

«Пожежна безпека»

(шифр)

**ДОСЛІДЖЕННЯ ПОЖЕЖОНЕБЕЗПЕЧНОСТІ ЛІСОВИХ ТЕРИТОРІЙ НА
ОСНОВІ ВИКОРИСТАННЯ КАПСУЛЬНИХ ТА ЗГОРТКОВИХ НЕЙРОННИХ
МЕРЕЖ**

Галузь: Інформаційні технології

2018/2019

АНОТАЦІЯ

У сучасному світі нейронні мережі інтенсивно розвиваються і використовуються в усіх сферах людської діяльності. Їх застосування для визначення пожежонебезпечності територій лісу може розпочати вирішення проблеми попередження лісових пожеж.

Метою роботи є дослідження можливості застосування нейронних мереж глибокого навчання для визначення пожежонебезпечних лісових територій на основі мультиспектральних зображень дистанційного зондування Землі.

Поставлена мета передбачає вирішення таких **завдань**: а) отримати вхідні дані з супутника Landsat 8; б) обробити мультиспектральні зображення та розрахувати спектральні індекси; в) реалізувати згорткову та капсульну нейронні мережі; г) провести порівняльний аналіз результатів роботи капсульної та згорткової нейронних мереж.

Методи дослідження: теорія нейронних мереж глибокого навчання, теорія розпізнавання мультиспектральних зображень, методи математичної статистики.

Загальна характеристика роботи. Робота присвячена вирішенню науково-прикладної задачі визначення пожежонебезпечних лісових територій із використанням нейронних мереж глибокого навчання на прикладі пожежі «Camp Fire». Реалізовано згорткову та капсульну нейронні мережі, що здатні детектувати пожежонебезпечні території на основі мультиспектральних зображень. Проведено їх порівняльний аналіз.

Обсяг роботи становить 30 сторінок, 14 рисунків, 3 таблиці, 7 формул і 25 джерел інформації.

Ключові слова: згорткові нейронні мережі, капсульні нейронні мережі, лісові пожежі, мультиспектральні зображення, спектральні індекси.

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1 АНАЛІЗ ПРОБЛЕМИ ПОПЕРЕДЖЕННЯ ЛІСОВИХ ПОЖЕЖ ТА ПОСТАНОВКА ЗАДАЧІ	6
РОЗДІЛ 2 ВИКОРИСТАННЯ МУЛЬТИСПЕКТРАЛЬНИХ ЗОБРАЖЕНЬ ДЛЯ РОЗРАХУНКУ СПЕКТРАЛЬНИХ ІНДЕКСІВ	9
2.1 Загальна характеристика дистанційного зондування Землі	9
2.2 Отримання даних дистанційного зондування Землі	10
2.3 Розрахунок спектральних індексів	12
РОЗДІЛ 3 ЗАСТОСУВАННЯ КАПСУЛЬНИХ ТА ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИЗНАЧЕННЯ ПОЖЕЖОНЕБЕЗПЕЧНОСТІ ЛІСОВИХ ТЕРИТОРІЙ	16
3.1 Огляд технології глибокого навчання	16
3.2 Згорткові нейронні мережі	17
3.3 Капсульні нейронні мережі	19
3.4 Архітектури згорткової та капсульної нейронних мереж, що використовувались у дослідженні	21
3.5 Підготовка навчального, затверджувального та тестового наборів даних	23
3.6 Результати дослідження та їх обговорення	24
ВИСНОВКИ	30
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	31

ВСТУП

У сучасному світі нейронні мережі інтенсивно розвиваються й проникають в усі сфери людської діяльності. Фахівці в галузі програмного забезпечення використовують нейронні мережі для боротьби з комп'ютерними вірусами, медичні працівники – для діагностики та прогнозування лікування, банкіри – для оцінки кредитних ризиків, курсів облігацій, акцій та валют. Застосовуються нейронні мережі й у автоматизованих системах керування, й у інформаційних системах для обробки мовної інформації, в автоматичному перекладі, при обробці текстів, для розв'язання задач кодування та декодування інформації та ін.

У даному дослідженні нейронні мережі глибокого навчання використовуються для вирішення науково-прикладної задачі визначення пожежонебезпечних лісових територій на прикладі пожежі «Camp Fire», що сталася в листопаді 2018 року (Каліфорнія, США). В останні роки пожежі набувають величезних масштабів, важко контролюються та, у разі виникнення, вимагають великої кількості ресурсів для їх усунення.

Метою роботи є дослідження можливості застосування нейронних мереж глибокого навчання для визначення пожежонебезпечних лісових територій на основі мультиспектральних зображень дистанційного зондування Землі.

Поставлена мета передбачає вирішення таких **завдань**:

- огляд територій, на яких відбулися наймасштабніші лісові пожежі за останній час, аналіз економічних та екологічних збитків від лісових пожеж;
- отримання та обробка мультиспектральних зображень території пожежі з супутника Landsat 8;
- розрахунок спектральних індексів NDVI, NDWI, PSRI;
- реалізація згорткової нейронної мережі з використанням бібліотеки Keras;
- реалізація капсульної нейронної мережі з використанням бібліотеки Keras та власних розробок;

– порівняльний аналіз результатів роботи капсульної та згорткової нейронних мереж.

Об’єкт дослідження – процес визначення пожежонебезпечних лісових територій з використанням капсульної та згорткової нейронних мереж.

Предмет дослідження – процес розпізнавання мультиспектральних зображень з використанням нейронних мереж глибокого навчання.

Методи досліджень – теорія нейронних мереж глибокого навчання, теорія розпізнавання мультиспектральних зображень, методи математичної статистики.

Наукова новизна – удосконалено метод розпізнавання мультиспектральних зображень за рахунок використання капсульних та згорткових нейронних мереж.

РОЗДІЛ 1

АНАЛІЗ ПРОБЛЕМИ ПОПЕРЕДЖЕННЯ ЛІСОВИХ ПОЖЕЖ ТА ПОСТАНОВКА ЗАДАЧІ

У 2017 році сталися лісові пожежі у Британській Колумбії (Канада). Згоріло 1.2 мільйони акрів лісу та спричинено збитків більше ніж на півмільярда доларів США. Окрім цього, забруднення атмосфери внаслідок лісових пожеж спричинило респіраторні захворювання у людей, що жили у радіусі сотень кілометрів.

У липні 2018 року в штаті Каліфорнії (США) трапилася одна з найбільших лісових пожеж під назвою «Mendocino Complex». Загальна площа пожежі становила майже 460,000 акрів. Менше ніж через півроку в цьому штаті знову відбулась пожежа, що отримала назву «Camp Fire», площею більше 150,000 акрів. Ця пожежа забрала 86 людських життів [1].

Аналізуючи статистику США [2] можна спостерігати, що порівняно з 1980-тими роками на даний момент кількість акрів землі, на якій відбувалися пожежі збільшилася майже вдвічі, у той час як кількість пожеж стала меншою (рис. 1.1).

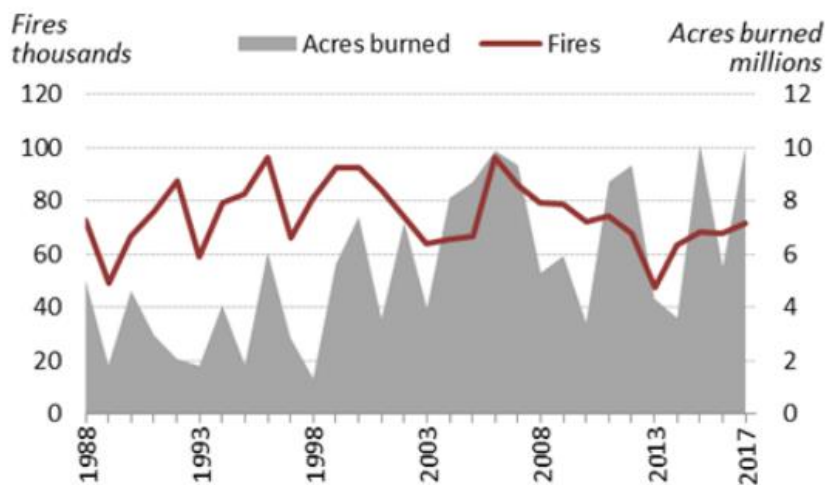


Рисунок 1.1 – Загальна кількість пожеж та згорілих територій у США (з 1988 по 2017 роки). Лінією зазначено кількість пожеж (у тисячах). Зафарбована площа – згоріла територія (у мільйонах акрів)

На рисунку 1.2 зображено гістограму за сумарною найбільшою площею пожеж у США з 1960 по 2017 роки. Найбільші площі лісів згоріли у 2015 та 2017 роках – більше 10 мільйонів акрів. У п'ятірці років з сумарною найбільшою площею пожеж не опинилося жодного року, меншого за 2006, що свідчить про те, що проблема лісових пожеж у сучасності завдає більших втрат, ніж у XX столітті.

