі "upsampling" зображення зменшується та інтерполюється для отримання більшої кількості пікселів, після чого застосовуються згортки для удосконалення деталей.

Основна перевага методу EDSR полягає в тому, що він дозволяє досягати високої якості удосконалення зображень, що забезпечує високу точність відновлення деталей та збереження кольорів та текстур на зображенні.

3.5.4 ESPCN

Метод ESPCN (Enhanced Super-Resolution Generative Adversarial Networks) є методом машинного навчання для збільшення роздільної здатності зображень, також відомий як метод super-resolution. Він був запропонований у 2016 році та використовується для генерації високоякісних зображень, що мають вищу роздільну здатність, ніж вихідні зображення[27].

Метод ESPCN використовує згорткову нейронну мережу (Convolutional Neural Network - CNN) для збільшення роздільної здатності зображень. Він призначений для збільшення роздільної здатності у 2, 3 або більше разів в порівнянні з вихідними зображеннями. Для досягнення цієї мети, метод ESPCN використовує метод upsampling, який збільшує розмір зображення, а потім використовує згорткову нейронну мережу для покращення якості зображення.

Ефективна субпіксельна згорткова нейронна мережа. Замість того, щоб створювати надвисоку роздільну здатність після масштабування низької за допомогою бікубічного фільтра, розробники вирішили отримувати “feature maps” об'єктів у малій кількості та використовувати складні фільтри роздільної здатності масштабування, щоб отримати результат.

Масштабування шарів розвертаються тільки в кінці мережі, що гарантує, що складні операції, що знаходяться в моделях, виконуються в менших вимірах, що робить їх швидшими, особливо в порівнянні з іншими методами. Алгоритм роботи ESPCN показано на рис. 3.10:

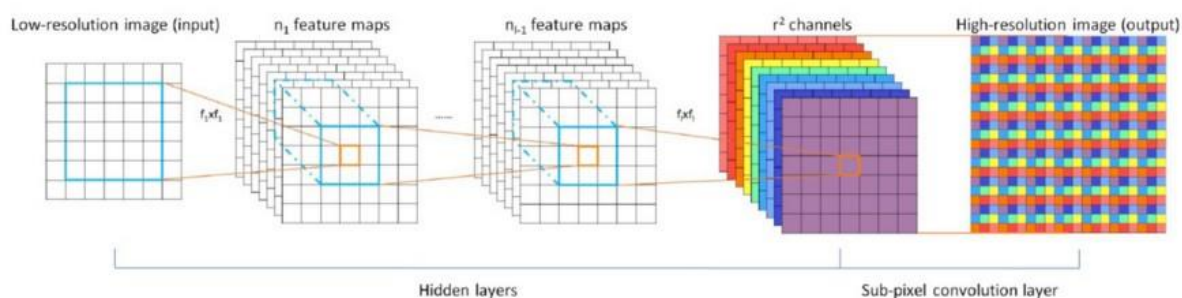


Рисунок 3.10 - Алгоритм роботи ESPCN [27]

Оригінальна структура ESPCN була натхненна SRCNN. Замість звичайних шарів згорткової мережі використовуються субпіксельні, що діють як шари деконволюції. Субпіксельні шари все ж використовуються в останньому шарі для створення “feature maps” високої роздільної здатності. Разом з цим до них була обрана функція активації Tanh, яка працює набагато краще, ніж стандартна функція ReLu.

Метод ESPCN має кілька переваг, серед яких є швидкість обробки та низький рівень витрат ресурсів. Крім того, метод ESPCN є ефективним для збільшення роздільної здатності зображень з високою якістю, такої як медичні зображення та зображення з високою деталізацією.

3.5.5 Вибір оптимального методу Upscaling

Для вибору оптимального методу Upscaling треба обрати фотографію та обрати метрику за котрими буде порівнюватись результат, це SSIM(Індекс структурної подібності) і PSNR(Пікове ставлення сигналу до шуму), а також час, за котрий метод буде опрацьовувати фотографію, бо таких фотографій більш ніж 1000, які використовуються при навчанні.

В якості прикладу було обрано дві фотографії T-72 у розмірі 138 на 139 пікселей. Фото прикладу оригінального розміру фото зображено на рис. 3.11 та 3.12:

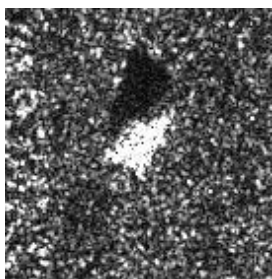


Рисунок 3.11 - Фото прикладу оригінального розміру фото

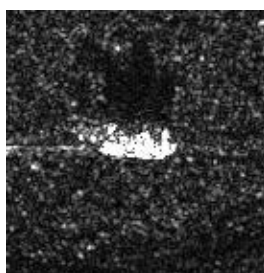


Рисунок 3.12 - Фото прикладу оригінального розміру фото

Після апскейлінгу розмір фото повинен збільшитися у 3 рази. Перше фото, рис. 3.13, буде оброблена методом ESPCN, а друге, рис. 3.14, EDSR:

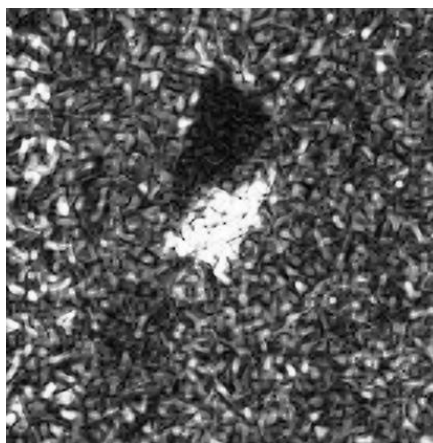


Рисунок 3.13 – Результат моді ESPCN

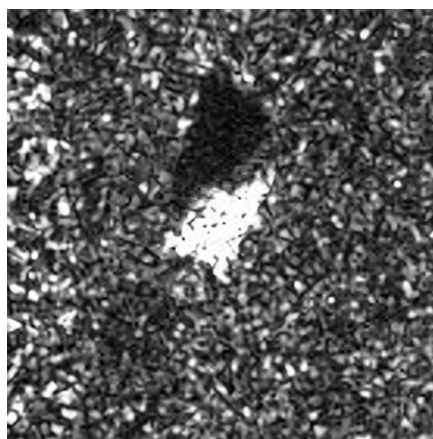


Рисунок 3.14 - Результат моді EDSR

Приклад роботи ESPCN та EDSR на першій фото відповідно на рисунках 3.15 та 3.16.

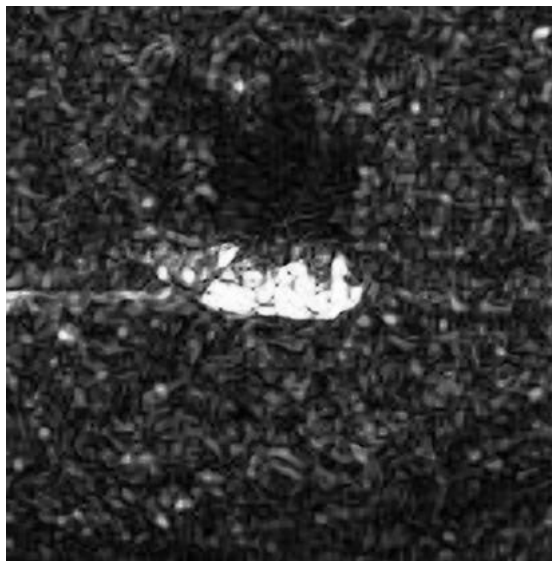


Рисунок 3.15 - Результат моді ESPCN

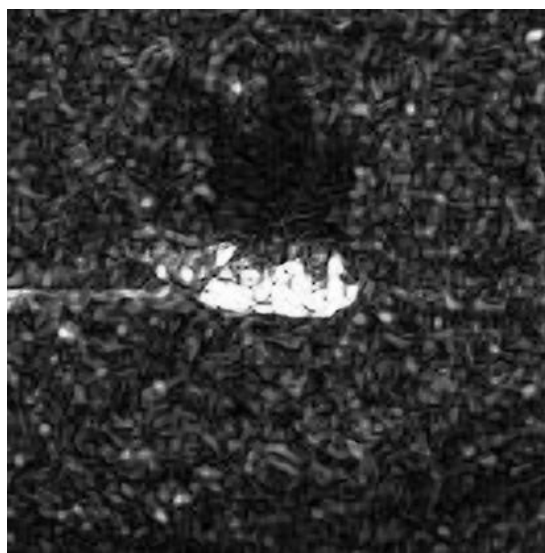


Рисунок 3.16 - Результат моді EDSR

Навіть якщо сильно придивитись, то різницю побачити людським оком майже неможливо, тоді подивимось на схожість нових фотографій з початковими.

Таблиця 3.1 - Середні значення PSNR

Назва моделі	3x
EDSR	26.554
ESPCN	25.854

Таблиця 3.2 - Середні значення SSIM

Назва моделі	3x
EDSR	0,831
ESPCN	0,787

Таблиця 3.3 - Середній час у секундах

Назва моделі	3x
EDSR	12,018
ESPCN	0,042

Звідси можна зробити висновок що, метод EDSR дає найкращі результати, але час його роботи є більшим порівняно з іншим, що значить для застосування підходить метод ESPCN, так як різниця у якості є невелика.

3.6 Висновки до розділу

У цьому розділі було розглянуто математичний апарат системи, метрики за якими буде проходити оцінка моделей та обрано функцію оптимізації для нейромережі, тому що це має дуже важливий вплив на систему в загалі. Для фотографії які мають меншу роздільну здатність було застосовано метод upscaling, які були описані та порівняні між собою. Надалі було обрано метод ESPCN.

У таблиці 3.4 було описано важливі характеристики кожного із методів оптимізації та записано у таблицю.

Таблиця 3.4 – Список властивості методів оптимізації

Метод оптимізації	Характеристики
SGD	Є простим та ефективним методом оптимізації, хоча він може страждати від проблеми меленої збіжності або застрягання в локальному мінімумі.
AdaGrad	Одним з головних переваг Adagrad є те, що він дозволяє ефективно навчати моделі з великою кількістю параметрів, таких як згорткові нейронні мережі. Однак, недоліком може бути те, що з часом сума квадратів градієнтів може стати дуже великою, що призводить до зменшення швидкості навчання, що може призвести до повільного навчання моделі.
RMSProp	Може уникнути проблеми "затухання градієнтів", також може допомогти прискорити збіжність моделі. Проте, як і в інших адаптивних методах оптимізації, може виникати проблема "перестрілювання", коли модель перестає генералізувати на нові дані через занадто агресивні налаштування швидкості навчання для кожного параметра

Метод оптимізації	Характеристики
Adam	Він здатний автоматично налаштовувати швидкість навчання на кожному кроці оптимізації та має вбудований механізм для регуляризації моделі, що може допомогти зменшити перенавчання та покращити загальну точність моделі

Також було з'ясовано, що метод оптимізації Adam – є одним із найвдалішим вибором для згорткової мережі, який буде використаний у подальших моделях.

4 КОМПОНЕНТНА МОДЕЛЬ СИСТЕМИ

4.1 Складові системи

Компонентна модель системи розпізнавання військової техніки складається з таких частин (рис. 4.1): вебзастосунок, сховище даних, алгоритм розпізнавання військової техніки та сервер, причому всі частини можуть бути інкапсульовані на віддалених хмарних сервісах наприклад Azure або AWS.

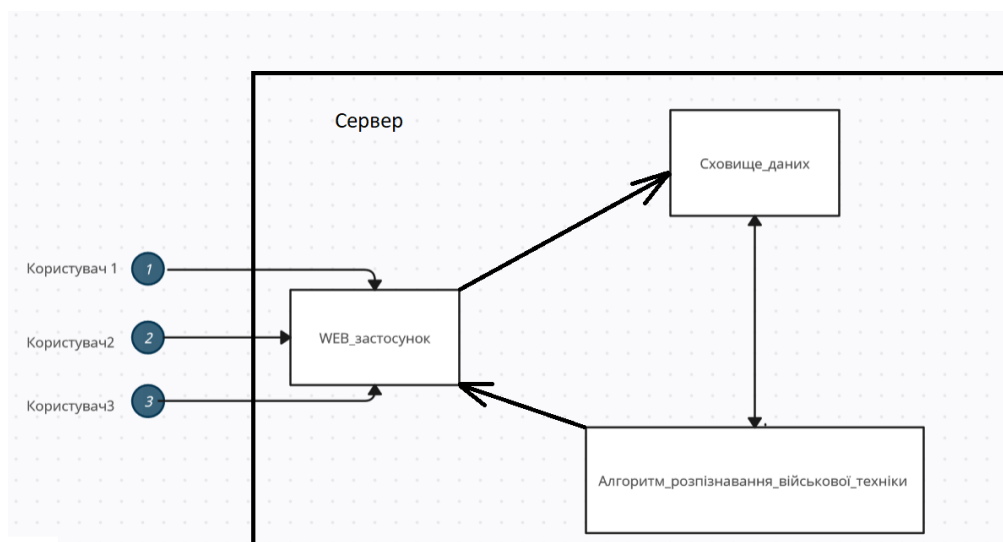


Рисунок 4.1 – Модель системи

У цій діаграмі Web-додаток взаємодіє зі Сховищем даних та Алгоритмом розпізнавання військової техніки через Сервер. Web-додаток використовує дані, які зберігаються в Сховищі даних, а також передає дані на Алгоритм розпізнавання військової техніки для їх обробки. Алгоритм розпізнавання військової техніки отримує дані з Сервера та повертає результати своєї роботи назад на Сервер.

4.2 Опис функціональних компонентів

Архітектура компонентної моделі:

- «Web застосунок» – призначений як основний інтерфейс для користувачів. За його допомогою можливо завантажити дані на сервер та отримати результат класифікації.
- «Сховище даних» – призначений для збереження фотографії та результатів їх обробки;
- «Алгоритм нормалізації» – використовується для обробки зображень та абскейлінгу, якщо це потрібно;
- «Алгоритм розпізнавання військової техніки» – призначений для розпізнавання військової техніки за фотографією;

5 ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

5.1. Архітектура розроблених програмних засобів

Так як система працює з вхідними зображеннями, то в першу чергу потрібно їх обробити, також якщо фотографія має іншу роздільну здатність ніж 368×368 , то потрібно змінити її розмір за допомогою моделі апскейлінгу, або навпаки зменшити, якщо фото більше потрібного розміру.

Розроблювальна система представлена як веб-додаток. Багато в чому його успіх залежить від того наскільки продуманим є архітектура - основа будь-якого веб-рішення. Складання частинок відповідно до мети вебрішення, приведе до того, що продукт буде працювати належним чином, та легкий в адаптуванні.

Під час розробки були використані декілька фреймворків Python, наприклад, для реалізації моделі нейронної мережі використовується TensorFlow, це бібліотека, яка забезпечує велику кількість алгоритмів машинного навчання та баз даних.

PostgreSQL, що забезпечує локальне зберігання даних та інші допоміжні інструменти.

Для подальшого дослідження було обрана три архітектури нейронної мережі:

- 1) Много шаровий перцептрон, натренований на датасеті Mstar
- 2) Згорткова нейронна мережа архітектури CNN натренована на Mstar.
- 3) Комбінована згорткова нейронна мережа архітектури CNN з додатковими шарами для уніфікації моделі, натренована на датасеті Mstar.

5.2 Навчання моделі

Дані розбиваються на тренувальні та тестові, у пропорції 80\20.

Навчання проходить у 20 епох (рис. 5.1):

```

Epoch 1/20
296/296 [=====] - 62s 95ms/step - loss: 1.1277 - categorical_accuracy: 0.5828 - MCC: 0.5224 - F2: 0.5809 - auc: 0.9202 - prc: 0.6689 - val_loss: 5.3709 - val_categorical_accuracy: 0.1041 - val_MCC: 0.0
Epoch 2/20
296/296 [=====] - 25s 84ms/step - loss: 0.5274 - categorical_accuracy: 0.8143 - MCC: 0.7880 - F2: 0.8144 - auc: 0.9809 - prc: 0.9003 - val_loss: 5.8609 - val_categorical_accuracy: 0.2218 - val_MCC: 0.2
Epoch 3/20
296/296 [=====] - 25s 83ms/step - loss: 0.3487 - categorical_accuracy: 0.8797 - MCC: 0.8027 - F2: 0.8798 - auc: 0.9908 - prc: 0.9498 - val_loss: 4.3817 - val_categorical_accuracy: 0.2976 - val_MCC: 0.2
Epoch 4/20
296/296 [=====] - 25s 83ms/step - loss: 0.2323 - categorical_accuracy: 0.9194 - MCC: 0.9080 - F2: 0.9195 - auc: 0.9957 - prc: 0.9751 - val_loss: 15.1381 - val_categorical_accuracy: 0.2244 - val_MCC: 0.0
Epoch 5/20
296/296 [=====] - 25s 84ms/step - loss: 0.2251 - categorical_accuracy: 0.9244 - MCC: 0.9137 - F2: 0.9244 - auc: 0.9953 - prc: 0.9758 - val_loss: 3.8431 - val_categorical_accuracy: 0.3337 - val_MCC: 0.3
Epoch 6/20
296/296 [=====] - 26s 88ms/step - loss: 0.1787 - categorical_accuracy: 0.9388 - MCC: 0.9381 - F2: 0.9389 - auc: 0.9972 - prc: 0.9843 - val_loss: 1.6062 - val_categorical_accuracy: 0.6198 - val_MCC: 0.6
Epoch 7/20
296/296 [=====] - 25s 83ms/step - loss: 0.1717 - categorical_accuracy: 0.9398 - MCC: 0.9312 - F2: 0.9398 - auc: 0.9970 - prc: 0.9850 - val_loss: 8.6577 - val_categorical_accuracy: 0.2798 - val_MCC: 0.2
Epoch 8/20
296/296 [=====] - 25s 84ms/step - loss: 0.1442 - categorical_accuracy: 0.9496 - MCC: 0.9425 - F2: 0.9496 - auc: 0.9980 - prc: 0.9891 - val_loss: 2.5324 - val_categorical_accuracy: 0.6203 - val_MCC: 0.6
Epoch 9/20
296/296 [=====] - ETA: 0s - loss: 0.1429 - categorical_accuracy: 0.9515 - MCC: 0.9446 - F2: 0.9515 - auc: 0.9977 - prc: 0.9892
Epoch 9: ReduceLROnPlateau reducing learning rate to 0.0008080808237487257.
296/296 [=====] - 24s 83ms/step - loss: 0.1429 - categorical_accuracy: 0.9515 - MCC: 0.9446 - F2: 0.9515 - auc: 0.9977 - prc: 0.9892 - val_loss: 11.8242 - val_categorical_accuracy: 0.2871 - val_MCC: 0.0
Epoch 10/20
296/296 [=====] - 24s 82ms/step - loss: 0.0778 - categorical_accuracy: 0.9736 - MCC: 0.9699 - F2: 0.9736 - auc: 0.9994 - prc: 0.9966 - val_loss: 0.1663 - val_categorical_accuracy: 0.9399 - val_MCC: 0.9
Epoch 11/20

```

Рисунок 5.1 – Навчання моделі

Результати метрик після навчання третьої моделі (рис. 5.2):

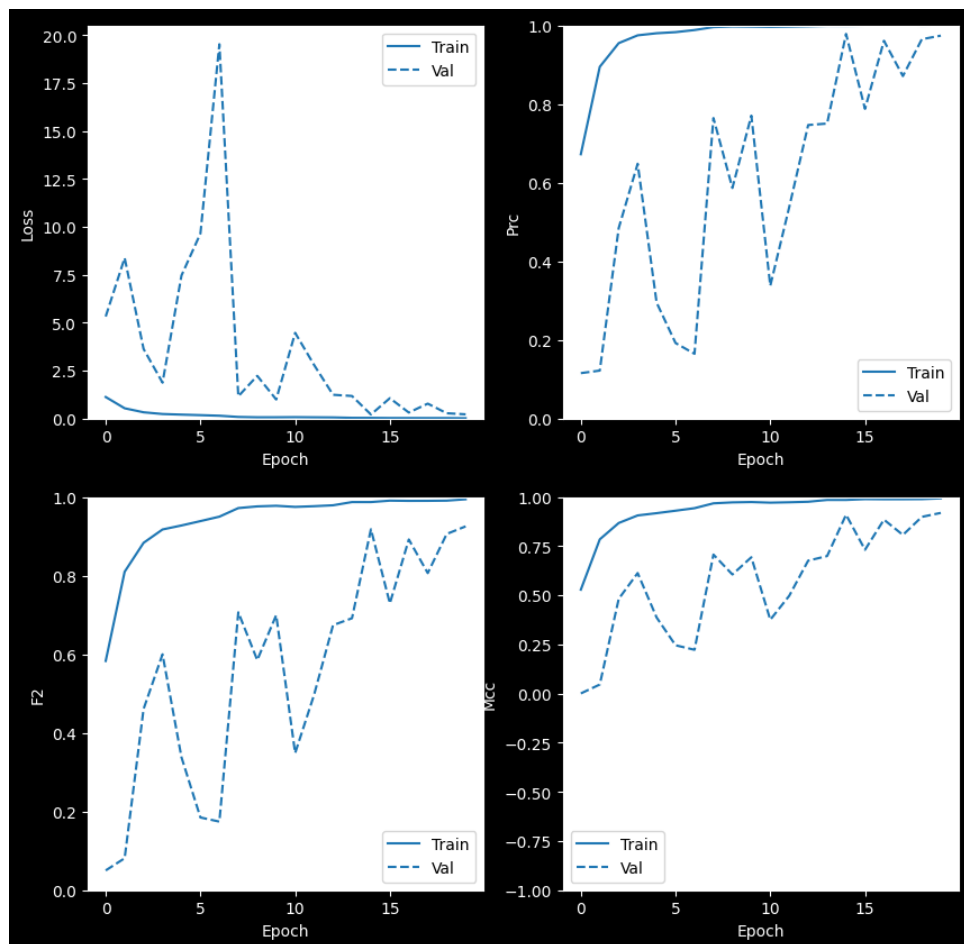


Рисунок 5.2- Графіки навчання моделі

5.3 Результат роботи

Інтерфейс програми був розроблений у вигляді web-додатку. Back-end частина реалізована на Flask. Також майбутню систему без проблем можливо масштабувати за допомогою конвенції Docker. Завдяки цьому програма буде працювати в окремому віртуальному оточенні, що дозволить встановити спеціалізовані бібліотеки потрібних версії для безперервної роботи без помилок.