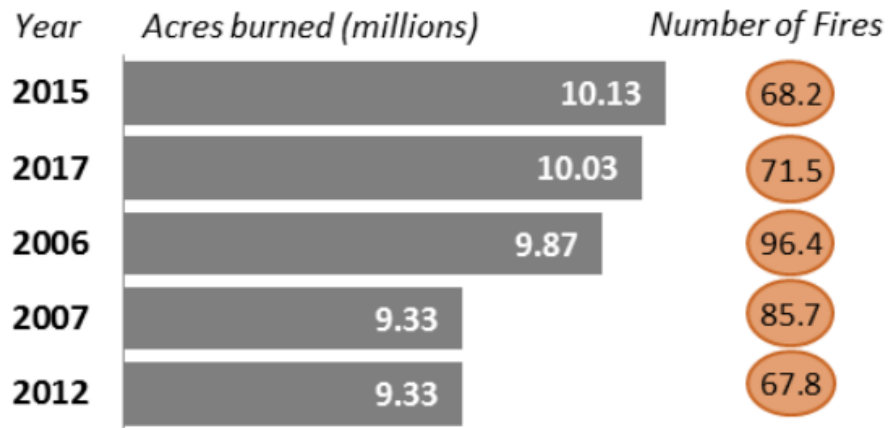


Рисунок 1.2 – Сумарна найбільша площа пожеж (з 1960 по 2017 роки) у мільйонах акрів

Першочерговими втратами від великої кількості лісових пожеж є людські життя та природні ресурси. З іншого боку, варто пам'ятати, що пожежі спричиняють чималі економічні витрати на їх подолання та усунення наслідків. На рисунку 1.3 зображено фінансові втрати США, спричинені пожежами. Наприклад, у 2017 році пожежі завдали збитків майже на 20 мільярдів доларів США. У той час як, наприклад, у 2008 та 2015 роках було витрачено приблизно 2,5 мільярди доларів, а у 2014, 2010 та 2009 – менше мільярда [3].

Wildfire Losses In The United States, 2008-2017 (1)

(2017 \$ millions)

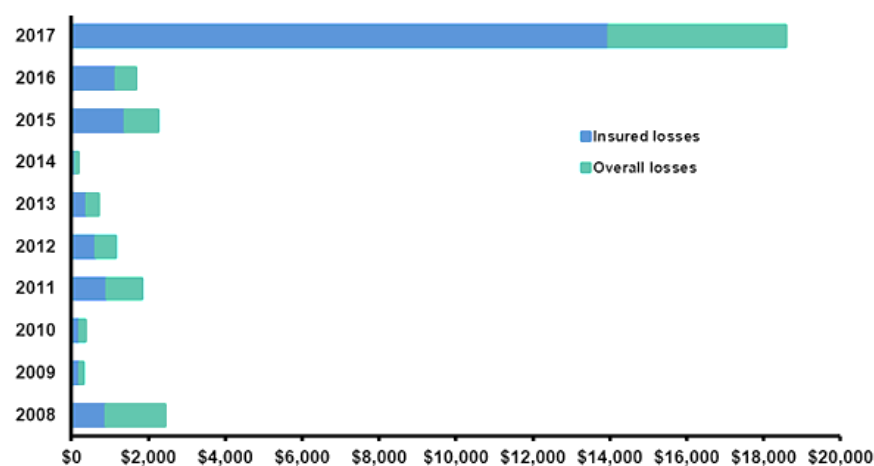


Рисунок 1.3 – Фінансові втрати у США, спричинені пожежами (у мільйонах доларів). (Світліший колір – загальні втрати, темніший колір – страхові збитки)

Отримавши змогу попереджувати появу лісових пожеж можна зберегти людські життя та здоров'я, природні багатства, екологію та економічні ресурси.

Для вирішення такого важливого завдання будуть використані нейронні мережі глибокого навчання (згорткова та капсульна), що матимуть здатність розпізнавати пожежонебезпечні території. Буде використано мультиспектральні зображення з супутника Landsat 8 [5]. На основі цих даних ми розрахуємо необхідні спектральні індекси, що допомагають виділити засушливу рослинність, вміст вологи та карбону (у вигляді лігніна та целюлози). Нейронні мережі визначатимуть імовірність пожежонебезпечності територій, що дозволить зробити відповідні висновки. Такий підхід може розпочати вирішення проблеми попередження лісових пожеж.

РОЗДІЛ 2

ВИКОРИСТАННЯ МУЛЬТИСПЕКТРАЛЬНИХ ЗОБРАЖЕНЬ ДЛЯ РОЗРАХУНКУ СПЕКТРАЛЬНИХ ІНДЕКСІВ

2.1 Загальна характеристика дистанційного зондування Землі

Дистанційне зондування Землі (ДЗЗ) визначають як метод вимірювання властивостей об'єктів на земній поверхні, в якому використовуються дані, отримані за допомогою повітряних літальних апаратів і штучних супутників Землі [5]. Сутність такого методу полягає в тому, щоб замість проведення вимірювань за місцем розташування об'єкта спробувати виміряти його характеристики на відстані. Апаратура ДЗЗ, особливо та, яка встановлюється на штучному супутнику Землі, дозволяє проводити періодичну зйомку земної поверхні і спостерігати за змінами навколишнього середовища та впливу на нього людини. Дистанційне зондування Землі застосовується у таких областях:

- моніторинг і оцінка стану навколишнього середовища;
- спостереження за поновлюваними природними ресурсами (заболочені території, ґрунт, ліси, океани);
- спостереження за глобальними змінами (виснаження озонового шару атмосфери, глобальне потепління);
- сільське господарство (аналіз стану сільськогосподарських культур, прогноз врожаю, дослідження ерозії ґрунтів);
- метеорологія (дослідження динаміки атмосфери, прогноз погоди);
- картографія (топографія, землекористування, цивільне будівництво) тощо.

Для задоволення потреб різних категорій фахівців була розроблена велика кількість систем ДЗЗ, за допомогою яких вивчають просторові, спектральні і часові параметри найрізноманітніших об'єктів. Кожна з цих систем орієнтована на певне коло завдань і має свої особливості. Так, наприклад, у метеорології

використовується частина території з відносно низькою роздільною здатністю. У картографії, навпаки, основною вимогою є максимальна просторова роздільна здатність, а періодичність зйомки може бути відносно низькою. Правильно відкалібровані дані ДЗЗ можна використовувати для перевірки глобальних кліматичних моделей, призначених для моделювання та прогнозу розвитку природного середовища.

Мультиспектральним називають зображення, яке захоплює дані на певних частотах по всьому електромагнітному спектру [6]. Довжини хвиль можуть бути розділені за допомогою фільтрів або за допомогою приладів, які чутливі до певних довжин хвиль світла, у тому числі частот, що знаходяться за межами діапазону видимого світла. Спектральне зображення може дозволити отримати додаткову інформацію, яку людське око не в змозі захопити. Мультиспектральні зображення – це основний тип зображень, що отримують за допомогою дистанційного зондування Землі. Як правило, супутники мають три або більше радіометрів. Кожен з них відноситься до одного цифрового зображення в невеликому діапазоні видимого спектру, починаючи від 0,7 мкм до 0,4 мкм і доходячи до інфрачервоних довжин від 0,7 мкм до 10 мкм. Спектральні зображення з більш чисельними діапазонами, кращою спектральною роздільною здатністю або більш широким спектральним діапазоном називаються гіперспектральними або ультраспектральними зображеннями.