При запуску програми вона заново створює нейронну мережу та завантажує

ваги попередньо навченої моделі які зберігаються в окремій папці. Після вдалої ініціалізації запускається сам сервер на локальній адресі : <http://127.0.0.1:5000>.

По проходженні за посиланням відкривається сайт для завантаження окремих фотографій, які надалі пройдуть через мережу.

Так виглядає початкова сторінка (рис.5.3):

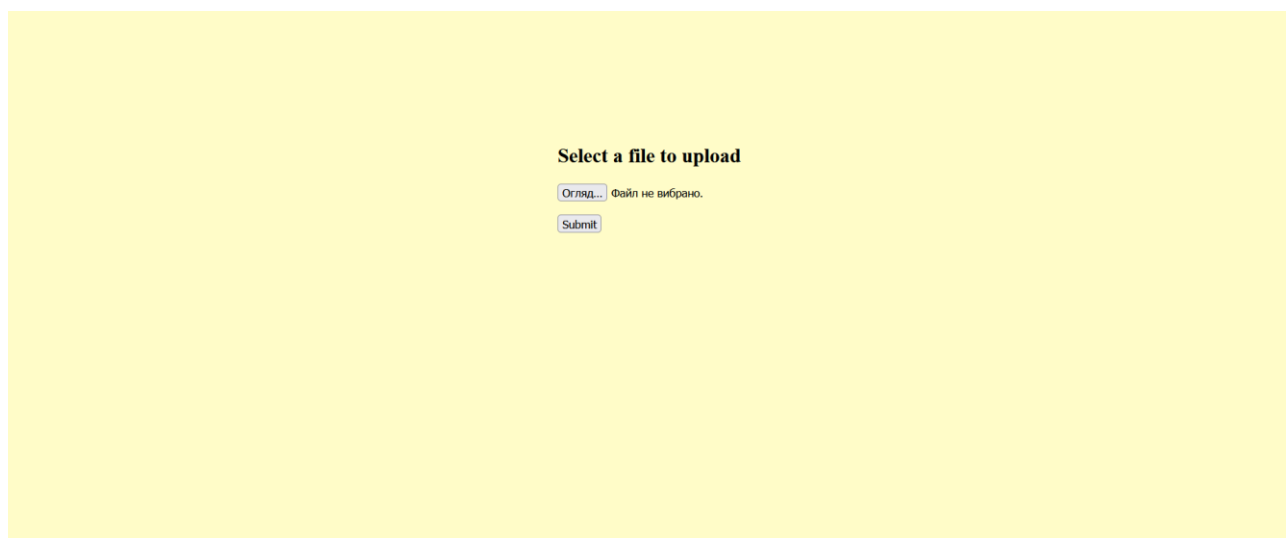


Рисунок 5.3 – Початковий екран програми

Потім користувачу дозволяється завантажити фотографію з військовою технікою. Після чого потрібно підтвердити завантаження фото. Наступним кроком

сервер зберігає нову фотографію у себе і пропускає її через модель, яка в свою чергу повертає найбільш (най)імовірніший результат техніки.

Результат роботи (рис. 5.4):

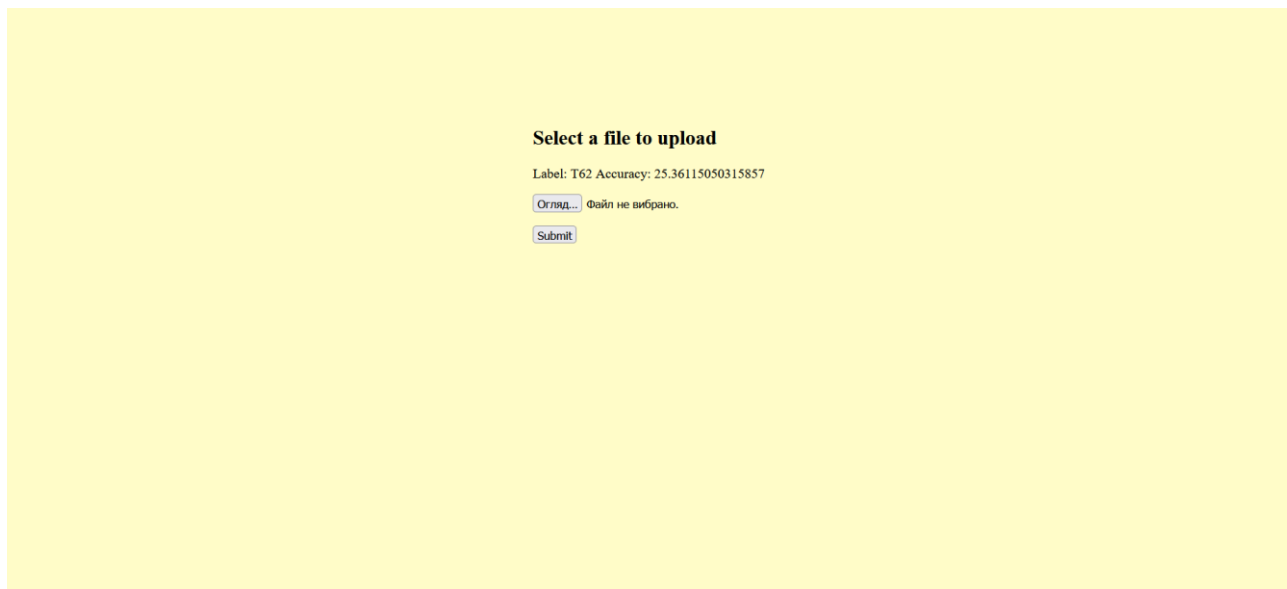


Рисунок 5.4 – Результат роботи алгоритму

Після чого можливо одразу вибрати нове фото для розпізнавання військової техніки (рис. 5.5):

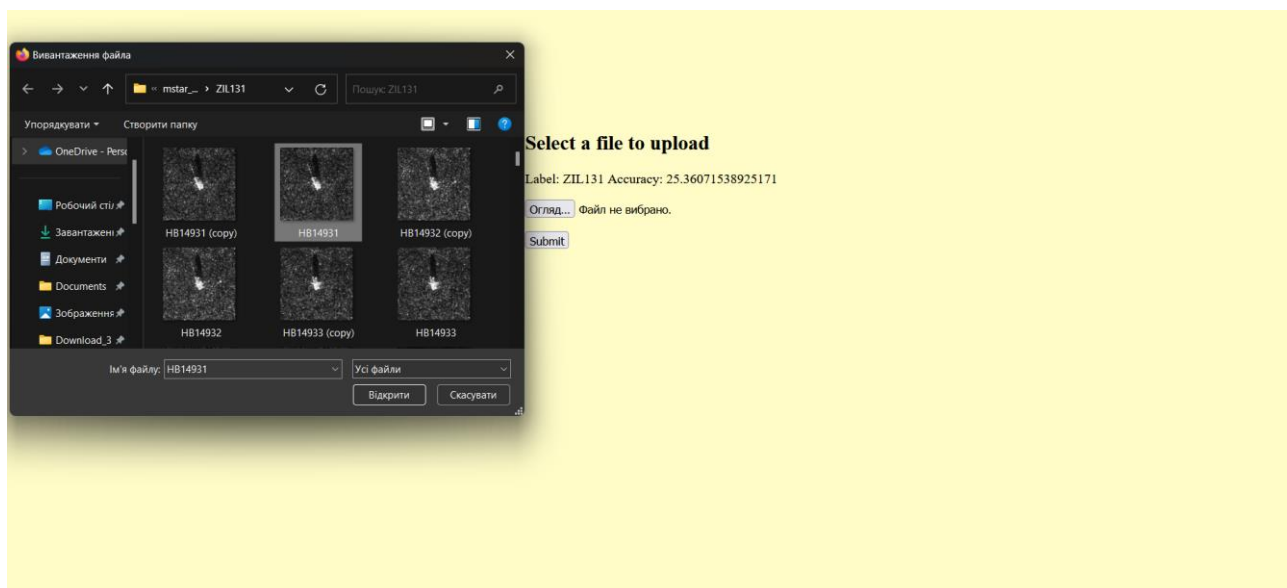


Рисунок 5.5 – Завантажування фото до системи

Якщо подивитися на точність моделі на тестових даних, то найкраще описує результат роботи матриця помилок, але для більш зрозумілої та інтерпретованої оцінки можна отримати результати метрик.

Матриця помилок на тестових даних має такі значення (рис. 5.6):

	Z51	BRDM_2	BTR_60	D7	SUICY	T62	T72	ZIL131	ZSU_23_4
Z51	0	0	0	0	0	0	0	0	0
BRDM_2	0	0	0	0	0	0	0	0	0
BTR_60	0	0	0	0	0	0	0	0	0
D7	0	0	0	0	0	0	0	0	0
SUICY	0	0	0	0	34	0	7	0	2
T62	0	0	0	0	0	0	39	0	0
T72	0	0	0	0	0	0	79	0	0
ZIL131	0	0	0	0	0	0	0	55	0
ZSU_23_4	0	0	0	0	0	0	0	0	36

Рисунок 5.6 – Матриця помилок першої моделі

Можна побачити що перша модель зовсім не відрізняє об'єктів на фотографіях. Метрики відповідно показують нульові результати. Ця мережа працює некоректно, через те що її архітектура проста, і її кількість внутрішніх параметрів не може вдало справлятися з цим завданням.

Наступна модель спрацювала набагато краще (рис. 5.7), але є невеличкі помилки з окремими класами, але вже на цьому рівні можна побачити що мережа може відрізняти класи техніки.

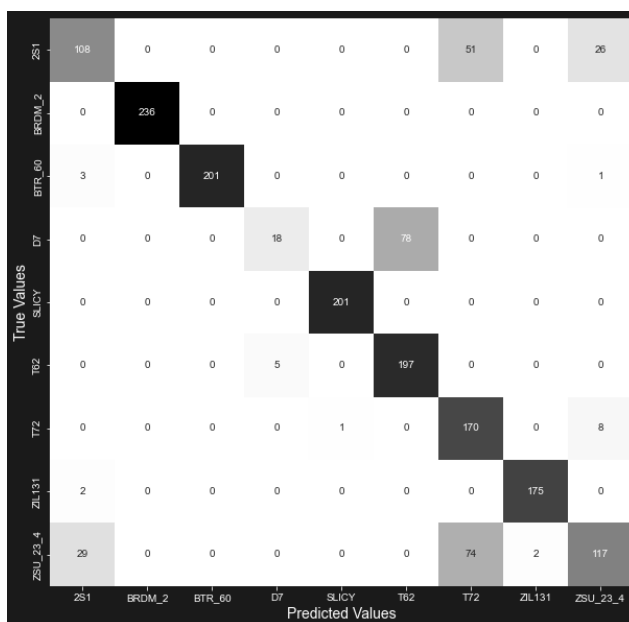


Рисунок 5.7 - Матриця помилок другої моделі

Для порівняння результатів роботи моделей будемо опиратися на результати метрик. Для другої мережі вони виглядають так:

loss: 0.362,

categorical_accuracy: 0.822,

MCC: 0.802,

F2: 0.813,

Auc: 0.991

Остання матриця помилок має наступний вигляд (рис. 5.8). Дана модель є найскладнішою, але результат роботи вийшов точніший. Завдяки додатковим шарам регуляризації і більшій кількості прихованих шарів, ця модель працює дуже вдало.

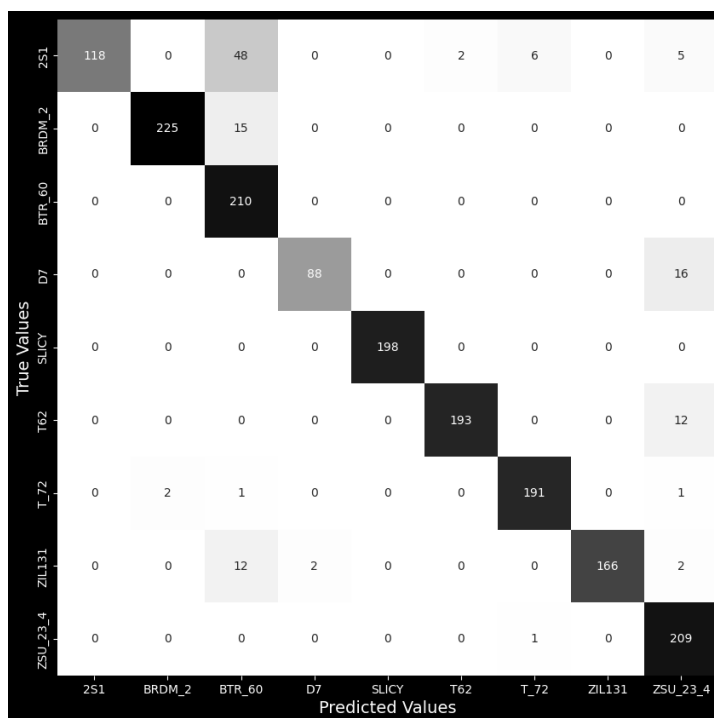


Рисунок 5.8 - Матриця помилок третьої моделі

а остаточні результати метрик були наступними:

loss: 0.119,

Categorical_accuracy: 0.961,

MCC: 0.956,

F2: 0.961,

Auc: 0.997

5.4 Висновки до розділу

В результаті проведеного дослідження встановлено, що для обраного датасету найкращим методом для розпізнавання військової техніки є нейрона мережа на основі CNN з додатковими шарами регуляризації. Головною перевагою такою системи є те що вона доволі точно відрізняє між собою техніку Т-62 і Т-72, які дуже схожі між

собою. Також вона вдало працює з макетами які є в оригінальному датасеті. А мережа персептрон зовсім не впоралась с завданням, бо її архітектура не дозволяти виконувати важкі задачі. Звичайна CNN працює добре, і може також відрізнати схожі танки між собою, але працює менш точно ніж остання модель.

Результуюча точність алгоритму складає за Categorical_accuracy: 0.961 та за MCC: 0.956.

ВИСНОВКИ

У дисертаційній роботі було отримано такі нові теоретичні та практичні результати:

1. Встановлено, що створення такої системи може підвищити якість розвідки, яка в свою чергу підвищить оборонну спроможність нашої країни. На початку роботи проведено огляд існуючих рішень, розглянуто, які інструменти можуть допомогти для збору нових даних. Розібрані рішення для розпізнавання об'єктів та їх недоліки для даної системи. Також було описані навчальні дані, та обґрунтовано їх важливість на сьогодні.

2. Реалізовано математичне та програмне забезпечення системи розпізнавання військової техніки за допомогою супутникових зображень. В системі застосовано низку методів машинного навчання, а саме згорткових нейронних мереж серед яких визначається найкращий метод для заданого набору вхідних даних. В процесі роботи було описано декілька методів попередньої обробки зображень, таких як Апскейлінг. Було розібрано декілька моделей для підвищення роздільної якості фотографій та обрано оптимальний з них. Також була описана компонента модель системи та її складові, включаючи Веб додаток, та відповідно сам алгоритм нейронної мережі.

3. В результаті проведеного дослідження встановлено, що для обраного датасету найкращим методом для розпізнавання військової техніки є НМ на основі CNN з додатковими шарами регуляризацій. Головною перевагою такою системи є те, що вона доволі точно відрізняє між собою техніку Т-62 і Т-72, які дуже схожі між собою. Також вона вдало працює з макетами, які є в оригінальному датасеті. Мережа персептрон зовсім не впоралась з завданням, бо її архітектура не дозволяє виконувати важкі задачі. Звичайна CNN працює добре, і може також відрізняти схожі танки між собою, але працює менш точно ніж остання модель. Результуюча точність алгоритму складає за Categorical_accuracy: 0.961 та за MCC: 0.956.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Golkar A., Cataldo G., Osipova K. Small satellite synthetic aperture radar (SAR) design: A trade space exploration model // *Acta Astronautica*. – 2021. – Vol. 187. – P. 458-474.
2. Ustin S. L., Middleton E. M. Current and near-term advances in Earth observation for ecological applications // *Ecological Processes*. — 2021. — № 1 (10). — С. 1–37.
3. ENVI [Електронний ресурс]: версія 5.6.3 / L3Harris Geospatial. – Режим доступу: <https://www.l3harrisgeospatial.com/Software-Technology/ENVI>.
4. Shafizadeh-Moghadam H., Khazaei M., Alavipanah S.K., Weng Q. Google Earth Engine for large-scale land use and land cover mapping: an object-based classification approach using spectral, textural and topographical factors // *GIScience & Remote Sensing*. – 2021. – Vol. 58. – No. 6. – P. 914-928.
5. Eineder M., Moreira A., Roth A. Ten Years of TerraSAR-X—Scientific Results // *Remote Sensing*. – 2019. – Vol. 11. – No. 3. – P. 364.
6. WorldView-3 [Електронний ресурс] // *Apogeo Spatial*. — 2014. — № 4 (29). — С. 14–17.
7. Li L., Han L., Ding M., Cao H., Hu H. A deep learning semantic template matching framework for remote sensing image registration // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2021. – Vol. 181. – P. 205-217.
8. OpenCV 4.x [Електронний ресурс] // *OpenCV documentation*. — 2021. — Режим доступу: <https://docs.opencv.org/4.x/>.
9. Shetty A.B., Deeksha B., JeevanRebeiro R., Ramyashree H. Facial recognition using Haar cascade and LBP classifiers // *Global Transitions Proceedings*. – 2021. – Vol. 2. – No. 2. – P. 330-335.
10. Kusumaa G.P., Wigatia E.K., Chandraa E. A Review of Recent Advancements in Appearance-based Object Recognition // *Procedia Computer Science*. – 2019. – Vol. 157. – P. 613-620.

11. Mittal H., Pandey A.C., Saraswat M., Kumar S., Pal R., Modwel G. A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets // *Multimedia Tools and Applications*. – 2022. – Vol. 81. – P. 35001-35026.
12. Strack R. Deep learning in imaging // *Nature Methods*. – 2019. – Vol. 16. – P. 17.
13. Krasnoff B. How to start experimenting with Google Lens // *The Verge*. – 2022.
14. Kumar A., Singh A., Singh R., Singh S. A comparative study of IBM Watson visual recognition and Microsoft Azure computer vision // *International Journal of Advanced Computer Science and Applications*. – 2020. – Vol. 11. – No. 4. – P. 1-7.
15. Defense News security global military army equipment technology industry - Army Recognition [Электронный ресурс] // *Army Recognition*. — 2023.
16. Wong S. High Fidelity Synthetic SAR Image Generation of a Canonical Target / S. Wong // *IEEE Transactions on Geoscience and Remote Sensing*. – 2020. – Vol. 58, No. 9. – P. 6475–6487. – DOI: 10.1109/TGRS.2020.2976698.
17. Wang, Jiayun, et al. "Orthogonal convolutional neural networks." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.
18. Galvez, Reagan L., et al. "Object detection using convolutional neural networks." *TENCON 2018-2018 IEEE Region 10 Conference*. IEEE, 2018.
19. Srivastava N. et al. Dropout: a simple way to prevent neural networks from overfitting // *The journal of machine learning research*. – 2014. – T. 15. – №. 1. – C. 1929-1958.
20. Haykin, S. (2009). *Neural networks and learning machines* (Vol. 3). Pearson.
21. Mohideen D.F.M., Raj J.S.S., Raj R.S.P., Raj P.R. Regression Imputation and Optimized Gaussian Naïve Bayes Algorithm for an Enhanced Diabetes Mellitus Prediction Model // *Brazilian Archives of Biology and Technology*. – 2021. – Vol. 64.
22. Ruder S. An overview of gradient descent optimization algorithms // *arXiv preprint arXiv:1609.04747*. – 2016.

23. Bera, Somenath, and Vimal K. Shrivastava. "Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification." *International Journal of Remote Sensing* 41.7 (2020): 2664-2683.

24. scikit-learn.org [електронний ресурс]: офіційний сайт scikit-learn. Режим доступу: https://scikit-learn.org/stable/modules/cross_validation.html

25. Третиник В. В., Давиденко М.А., Попереднє покращення якості зображення для підвищення точності розпізнавання об'єктів // Прикладна математика та комп'ютинг. ПМК, 2022 : п'ятнадцята наук. конф. магістрантів та аспірантів, Київ, 16—18 лист. 2022 р. : зб. тез доп. / [редкол.: Дичка І. А. та ін.]. — К. : Просвіта, 2022. — с.169-173 ISBN 978-617-7010-23-3.