2.2 Отримання даних дистанційного зондування Землі

Програма Landsat - найбільш тривалий проект по отриманню супутникових фотознімків планети Земля. Дані, отримані за допомогою Landsat, використовуються при вирішенні великої кількості тематичних завдань, таких як виміри протяжності та класифікацію рослинного покриву, визначення стану сільськогосподарських культур, геологічне картування, контроль ерозії ґрунтів у

береговій зоні та ін. Найбільш актуальні дані отримують з супутника Landsat 8, американського супутника дистанційного зондування Землі. Дані з цього супутника використовуються для:

- оперативного супутникового контролю природних ресурсів;
- дослідження динаміки протікання природних процесів і явищ;
- аналізу причин, прогнозування можливих наслідків і вибору способів запобігання надзвичайним ситуаціям.

Landsat 8 отримує зображення з роздільною здатністю знімків від 15 до 100 метрів на точку (піксель). Проводиться зйомка суші і полярних регіонів. Кожну добу знімається близько 400 сцен. Періодичність спостереження однієї і тієї ж території становить 16 днів. Знімки Landsat 8 знаходяться у відкритому доступі [7]. Спектральні діапазони супутника Landsat 8 наведено у таблиці 2.1.

Таблиця 2.1 – Спектральні діапазони супутника Landsat 8

Номер каналу	Назва каналу	Довжина хвиль, нм	Роздільна здатність, м
1	Ultra-blue	435-451	30
2	Blue	452-512	30
3	Green	533-590	30
4	Red	636-673	30
5	NIR	851-879	30
6	SWIR1	1566-1651	30
7	SWIR2	2107-2294	30
8	Panchromatic	503-676	15
9	Cirrus	1363-1384	30
10	Thermal infrared 1	1060-1119	100·(30)
11	Thermal infrared 2	1150-1251	100·(30)

В якості вхідних даних з супутника Landsat 8 було обрано територію пожежі «Camp Fire» за 7 жовтня 2018 року. Пожежа розпочалась 8 листопада 2018 року [8]. Координати пожежі - $39^{\circ}50'51''N, 121^{\circ}23'42''W$. На рисунку 2.1 зображено периметр пожежі [9].

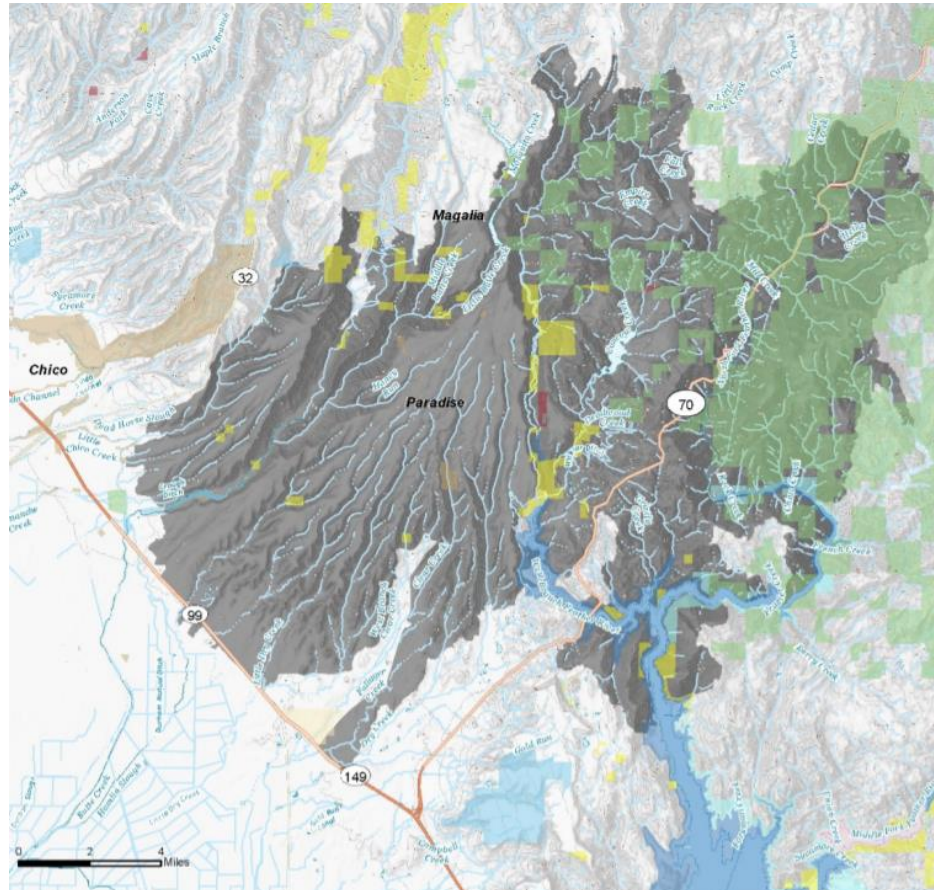


Рисунок 2.1 – Периметр пожежі «Camp Fire»

2.3 Розрахунок спектральних індексів

Для роботи зі спектральною інформацією часто вдаються до створення індексних зображень. На основі комбінації значень яскравості в певних каналах, інформативних для виділення досліджуваного об'єкта, і розрахунку за цими значеннями спектрального індексу об'єкта будується зображення, що відповідає значенню індексу в кожному пікселі, що і дозволяє виділити досліджуваний об'єкт або оцінити його стан.

Для аналізу спектральних зображень у роботі було використано три спектральних індекси – NDVI, NDWI, PSRI.

NDVI (англ. Normalized Difference Vegetation Index) – найбільш відомий та широко використовуваний вегетаційний індекс. Він застосовується у багатьох задачах, включаючи виявлення зеленої рослинності у мультиспектральних зображеннях. На червону зону спектра припадає максимум поглинання сонячної радіації хлорофілом, а на ближню інфрачервону зону - максимальне відбиття енергії. Висока фотосинтетична активність веде до більш низьких значень коефіцієнтів відбиття в червоній зоні спектра і великим значенням в ближній інфрачервоній. Зіставлення показників в даних спектральних діапазонах дозволяє виділяти рослинність. Індекс розраховується за такою формулою:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}, \quad (2.1)$$

де ρ_{NIR} – відображення у ближній інфрачервоній області спектра, а ρ_{RED} – відображення у червоної області спектра [10]. Після застосування формули 2.1 та нормалізації отримано індексне зображення, яке наведене на рисунку 2.2.

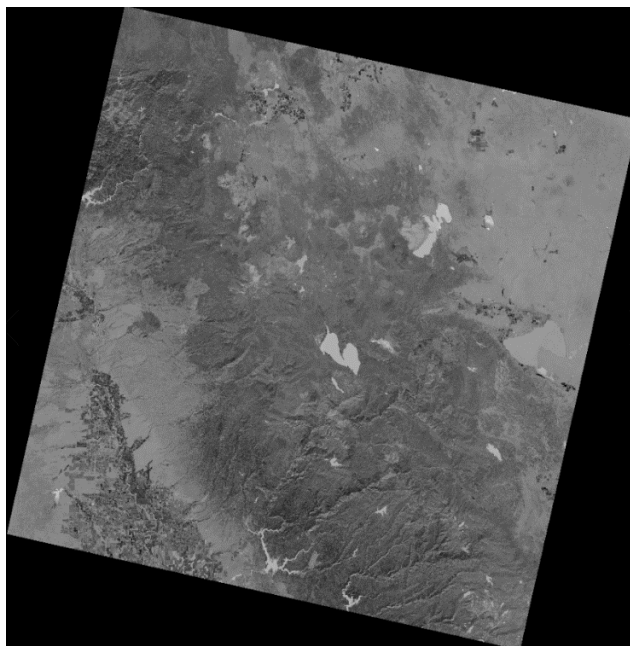


Рисунок 2.2 – Розподіл значень індексу NDVI (після нормалізації) у межах досліджуваної території. Темніший відтінок відповідає зеленій рослинності

NDWI (англ. Normalized Difference Water Index) було сформовано за принципом NDVI. Цей індекс характеризує вміст води – дуже важливий показник, оскільки високий вміст вологи характерний для здорової рослинності, що швидше росте та є стійкішою до пожеж. Розраховують NDWI таким чином:

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}}, \quad (2.2)$$

де ρ_{NIR} – відображення у ближній інфрачервоній області спектра, а ρ_{SWIR} – відображення у короткохвильовій інфрачервоній області спектра.