26. Lim B. Enhanced Deep Residual Networks for Single Image Super-Resolution [Електронний ресурс] / Lim B., Son S., Kim H., Nah S., Lee K.M. // arXiv:1707.02921 [cs.CV]. — Режим доступу: <https://arxiv.org/abs/1707.02921>.

27. Wang X. Enhanced Super-Resolution Generative Adversarial Networks [Електронний ресурс] / Wang X., Yu K., Wu S., Gu J., Liu Y., Dong C., Loy C.C., Qiao Y., Tang X. // arXiv:1809.00219 [cs.CV]. — Режим доступу: <https://arxiv.org/abs/1809.00219>.


```

        validation_split=.2,
        seed=10,
        label_mode='categorical',
        color_mode='grayscale',
        batch_size=batch_size)

val_batches = tf.data.experimental.cardinality(val_ds)
test_ds = val_ds.take(val_batches // 5)
val_ds = val_ds.skip(val_batches // 5)

print('Number of validation batches: %d' % tf.data.experimental.cardinality(val_ds))
print('Number of test batches: %d' % tf.data.experimental.cardinality(test_ds))

AUTOTUNE = tf.data.AUTOTUNE

train_ds = train_ds.cache().shuffle(1000).prefetch(buffer_size=AUTOTUNE)
val_ds = val_ds.cache().prefetch(buffer_size=AUTOTUNE)
test_ds = test_ds.cache().prefetch(buffer_size=AUTOTUNE)

num_classes=9

metrics = [
    keras.metrics.CategoricalAccuracy(name='categorical_accuracy'),
    tfa.metrics.MatthewsCorrelationCoefficient(num_classes=9, name='MCC'),
    tfa.metrics.FBetaScore(num_classes=9, average='weighted', beta=2.0, name='F2'),
    keras.metrics.AUC(name='auc'),
    keras.metrics.AUC(name='prc', curve='PR'),
]

class_names = ['2S1', 'BRDM_2', 'BTR_60', 'D7', 'SLICY', 'T62', 'T72', 'ZIL131', 'ZSU_23_4']

resize_and_rescale = tf.keras.Sequential([
    keras.layers.Resizing(128,128),
    keras.layers.Rescaling(1./255)])
data_augmentation = keras.Sequential(
    [
        keras.layers.RandomFlip("horizontal_and_vertical"),
        keras.layers.RandomRotation(0.2),
        keras.layers.RandomZoom(0.1),
    ]
)

def make_model(input_shape, num_classes):
    inputs = keras.Input(shape=input_shape)

    # Image augmentation block
    x = resize_and_rescale(inputs)
    x = data_augmentation(x)

```

```

# Entry block
x = keras.layers.Conv2D(32, 3, strides=2, padding="same")(x)
x = keras.layers.BatchNormalization()(x)
x = keras.layers.Activation("relu")(x)

x = keras.layers.Conv2D(64, 3, padding="same")(x)
x = keras.layers.BatchNormalization()(x)
x = keras.layers.Activation("relu")(x)

previous_block_activation = x # Set aside residual

for size in [128, 256, 512, 728]:
    x = keras.layers.Activation("relu")(x)
    x = keras.layers.SeparableConv2D(size, 3, padding="same")(x)
    x = keras.layers.BatchNormalization()(x)

    x = keras.layers.Activation("relu")(x)
    x = keras.layers.SeparableConv2D(size, 3, padding="same")(x)
    x = keras.layers.BatchNormalization()(x)

    x = keras.layers.MaxPooling2D(3, strides=2, padding="same")(x)

# Project residual
residual = keras.layers.Conv2D(size, 1, strides=2, padding="same")(
    previous_block_activation
)
x = keras.layers.add([x, residual]) # Add back residual
previous_block_activation = x # Set aside next residual

x = keras.layers.SeparableConv2D(1024, 3, padding="same")(x)
x = keras.layers.BatchNormalization()(x)
x = keras.layers.Activation("relu")(x)

x = keras.layers.GlobalAveragePooling2D()(x)
if num_classes == 2:
    activation = "sigmoid"
    units = 1
else:
    activation = "softmax"
    units = num_classes

x = keras.layers.Dropout(0.5)(x)
outputs = keras.layers.Dense(units, activation=activation)(x)
return keras.Model(inputs, outputs)

xception = make_model(input_shape=image_size + (1,), num_classes=9)

callbacks=[
    keras.callbacks.EarlyStopping(patience=10, verbose=1, monitor='val_F2', mode='max', restore_best_weights=True),
    keras.callbacks.ReduceLROnPlateau(factor=.5, patience=3, verbose=1),

```

```

    ]
xception.compile(
    optimizer=tf.keras.optimizers.Adam(1e-3),
    loss="categorical_crossentropy",
    metrics=metrics,
)

xception.load_weights('Weights25')

def getPrediction(filename):

    image = keras.utils.load_img('uploads/'+filename, color_mode='grayscale', target_size=(138,139))
    image_array = keras.utils.img_to_array(image)
    image_array = tf.expand_dims(image_array, 0)

    predictions = xception.predict(image_array)
    score = tf.nn.softmax(predictions[0])

    return class_names[np.argmax(score)], 100 * np.max(score)

```

ДОДАТОК Б ІЛЮСТРАТИВНИЙ МАТЕРІАЛ

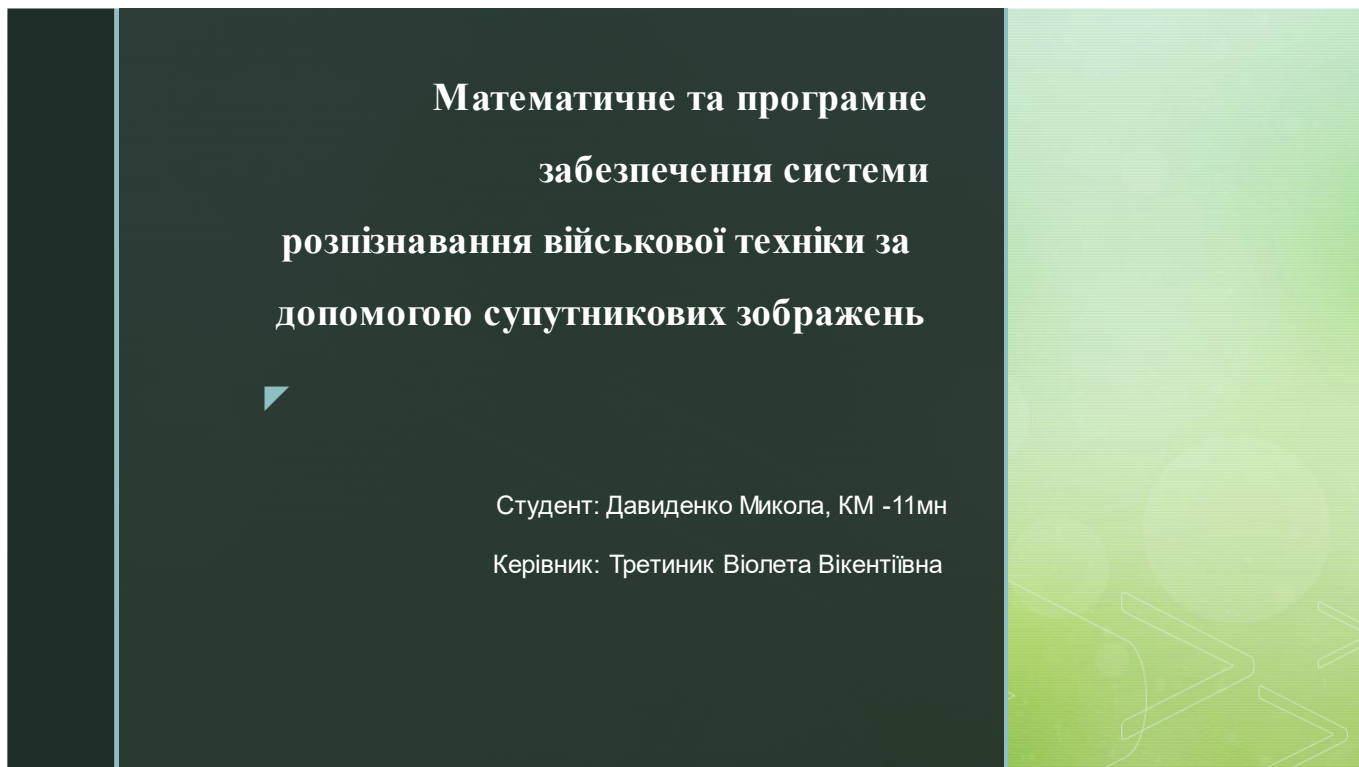


Рисунок Б.1 – Слайд 1

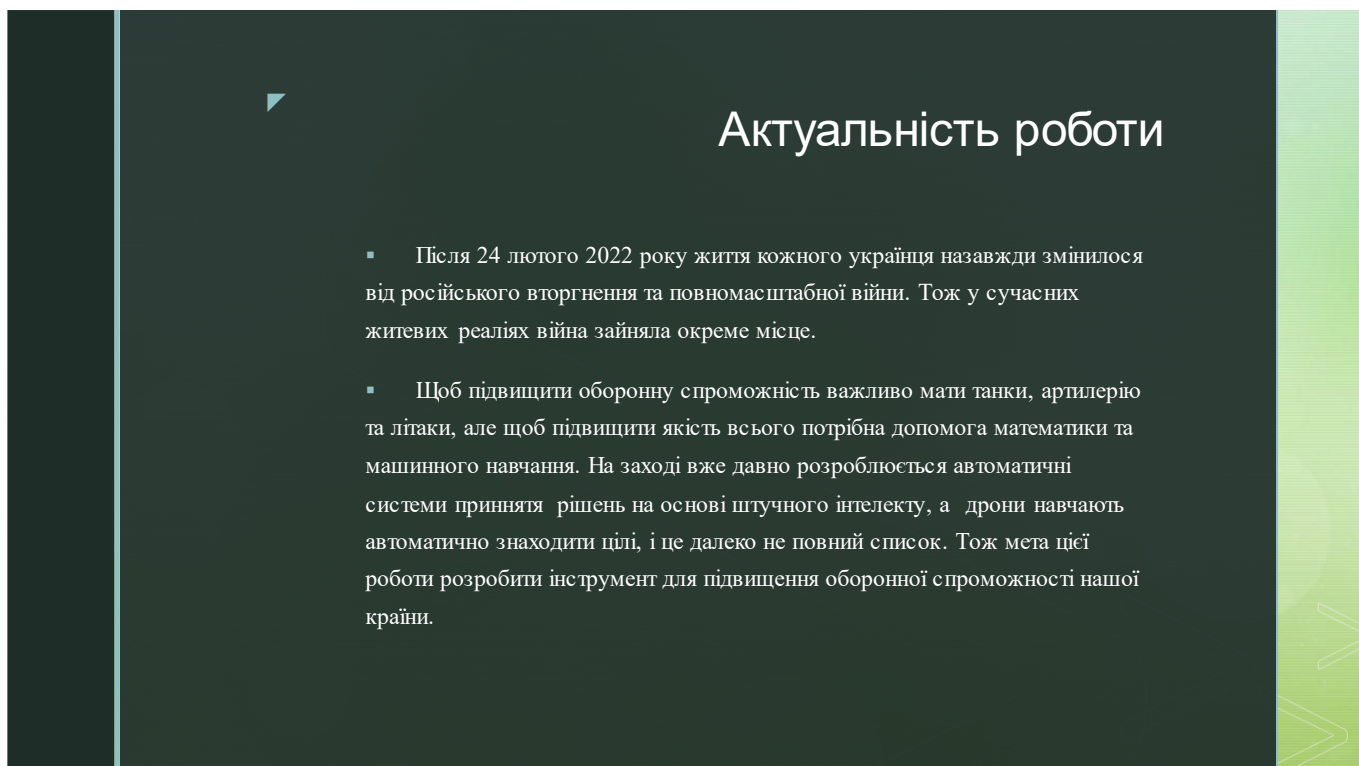


Рисунок Б.2 – Слайд 2

Мета і задачі дослідження

- Розробка математичного та програмного забезпечення системи розпізнавання військової техніки за допомогою супутникових зображень на основі нейромережевої моделі для інтеграції з сайтами. Для розширення функціональних можливостей і адаптації системи до використання людей без профільної освіти. Програма має бути інтегрована з сайтом, додатком або впровадження в робочий процес.