Під час розрахунку NDWI при обробці мультиспектрального супутникового зображення, водні об'єкти мають позитивні значення, а об'єкти ґрунтової і наземної рослинності мають нульові або негативні значення, внаслідок їх вищого коефіцієнту відбиття в інфрачервоній області спектру, ніж в області зеленого світла. При цьому необхідно провести нормалізацію та виключити від'ємні значення для збереження інформації про водяні об'єкти [11]. Отримане за формулою 2.2 зображення наведене на рисунку 2.3.

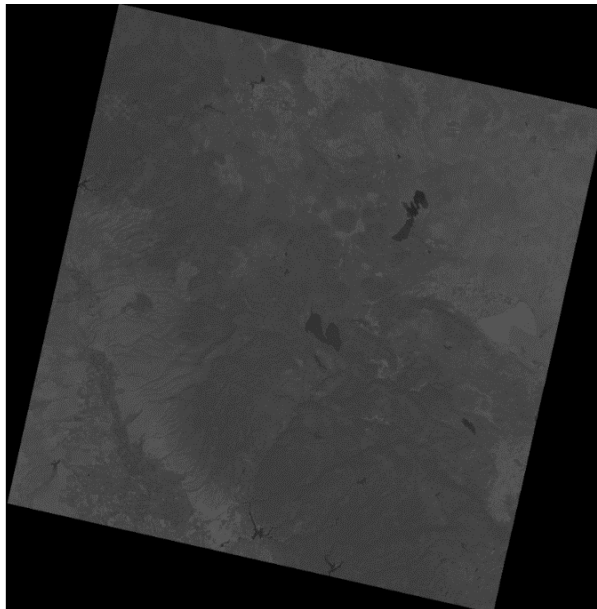


Рисунок 2.3 – Розподіл значень індексу NDWI (після нормалізації) у межах досліджуваної території. Темніший відтінок відповідає водним об'єктам

Індекс PSRI (англ. Plant Senescence Reflectance Index) [12] належить до групи Dry or Senescent Carbon. Ці індекси розроблені для знаходження загальної кількості

«сухого» вуглецю у вигляді лігніну і целюлози. Такий вуглець у великих кількостях присутній у деревині, мертвих або сухих рослинних тканинах. Збільшення цього показника може відображати процес «старіння» і відмирання рослин. Для розрахунку відносного вмісту азоту в рослинному покриві використовується середній інфрачервоний діапазон. Індокси, що належать до групи Dry or Senescent Carbon широко використовуються при оцінці пожежонебезпечності території. PSRI розраховують за формулою 2.3:

$$PSRI = \frac{\rho_{RED} - \rho_{GREEN}}{\rho_{NIR}} \quad (2.3)$$

де ρ_{NIR} – відображення у ближній інфрачервоній області спектра, а ρ_{GREEN} – відображення зеленої області спектра, ρ_{RED} - відображення червоної області спектра. Розраховане індексне зображення наведено на рисунку 2.4.

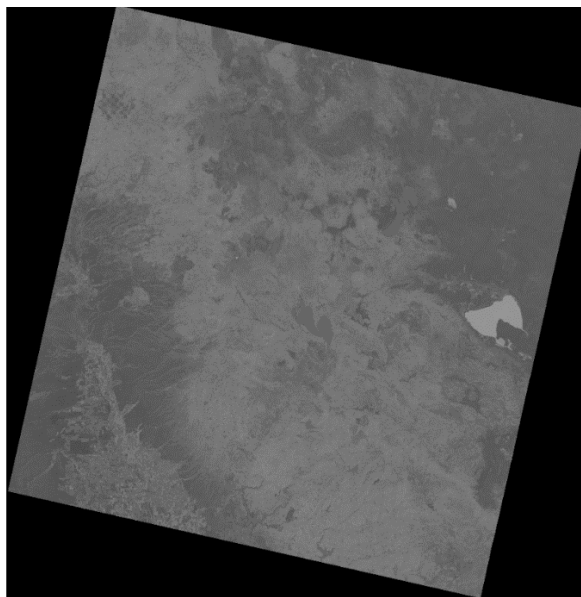


Рисунок 2.4 – Розподіл значень індексу PSRI (після нормалізації) у межах досліджуваної території. Темніший відтінок відповідає за більший вміст карбону

Індексні зображення будуть об'єднані та використовуватимуться нейронними мережами для визначення пожежонебезпечності лісових територій.

РОЗДІЛ 3

ЗАСТОСУВАННЯ КАПСУЛЬНИХ ТА ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИЗНАЧЕННЯ ПОЖЕЖОНЕБЕЗПЕЧНОСТІ ЛІСОВИХ ТЕРИТОРІЙ

3.1 Огляд технології глибокого навчання

Глибоке навчання - це сукупність методів, що дозволяють багатошаровим обчислювальним моделям працювати з даними, що мають ієрархію абстракцій. У роботі [13] наведено огляд основних тенденцій глибокого навчання. Ця сукупність методів використовується, наприклад, у згорткових нейронних мережах. Глибокі методи навчання представляють собою методи, що складаються з декількох простих, але нелінійних модулів - шарів. Наприклад, зображення представляється у формі масиву значень пікселів і вивчені властивості у першому шарі передаються наступному. Другий шар уточнює наявні характеристики досліджуваного об'єкта. Третій шар може збирати дані вже у вигляді частин об'єктів, а наступні шари будуть виявляти об'єкти як комбінації цих частин. Глибоке навчання робить великі успіхи у вирішенні багатьох задач. Воно використовується для розпізнавання мови [14-16], для прогнозування активності потенційних молекул лікарських засобів [17], аналізу даних прискорювача частинок [18].

Навчання з учителем (англ. Supervised learning) є найбільш поширеною формою машинного навчання. Для класифікації об'єктів збирається велика кількість даних (зображень), що розподіляються у певні категорії за відповідними властивостями. Під час навчання система отримує зображення на вхід та видає результат на виході у вигляді ймовірності знаходження шуканого об'єкта. Після отримання перших результатів, цільова функція вимірює допущені помилки між отриманими результатами та вихідними значеннями. Система регулює свої внутрішні параметри для зменшення кількості помилок та знову намагається

класифікувати вхідні дані. На практиці використовується поняття градієнтного спуску (англ. SGD) – знаходження локального екстремуму функції, за рахунок руху уздовж градієнта функції. У нейронних мережах це полягає в обчисленні середнього градієнта та коригуванні вагових коефіцієнтів. Цей процес продовжуються, поки цільова функція (функція помилки) не припинить своє зменшення.

3.2 Згорткові нейронні мережі

Згорткова нейронна мережа (Convolutional neural network, CNN) була запропонована Яном Лекуном у 1988 році [19]. Основний практичний успіх згорткових нейронних мереж - це розпізнавання зображень. Продуктивність систем розпізнавання на основі згорткових мереж викликала використання їх більшістю великих технологічних компаній, включаючи Google, Facebook, Microsoft, IBM, Yahoo!, Twitter і Adobe, а також швидкозростаюче число стартапів, проектів і розгортання продуктів для розпізнавання на основі згорткової мережі [13]. Структура CNN не включає в себе зворотні зв'язки, є односпрямованою та багат шаровою. Загальна ідея згорткових нейронних мереж полягає в чергуванні згорткових шарів і шарів підвибірки. Для навчання такої нейронної мережі найчастіше використовують метод зворотного поширення помилки [20].

Згортка – це математична операція, що застосовується до двох функцій $f(x)$, $g(x)$ і породжує третю функцію [21]. Матрицю K розмірності $k \times w$ називають ядром згортки. Воно кодує деяку виявлену ознаку. Маючи зображення I , обчислюється згорнуте зображення $I * K$, накладаючи ядро на зображення, і записується сума добутків елементів вихідного зображення і ядра:

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} \times I_{x+i-1,y+j-1} \quad (3.1)$$

Звичайну архітектуру згорткової нейронної мережі для розпізнавання зображення по k класам можна розділити на дві частини: послідовність шарів

згортки чи підвибірки, що чергуються, та декілька повнозв'язних шарів (шари, які приймають кожний піксель зображення, як незалежне значення).