Рисунок Б.3 – Слайд 3

Об'єкт дослідження

- Методи та моделі розпізнавання об'єктів за супутниковими знімками на основі методів та моделей нейронних мереж. Моделі підвищення якості зображення, апскейлінг. Різні архітектури нейронних мереж: згорткові нейронні мережі, глибока згорткова нейронна мережа, CNN, Залишкова нейронна мережа. Методи оптимізації та параметри нейронної мережі.

Рисунок Б.4 – Слайд 4

Предмет дослідження

- Математичне та програмне забезпечення системи розпізнавання військових об'єктів за супутниковими знімками на основі нейромережевої моделі. Застосування нейронних мереж на етапі попереднього тренування для розпізнавання військових об'єктів за супутниковими знімками, порівняльний аналіз методів оптимізації, що використовуються при навчанні нейронних мереж.

Рисунок Б.5 – Слайд 5

МЕТОДИ ФІКСАЦІЇ

- Радар із синтезованою діафрагмою - це бічна когерентна радіолокаційна система бічного огляду, як правило встановлюють на літаки чи супутники, в якій рух платформи використовується для електронного моделювання надзвичайно великої діафрагми
- Суть методу синтезу діафрагми полягає в тому, що радіолокаційні дані, отримані в кожному окремому періоді зонда, залишаються спільно з даними про розташування платформи, що відповідають цим періодам

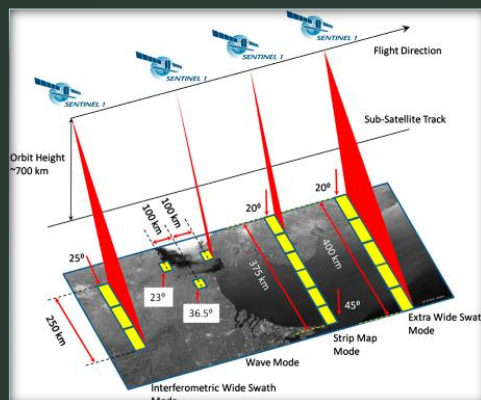
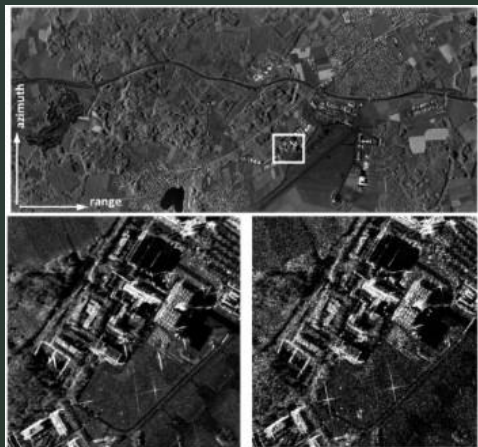


Рисунок Б.6 – Слайд 6

TerraSAR-X



- TerraSAR-X - це німецький космічний радар, який був запущений на орбіту у червні 2007 року. Це супутник зі штучним діелектричним відмінністю, що дозволяє йому досліджувати Землю, використовуючи мікрохвильові радарні сигнали. TerraSAR-X був розроблений з метою отримання високоточних даних для застосування в таких галузях, як картографія, геологія, метеорологія, а також військові дослідження.

Рисунок Б.7 – Слайд 7

WorldView-3

- WorldView -3 - це комерційний супутник високої роздільної здатності, запущений компанією DigitalGlobe в серпні 2014 року. Він є одним з найбільш потужних супутників, які в даний час перебувають у космосі, і має здатність робити знімки з висоти 617 км з роздільною здатністю до 31 см на піксель.



Рисунок Б.8 – Слайд 8

Набір даних

- В якості набору даних було обрано датасет Mstar (для виявлення та розпізнавання рухомих і стаціонарних цілей), створений Агентством передових оборонних дослідницьких проєктів США (DARPA) і Дослідницькою лабораторією ВПС США з 1995 по 1997 рік. Він має 9 класів по 1000 фотографії кожного класу.

Рисунок Б.9 – Слайд 9

Набір даних

- Дані містять такі 9 класів:
 - 2С1 Гвоздика — Самохідна артилерійська установка
 - ЗСУ-23—4 Шилка — зенітна самохідна установка
 - БРДМ-2 — броньована машина -розвідник -амфібія
 - БТР-60 — бронетранспортер
 - D7 — Бульдозер Caterpillar
 - ЗІЛ-131 — Військово-вантажний автомобіль
 - Т-62 — Основний бойовий танк
 - Т-72 — основний бойовий танк 2 -го покоління
 - SLICY — структура, що діє як муляж(а не транспортний засіб)

Рисунок Б.10 – Слайд 10

Приклади техніки



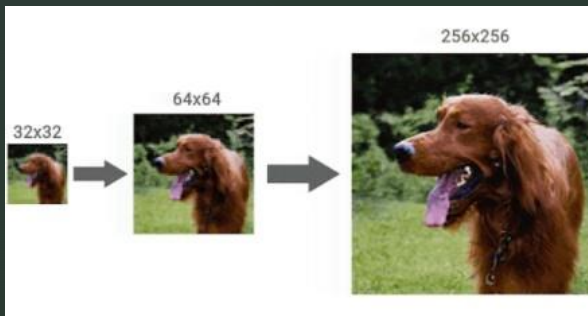
Рисунок Б.11 – Слайд 11



Приклад зображень

Рисунок Б.12 – Слайд 12

Опис існуючих методів підвищення якості фотографії (Upscaling)



Стандартні алгоритми:

- Бікубічна інтерполяція
- Білінійна інтерполяція

Нейронні мережі:

- CNN
- GAN

Рисунок Б.13 – Слайд 13

EDSR

EDSR (Enhanced Deep Residual Networks for Single Image Super-Resolution) - це модель яку можна використовувати для збільшення роздільної здатності зображення

Основні характеристики:

- Архітектура у виді ResNet
- Працює без шарів пакетної нормалізації
- Використовується для коефіцієнта масштабування 2x. Потім ці попередньо підготовлені ваги використовуються під час навчання для коефіцієнта масштабування 3x і 4x.

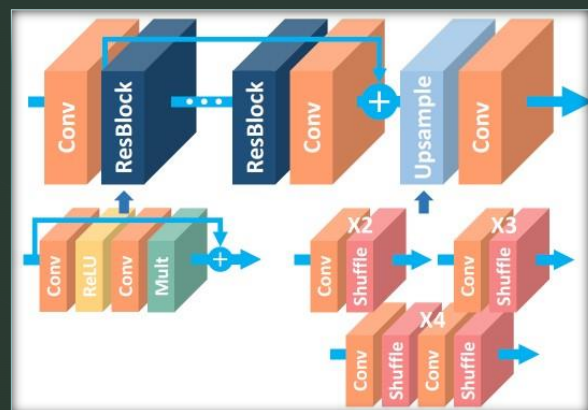


Рисунок Б.14 – Слайд 14

ESPCN

ESPCN (Ecient sub-pixel convolutional neural network)-ефективна субпіксельна згорткова нейронна мережа

Основні характеристики :

- Архітектура подібна CNN але є додаковий ефективний субпіксельний згортковий шар
- Використовує функцію активації Tanh, що працює краще ніж звичайна ReLu
- Масштабування шарів розвертаються тільки в кінці мережі, що робить модель швидшою

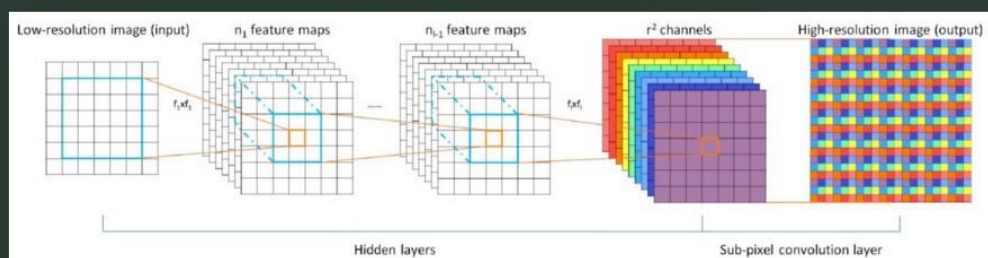
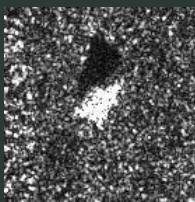
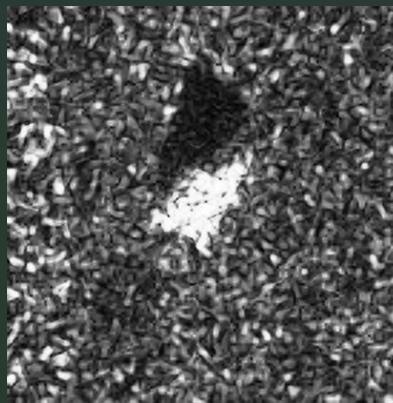


Рисунок Б.15 – Слайд 15

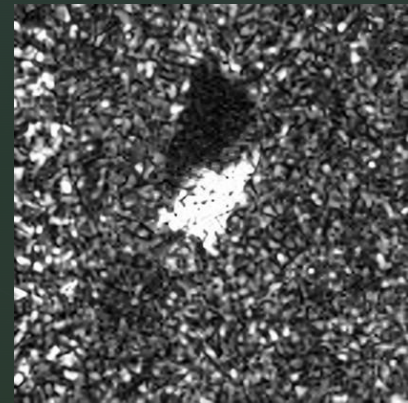
Результати Upscaling



Оригінальне
зображення



ESPCN



EDSR

Рисунок Б.16 – Слайд 16

Результати Upscaling

Таблиця 1 - Середні
значення PSNR

Назва моделі	3x
EDSR	26.554
ESPCN	25.854

Таблиця 2 - Середні
значення SSIM

Назва моделі	3x
EDSR	0,831
ESPCN	0,787

Таблиця 3 - Середній час у
секундах

Назва моделі	3x
EDSR	12,018
ESPCN	0,042

Рисунок Б.17 – Слайд 17

Математична модель

- В якості моделі було обрано нейронні мережі різних видів. Основну увагу було зроблено на згорткових нейронних мережах, так як вони зроблені для роботи з зображеннями.
- Згорткові нейронні мережі (CNN) аналогічні традиційним нейронним мережам в тому, що вони складаються з нейронів, які самооптимізуються через навчання.
- Кожен нейрон все одно отримуватиме вхідні дані та виконуватиме операцію (наприклад, скалярний продукт, за яким слідує нелінійна функція) - основа незліченних ШНМ. Від вхідні необроблені вектори зображень до остаточного результату оцінки класу, вся мережа все ще виражатиме єдину перцептивну функцію оцінки (вагу).
- Останній рівень міститиме функції втрати, пов'язані з класами та всіма регулярні поради та підказки, розроблені для традиційних ШНМ, все ще застосовуються.