Основна задача шару підвибірки (пулінг, англ. Pooling) є отримання на вході окремих фрагментів зображення та їх об'єднання в одне значення. Зазвичай використовуються функція максимізації (такий шар називають MaxPooling) і обирається значення з найбільшою величиною. Шар підвибірки не враховує деталі зображення, що обробляється, а звертає увагу тільки на ті дані, що дають найбільший внесок.

Один прохід шарів згортки та пулінгу впливає таким чином: скорочується довжина і ширина певного каналу, але збільшується його глибина. На вихідному шарі нейронної мережі використовується функція активації нейронів. У нашому дослідженні це – сигмоїдальна функція активації:

$$f(x) = \frac{1}{1+e^{-x}}. \quad (3.2)$$

У ході навчання нейронної мережі може виникнути проблема перенавчання, яка полягає в надмірно точній відповідності нейронної мережі конкретному набору навчальних прикладів. Дана проблема виникає у випадку, якщо модель мережі змогла вивчити навчальну множину (разом із шумом, який присутній в ній), проте не змогла розпізнати процеси, які цю множину породили. Для вирішення проблеми перенавчання використовують додатковий шар під назвою Dropout. Під час навчання, використовується параметр p . Проходячи по нейронам одного шару за одну епоху із ймовірністю p повністю виключає їх із мережі на час ітерації. Такі обмеження використовують замість скорочення кількості параметрів у нейронній мережі.

3.3 Капсульні нейронні мережі

За останні роки згорткові нейронні мережі досягли високих результатів, але у своїй архітектурі мають недоліки, на які наголошує провідний науковець у галузі штучних нейронних мереж Джеффрі Хінтон [22].

Як зазначалося у розділі 3.2 CNN мають шар MaxPooling. При використанні такого шару, втрачається частина інформації. Ця інформація може бути дуже корисною для нейронної мережі. Наприклад, ConvNet не здатна врахувати просторові та ієрархічні зв'язки між складовими, що має досліджуваний об'єкт. Мережа може помилково ідентифікувати об'єкт на зображенні, якщо його структурні елементи будуть розташовані у неправильному положенні відносно один одного. Тим самим, згорткові нейромережі необхідно навчати на великій кількості навчальних даних, щоб вона змогла вирізнити один і той самий об'єкт, трансльований на відповідну величину чи повернений у просторі [23].

Капсульні нейронні мережі спрямовані на усунення слабкості сучасних систем машинного навчання. Капсульна нейронна мережа – штучна нейронна мережа, яка створена для покращення моделювання ієрархічних зв'язків між об'єктами різних рівнів [24]. Капсула – це функція, яка намагається передбачити наявність та конкретні параметри об'єкта у певній області. CapsNet працює з цілими наборами нейронів, тобто з векторами. Значення вектора описують характеристики шуканого об'єкта. Капсули нижнього рівня (первинні) описують низькорівневі об'єкти, а капсули верхніх рівнів описують високорівневі об'єкти. Для визначення ймовірності знаходження об'єкта у зображенні капсули використовують нормовану довжину вихідного вектора. Напрямок отриманого вектора визначає просторове розміщення досліджуваного об'єкта відносно інших об'єктів. Це дозволяє правильно розпізнати об'єкт, навіть якщо він був зміщений у зображенні, адже ймовірність результуючого вектора залишиться постійною, а зміниться тільки його напрям. Така поведінка нейронів у була названа еквіваріантністю: ймовірності

виявлення об'єкта залишаються незмінними при зміні його положення.

Результуючий вектор капсули:

$$\mathbf{v}_j = \frac{\|\mathbf{s}_j\|^2}{1 + \|\mathbf{s}_j\|^2} \frac{\mathbf{s}_j}{\|\mathbf{s}_j\|}, \quad (3.3)$$

де \mathbf{s}_j – вхідний вектор капсули j .

Роутинг між капсулами відбувається за допомогою алгоритму динамічної маршрутизації [22]. Алгоритм динамічної маршрутизації між капсулами допомагає відрегулювати вагові коефіцієнти так, щоб капсулі вищого рівня передати більше «правильної» інформації, враховуючи дані, що були передані від інших капсул нижчого рівня. Для цього розраховується середній вектор серед усіх векторів капсул у шарі. Знаходяться відстані від кожного вектора на нижньому рівні до отриманого середнього вектора. На основі розрахованих відстаней обираються ті вектори, які є найближчими до середнього вектора. Оновлюються ваги векторів капсул так, щоб найближчі вектори мали більші значення. Таку процедуру виконують декілька разів, щоб правильно визначити кластер, до якого необхідно спрямувати інформацію про об'єкт.

Після виконання алгоритму динамічної маршрутизації результат передається капсулі верхнього рівня, яка відповідає за конкретну характеристику об'єкта. При навчанні капсульної мережі застосовується гранична функція втрат (англ. margin-loss function):

$$L_k = T_k \max(0, m^+ - \|\mathbf{v}_k\|)^2 + \lambda(1 - T_k) \max(0, \|\mathbf{v}_k\| - m^-)^2, \quad (3.4)$$

де k – клас об'єкта, $T_k = 1$, якщо об'єкт присутній на зображенні, $T_k = 0$, якщо об'єкт відсутній на зображенні, \mathbf{v}_k – результуючий вектор капсули верхнього рівня, $\lambda = 0.5$ – коефіцієнт зниження втрати, $m^+ = 0.9$ – очікуване значення довжини вектора при присутності об'єкта, $m^- = 0.1$ - очікуване значення довжини вектора при відсутності об'єкта. Загальний результат функції втрат для нейромережі обчислюється як різниці між отриманими значеннями та очікуваними.

3.4 Архітектури згорткової та капсульної нейронних мереж, що використовувались у дослідженні

Архітектура згорткової нейронної мережі для нашого дослідження представлена у таблиці 3.1.

Таблиця 3.1 – Архітектура згорткової нейронної мережі

№	Назва шару	Розмір вхідної матриці	Розмір вихідної матриці	Кількість навчальних параметрів
1	Згортковий (Convolutional)	(32x32x3)	(30x30x32)	896
2	Субдискретизуючий (MaxPooling)	(30x30x32)	(15x15x32)	0
3	Згортковий (Convolutional)	(15x15x32)	(13x13x32)	9248
4	Субдискретизуючий (MaxPooling)	(13x13x32)	(6x6x32)	0
5	Згортковий (Convolutional)	(6x6x32)	(4x4x64)	18496
6	Субдискретизуючий (MaxPooling)	(4x4x64)	(2x2x64)	0
7	Шар зведення (Flattening)	(2x2x64)	256	0
8	Повнозв'язний шар (Fully Connected)	256	64	16448
9	Dropout	64	64	0
10	Повнозв'язний шар (Fully Connected)	64	1	65

Загальна кількість параметрів у згортковій нейронній мережі склала 45,153.

Архітектура капсульної нейронної мережі представлена у таблиці 3.2. Загальна кількість параметрів у капсульній нейронній мережі склала 9,582,336.

Таблиця 3.2 – Архітектура капсульної нейронної мережі

№	Назва шару	Розмір вхідної матриці	Розмір вихідної матриці	Кількість навчальних параметрів
1	Згортковий (Convolutional)	(32x32x3)	(24x24x256)	62464
2	Капсули нижчого рівня (первинні капсули) (PrimaryCaps)	(24x24x256)	(2048x8)	5308672
3	Капсули вищого рівня (складні капсули) (RouteCaps)	(2048x8)	(2x16)	528384
4	Маскуючий (Mask)	(2x16)	16	0
5	Повнозв'язний шар (Fully Connected)	16	512	8704
6	Повнозв'язний шар (Fully Connected)	512	1024	525312
7	Повнозв'язний шар (Fully Connected)	1024	3072	3148800
8	Вихідний, що перетворює у необхідну розмірність (Reshape)	3072	32x32x3	0

Для проведення дослідження використовувалася мова програмування Python 3.6. Основне середовище розробки програм – Spyder IDE. Вибір даної мови пов'язаний з тим, що Python володіє зрозумілим синтаксисом. Це дозволяє зосередитися на фактичному дослідженні та використанні інструментів математичного моделювання. Більшість продуктів, бібліотек і модулів для програмування нейронних мереж написані на Python.