Рисунок Б.18 – Слайд 18

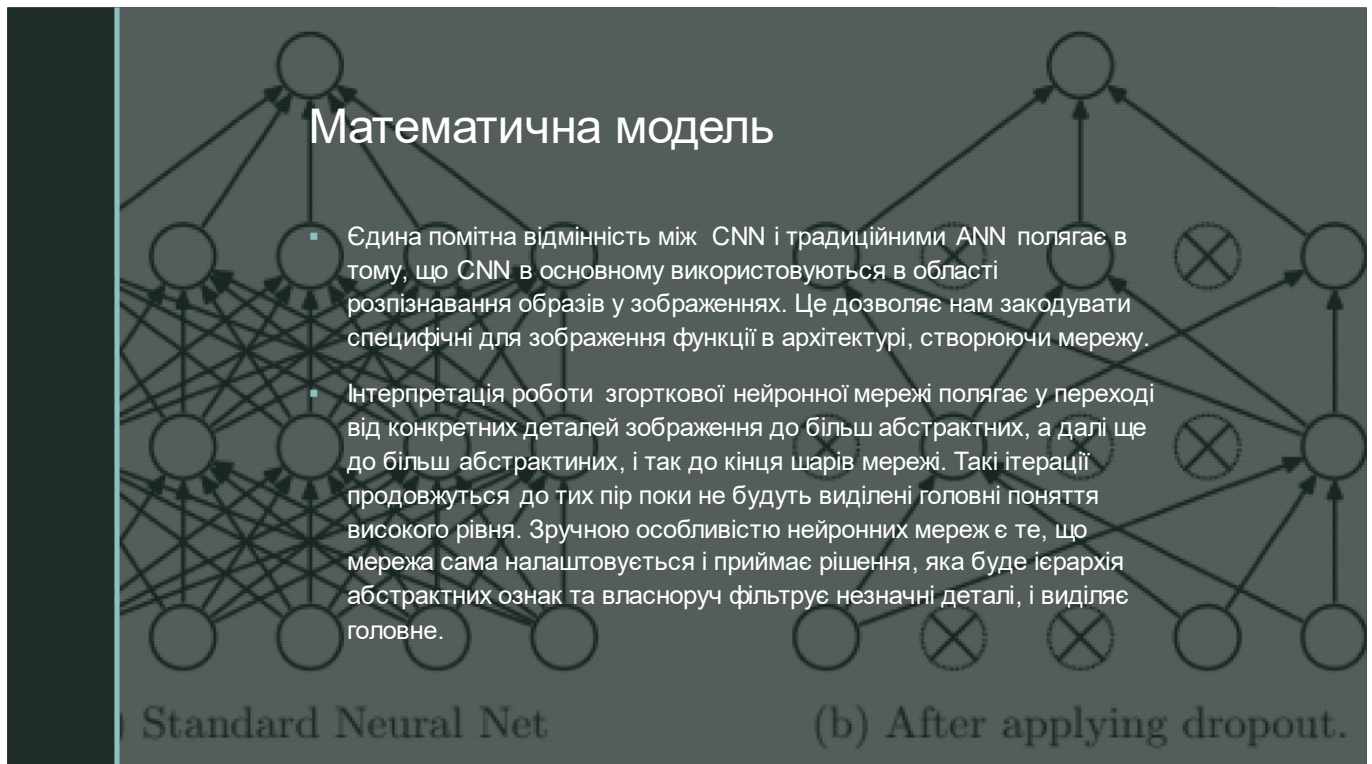


Рисунок Б.19 – Слайд 19

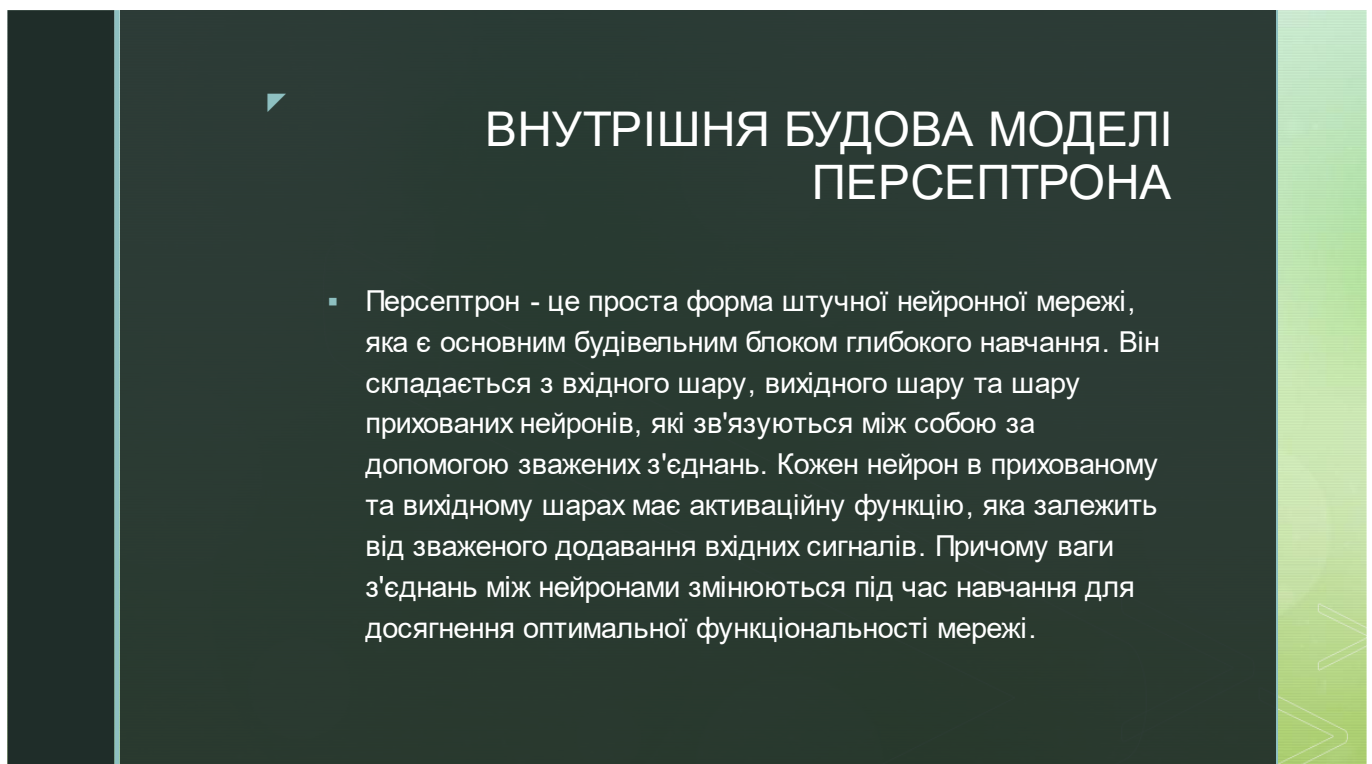
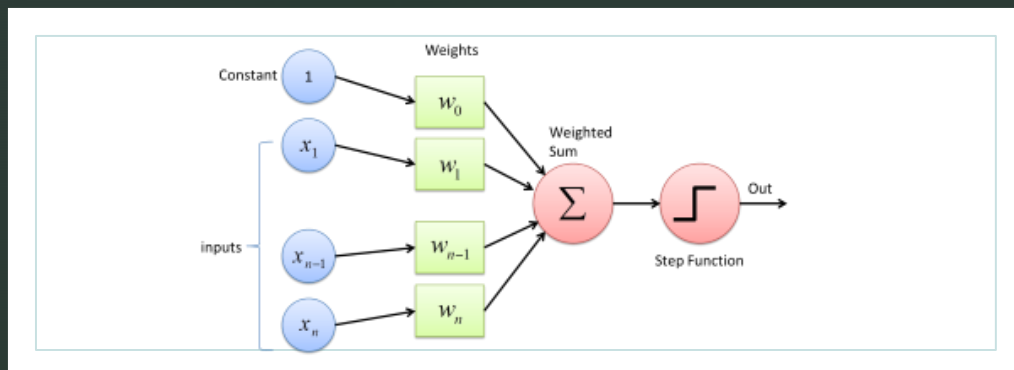


Рисунок Б.20 – Слайд 20



Архітектура Персептрона

Рисунок Б.21 – Слайд 21

Метрики вимірювання точності

- Categorical Accuracy (Категорична точність обчислює відсоток прогнозованих значень, які збігаються з фактичними значеннями для одноразових міток). Середнє значення показників точності для 9 класів.
- Середньозважена оцінка F2. Це зважене гармонічне середнє значення точності та відкликання. Вихідний діапазон становить [0, 1]. Працює як для класифікації з кількома класами, так і з кількома мітками.
- Коефіцієнт кореляції Метьюза. Це числове представлення ефективності класифікатора. 1 = Ідеально, 0 = Випадкове вгадування.

Рисунок Б.22 – Слайд 22

Опис розробленої системи

Реалізована система представляє собою веб-застосунок, побудований за допомогою наступних технологій:

- HTML5, CSS, Bootstrap
- Javascript (MathJax)
- Python (Flask, Sympy, Numpy, Scipy, Matplotlib, Scikit-learn, Pandas)

Спершу дані перевіряються на вхідний розмір зображення. Якщо розмір підходить, фото проходять через мережу, яка вже була навчена, та виводить результат класифікації.