У нашому дослідженні використовувались такі бібліотеки та модулі:

- Keras 2.2.4 - відкрита бібліотека для програмування нейронних мереж, написана мовою Python;
- Tensorflow 1.6 - відкрита програмна бібліотека для машинного навчання, що розроблена компанією Google. Застосовується як для досліджень, так і для розробки програмних продуктів;
- Numpy 1.15.4 - підтримує можливість обчислення багатовимірних матриць;
- GDAL 2.4.0 – відкрита бібліотека для читання і запису растрових і векторних гео-просторових форматів даних.

Під час навчання, тестування та перевірки результатів, отриманих від нейронних мереж використовувався ноутбук фірми ASUS. Даний пристрій має 16Gb оперативної пам'яті, центральний процесор Intel Core i7 сьомого покоління (до 3.5 GHz) та одну відеокарту типу Nvidia GeForce 940MX.

3.5 Підготовка навчального, затверджувального та тестового наборів даних

Розраховані індексні зображення (див. п. 2.3) були об'єднані в єдине зображення, що представлене у трьох RGB каналах. Червоний канал відповідає за індексне зображення PSRI, зелений - NDVI блакитний – NDWI. Знаючи точні координати пожежонебезпечної місцевості, із отриманого результуючого зображення було виділено територію, на якій відбулася пожежа «Camp Fire».

Загальна кількість підготовлених зображень становить 40,000 одиниць. 20,000 зображень (50% від загальної кількості) формують пожежонебезпечну територію. Друга половина вибірки складає зображення, що відображають територію, на якій пожежі не було.

Усі зображення були розподілені на три набори даних:

- навчальний набір даних (28000 зображень – 70% від загальної кількості);
- затверджувальний набір даних (6000 зображень – 15% від загальної кількості);
- тестовий набір даних (6000 зображень – 15% від загальної кількості).

Кожне зображення має розмір (32×32) пікселів та три канали RGB. Розмір підвибірки склав 16. На рисунку 3.1 подано приклади зображень, що були вхідними даними для нейромереж.

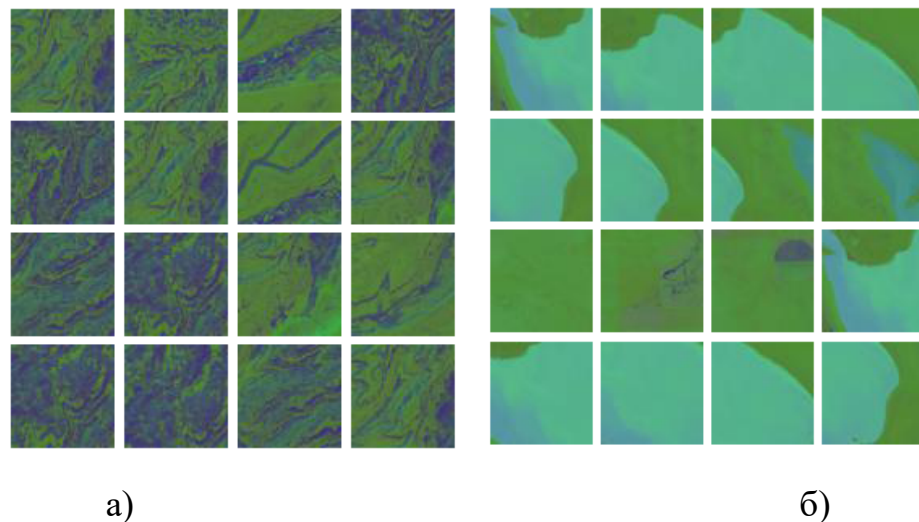


Рисунок 3.1 – Приклади зображень для обробки нейромережею:

а – фрагменти території пожежі, *б* – фрагменти поза територією пожежі

3.6 Результати дослідження та їх обговорення

На вхід згортковій та капсульній нейронним мережам були подані зображення розміром (32×32) з трьома спектральними каналами, а отже в загальному випадку на вхід подається матриця розміром $(32, 32, 3)$. Навчалися обидві нейромережі по 50 епох. На рисунку 3.2 візуалізовано процес розпізнавання зображень згортковою нейронною мережею після першого згорткового шару (обробка однією з карток ознак).

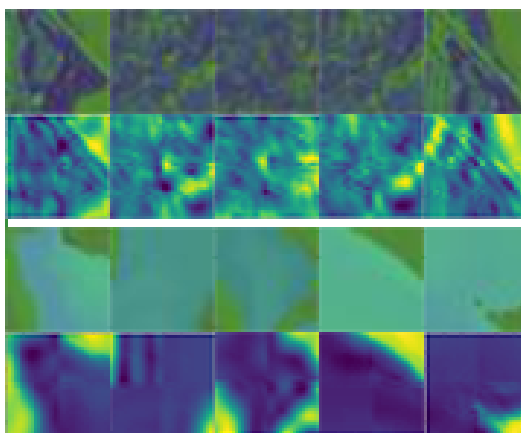


Рисунок 3.2 – Візуалізація процесу розпізнавання згортковою нейронною мережею (п'ять зображень з пожежонебезпечною територією, п'ять зображень з територією без пожежної небезпеки)

На рисунку 3.3 візуалізовано процес розпізнавання зображень капсульною нейронною мережею.

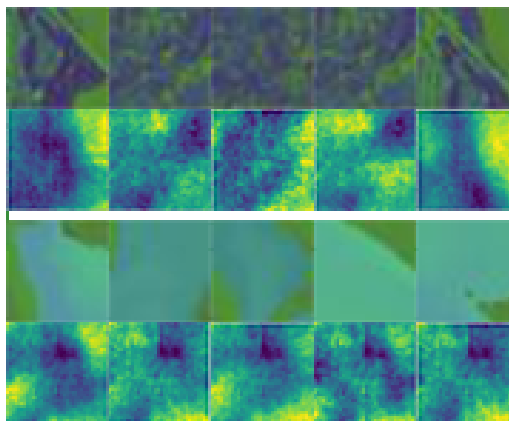


Рисунок 3.3 – Візуалізація процесу розпізнавання капсульною нейронною мережею (п'ять зображень з пожежонебезпечною територією, п'ять зображень з територією без пожежної небезпеки)

Для згорткової мережі середній час проходження однієї епохи складав близько 60 секунд. Загальний час навчання - майже 60 хвилин. Для капсульної нейронної мережі середній час проходження однієї епохи склав близько 1020 секунд. Загальний час навчання – майже 14 годин. Таку велику різницю у часі навчання можна пояснити набагато більшою кількістю навчальних параметрів у капсульній нейронної мережі та обчислювальними потужностями (див. п. 3.4). Побудовані

графіки залежності точності розпізнавання та функцій помилок від кількості епох для навчальної та затверджувальної вибірок зображені на рисунках 3.4-3.7.

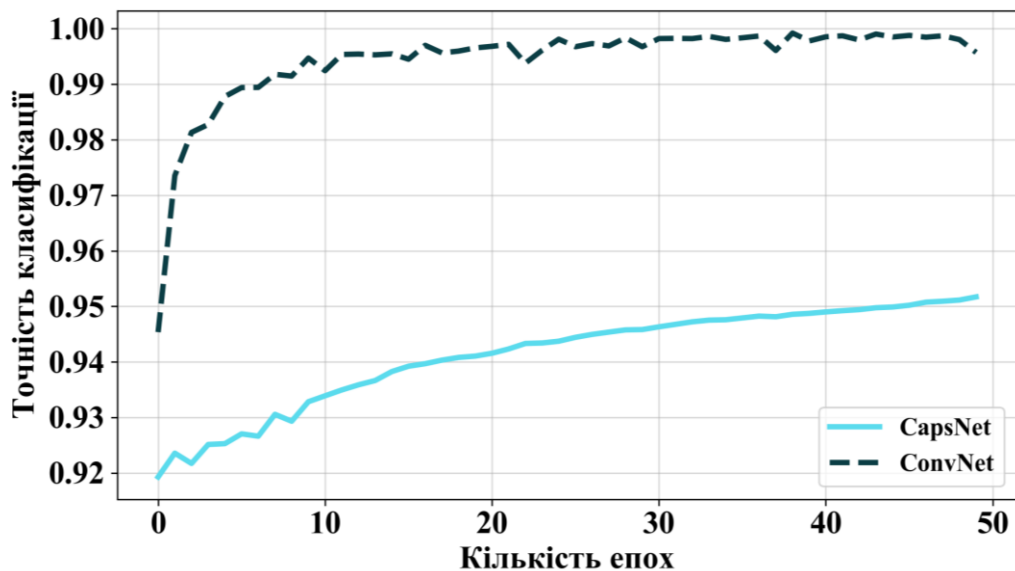


Рисунок 3.4 – Графік залежності точності класифікації на навчальній вибірці від кількості епох навчання для капсульної та згорткової нейронних мереж

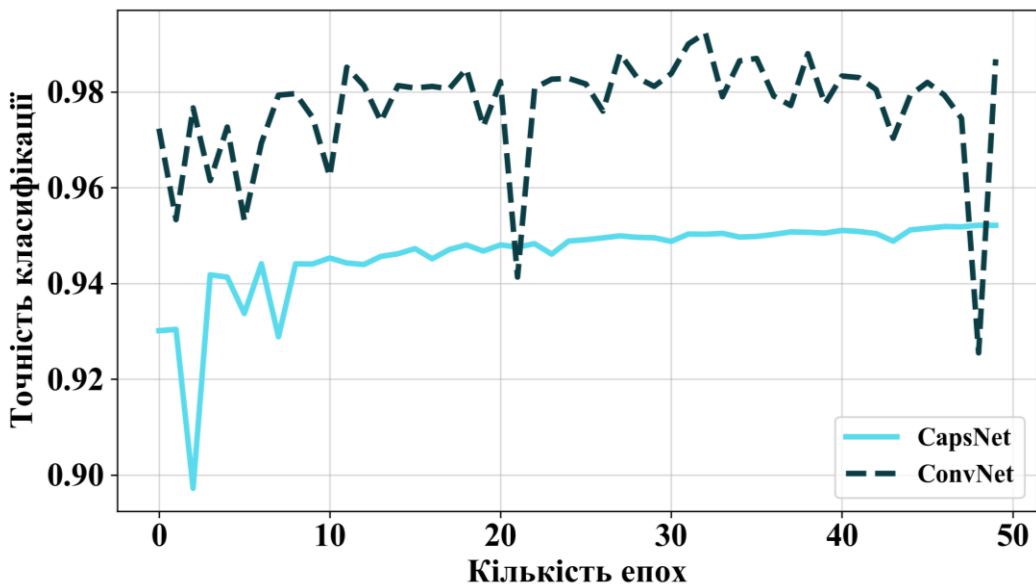


Рисунок 3.5 – Графік залежності точності класифікації на затверджувальній вибірці від кількості епох навчання для капсульної та згорткової нейронних мереж

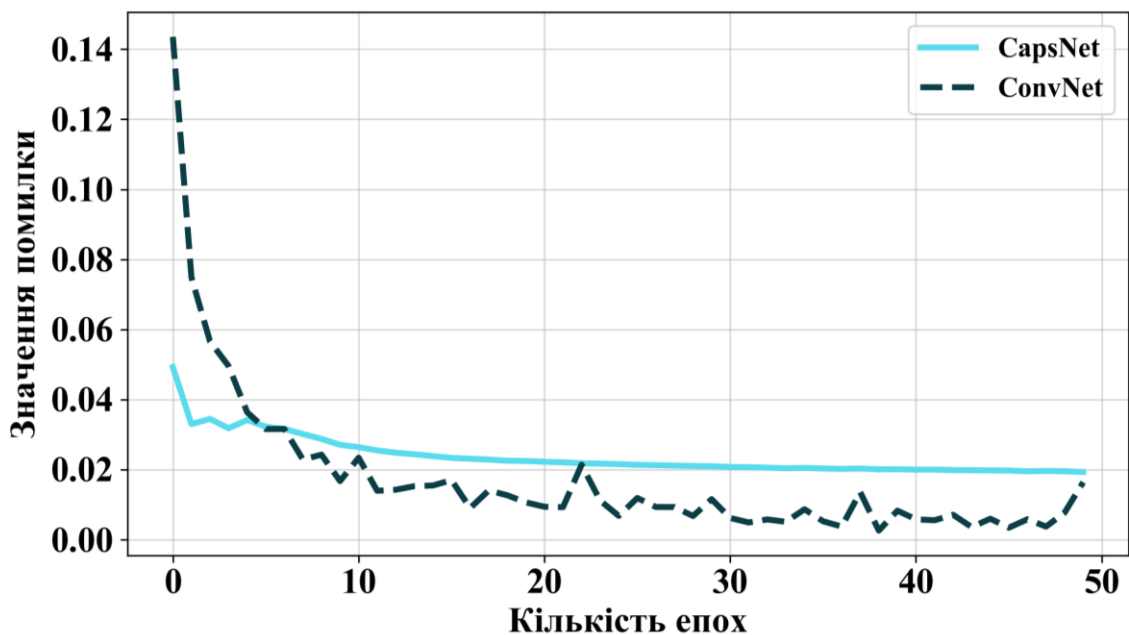


Рисунок 3.6 – Графік залежності значення помилки на навчальній вибірці від кількості епох навчання для капсульної та згорткової нейронних мереж

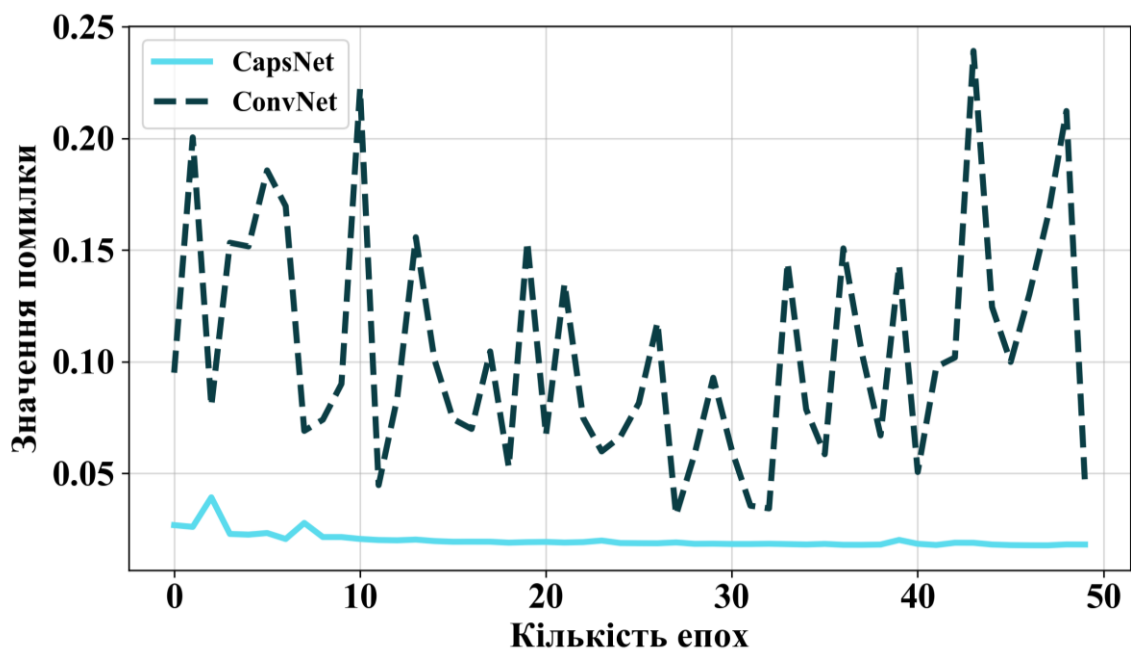


Рисунок 3.7 – Графік залежності значення помилки на затверджувальній вибірці від кількості епох навчання для капсульної та згорткової нейронних мереж

Виявлено, що при використанні не дуже складної архітектури та невеликої кількості гіперпараметрів, згорткові та капсульні нейронні мережі здатні ефективно класифікувати об'єкти, що відносяться до двох класів.

CNN та CapsNet показали майже однакові результати на тестовому наборі даних. Точність класифікації для згорткової нейронної мережі склала 94.27%, а для капсульної – 94.89%.