Рисунок Б.23 – Слайд 23

Складові системи

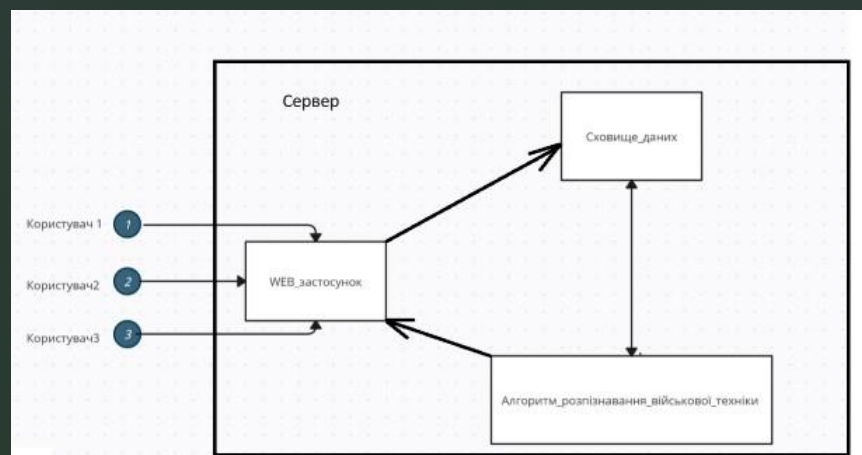


Рисунок Б.24 – Слайд 24

Интерфейс програми

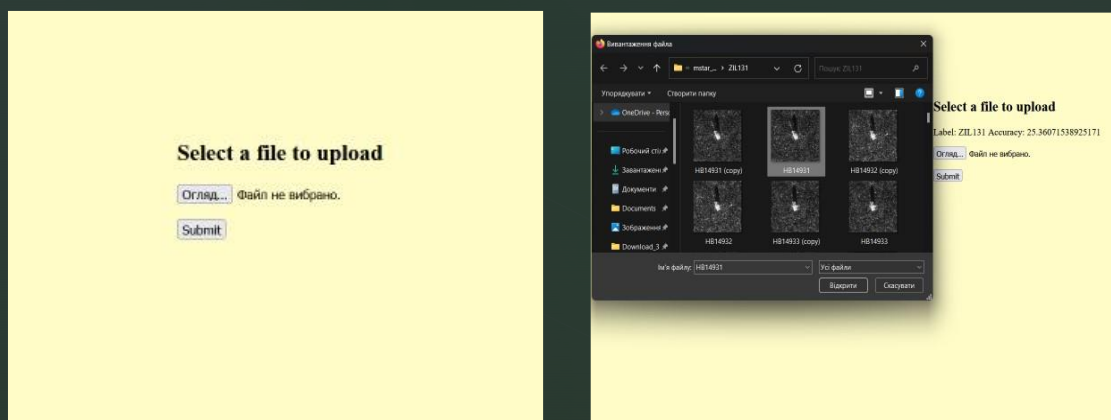


Рисунок Б.25 – Слайд 25

Интерфейс програми



Рисунок Б.26 – Слайд 26

Демонстрація результатів роботи програми

На малюнку видно що
багатошаровий перцептрон
не зміг відрізнити класи у
тестовій вибірці, через свою
просту архітектуру

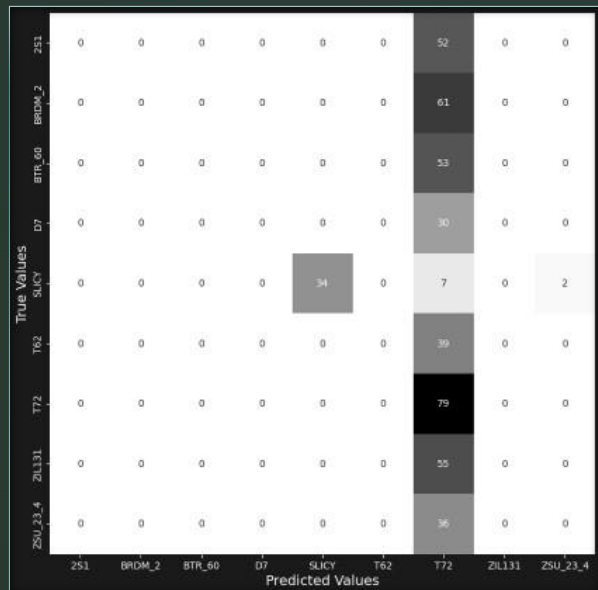


Рисунок Б.27 – Слайд 27

Демонстрація результатів роботи програми

Після навчання на
тренувальній вибірці, модель
CNN показала наступні
результати:

loss: 0.362,

Categorical_accuracy: 0.822,

MCC: 0.802,

F2: 0.813,

Auc: 0.991

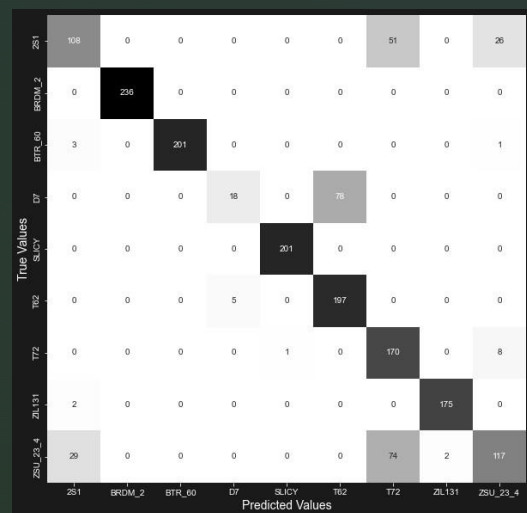


Рисунок Б.28 – Слайд 28

Демонстрація результатів роботи програми

Після навчання на тренувальній вибірці, моделі на основі CNN з додатковими шарами регуляризації показали наступні результати:

loss: 0.119,

Categorical_accuracy: 0.961,

MCC: 0.956,

F2: 0.961,

Auc: 0.997

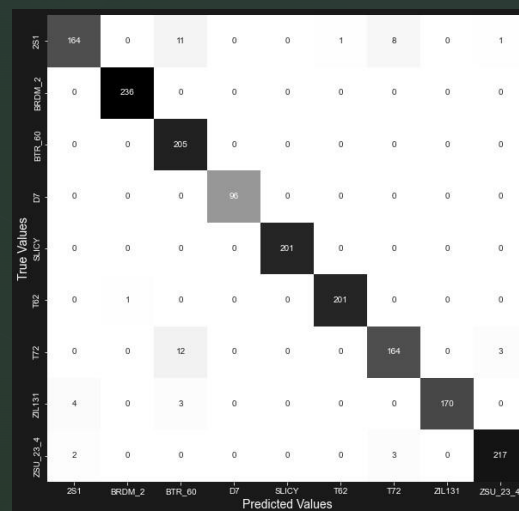


Рисунок Б.29 – Слайд 29

Висновки

- Розроблено математичне та програмне забезпечення системи розпізнавання військової техніки за допомогою супутникових зображень. В системі застосовано низку класичних методів машинного навчання, а саме згорткових нейронних мереж серед яких визначається найкращий метод для заданого набору вхідних даних.

В результаті проведеного дослідження встановлено, що для обраного датасету найкращим методом для розпізнавання військової техніки є нейронна мережа на основі CNN з додатковими шарами регуляризації. Результуюча точність алгоритму складає з Categorical_accuracy: 0.961 та з MCC: 0.956.

Рисунок Б.30 – Слайд 30

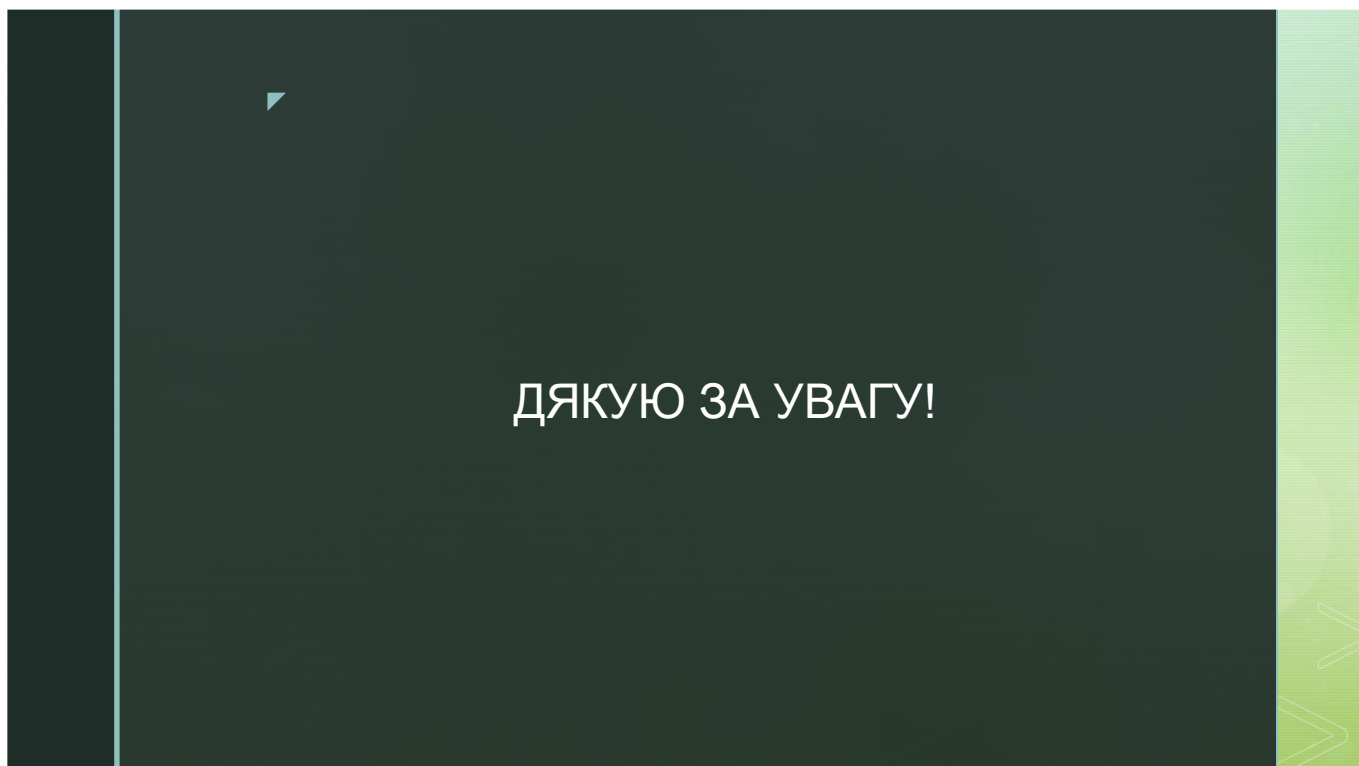


Рисунок Б.31 – Слайд 31