Аналіз отриманих залежностей (рис. 3.4-3.5) показав, що найбільша точність згорткової нейронної мережі на навчальних даних склала – 99.92% на 39 епосі, найменша – 94.53% на першій епосі, а для капсульної мережі найбільша точність склала 95.17% на 50 епосі, найменша – 94.53% на першій епосі. На затверджувальному наборі даних найбільша точність для ConvNet становила 99.23% на 33 епосі, а найменша – 92.55% на 49 епосі. Для CapsNet ці показники склали відповідно 95.12% на 49 епосі та 89.72% на третій епосі.

На графіках залежності значення помилки від кількості епох (рис. 3.6-3.7) видно, що найбільша помилка на навчальному наборі даних для CNN – 14.35% на першій епосі, найменша – 0.26% на 39 епосі. Для CapsNet найбільша помилка спостерігається на першій епосі – 4.92%, найменша – 1.93% на останній епосі. На затверджувальному наборі даних – для згорткової нейромережі найбільша помилка – це 23.91% на 44 епосі, найменша – 3.12% на 28 епосі. Для капсульної мережі найбільше значення помилки було – 3.92% на третій епосі, найменше – 1.77% на 48 епосі.

Для згорткової нейронної мережі характерне швидке зростання точності на навчальному наборі даних вже при перших епохах. Для капсульної – зростання відбувається плавно та не має великих стрибків протягом усіх епох.

Точність класифікації на затверджувальній вибірці (рис. 3.5) у CNN вища ніж у CapsNet у середньому на 4%, проте видно великі стрибки на деяких епохах, що не характерно для капсульної мережі. Ми припускаємо, що такі результати можуть свідчити про те, що капсульні нейронні мережі навчаються узагальнювати отриману інформацію, більш стійкі до проблеми перенавчання та мають тенденцію до плавного росту точності класифікації, що може бути зумовлено їх архітектурою.

На графіках (див. рис. 3.6) залежності значення помилки від кількості епох (навчальна вибірка) для ConvNet можна спостерігати різкий спад значення помилки до десятої епохи та стрибки на наступних епохах. Загалом, значення помилки у ConvNet дещо нижче, ніж для CapsNet, яка має гладку прямо пропорційну тенденцію до зменшення помилки при збільшенні кількості епох, починаючи з п'ятої. З іншого боку, аналіз значення помилки на затверджувальній вибірці показує, що капсульні нейронні мережі демонструють кращий результат (значення помилки менша приблизно на 12%) за згорткові та не мають великого діапазону стрибків, що характерно для CNN. Припускаємо, що із збільшенням кількості епох, значення помилки для капсульної нейронної мережі буде мати тенденцію до зменшення, хоч і не стрімку.

У [22] дослідниками була протестована капсульна модель на CIFAR10 з сімома різними класами та досягнуте значення помилки склало 10,6%. Розмір зображення становив (24×24) пікселів. Було використано 64 різних типів первинних капсул (PrimaryCaps). Така ж помилка була досягнута і для згорткової мережі [25]. Для нашого випадку значення помилки менше, ніж у наведених вище дослідженнях. Це пов'язано з тим, що ми використовували меншу кількість первинних капсул та класифікація відбувалася за двома класами об'єктів.

ВИСНОВКИ

У роботі вирішена науково-прикладна задача визначення пожежонебезпечних лісових територій з використанням технологій нейронних мереж глибокого навчання на прикладі пожежі «Camp Fire».

1. Розглянуто території, на яких відбулись наймасштабніші лісові пожежі за останній час. Проаналізовано економічні та екологічні збитки від лісових пожеж.

2. Розраховані спектральні індекси та побудовані індексні зображення на основі мультиспектральних зображень Landsat 8 для заданої області дослідження.

3. Встановлено доцільність використання нейронних мереж глибокого навчання як комплексного підходу до вирішення поставленої задачі.

4. Реалізовані згорткова та капсульна нейронні мережі для вирішення задачі знаходження пожежонебезпечних територій.

5. Проведено порівняльний аналіз згорткової та капсульної нейронних мереж. Визначено, що для сформованої вибірки точність класифікації на тестовій вибірці згорткової нейронної мережі склала 94.27%, а капсульної – 94.89%.

6. Виявлено, що при використанні відносно нескладної архітектури, розглянуті нейронні мережі здатні впоратися з поставленим завданням.

7. Встановлено, що капсульні нейронні мережі демонструють менше значення помилки (приблизно на 12%) на затверджувальній вибірці, ніж згорткові нейронні мережі.

8. Отримані результати можна використовувати для визначення пожежонебезпечної території.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. The Department of Forestry and Fire Protection of California: Top 20 Largest California Wildfires [Electronic resource]. - [Cited 2018, 15 December]. - Available from: https://www.fire.ca.gov/communications/downloads/fact_sheets/Top20_Acres.pdf.
2. Hoover K. Wildfire Statistics / К. Hoover // Congressional Research Service [Electronic resource]. – 2018. - [Cited 2018, 15 December]. - Available from: <https://fas.org/sgp/crs/misc/IF10244.pdf>.
3. Insurance Information Institute: Facts & Statistics: Wildfires: Wildfire Losses In The United States, 2008-2017 [Electronic resource]. – 2017. - [Cited 2018, 15 December]. - Available from: <https://www.iii.org/fact-statistic/facts-statistics-wildfires>.
4. U.S. Geological Survey [Electronic resource]. - [Cited 2018, 15 December]. - Available from: <https://earthexplorer.usgs.gov/>
5. Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. Часть 1 / Р.А. Шовенгердт // - М.: Техносфера, 2010. – Т. 560 с.
6. Любченко Н., шифр «Пожежна безпека», Бондарчук В. Нейромережевий метод інтелектуальної обробки мультиспектральних зображень / Н. Любченко, шифр «Пожежна безпека», В. Бондарчук // Сучасні інформаційні системи. – 2017. – Т. 1. – №. 2. – С. 39-44.
7. USGS: Landsat 8 Instruments [Electronic resource]. – [Cited 2018, 15 December]. - Available from: <https://landsat.usgs.gov/landsat-8>.
8. The Department of Forestry and Fire Protection of California: Public Information Map [Electronic resource]. - [Cited 2018, 15 December]. - Available from: http://cdfdata.fire.ca.gov/pub/cdf/images/incidentfile2277_4287.pdf
9. The Department of Forestry and Fire Protection of California: CAMP FIRE Watershed Emergency Response Team Final Report [Electronic resource]. – [Cited 2018,

15 December]. - Available from:
http://cdfdata.fire.ca.gov/pub/cdf/images/incidentfile2277_4330.pdf

10. Бардиш Б. Використання вегетаційних індексів для ідентифікації об'єктів земної поверхні / Б. Бардиш, Х. Бурштинська // Сучасні досягнення геодезичної науки та виробництва. – Національний університет «Львівська політехніка» – №. 2 (28). – 2014. – С. 82-88.

11. McFeeters S. K. Remote Sensing The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features / S. K. McFeeters // International Journal of Remote Sensing. – V. 17. - No 7. – 1996. – P. 1425-1432.

12. Merzlyak M. N. et al. Non-destructive Optical Detection of Pigment Changes During Leaf Senescence and Fruit Ripening. / M. N. Merzlyak, A. A. Gitelson et al. // Physiologia Plantarum. – 1999. – P. 135-141.

13. LeCun Y., Bengio Y., Hinton G. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // nature. – 2015. – V. 521. – No 7553. – P. 436 – 444.

14. Mikolov T. et al. Strategies for training large scale neural network language models / T. Mikolov, A. Deoras et al. // Automatic Speech Recognition and Understanding (ASRU), IEEE Workshop on. – 2011. – P. 196-201.

15. Hinton G. et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups / G. Hinton, L. Deng et al. // IEEE Signal processing magazine. – 2012. – V. 29. – No 6. – P. 82-97.

16. Sainath T. N. et al. Deep convolutional neural networks for LVCSR / T. N. Sainath, A. Mohamed et al. // Acoustics, speech and signal processing (ICASSP), IEEE. – 2013. – P. 8614-8618.

17. Ma J. et al. Deep neural nets as a method for quantitative structure–activity relationships / J. Ma, R. P. Sheridan et al. // Journal of chemical information and modeling. – 2015. – V. 55. – No 2. – P. 263-274.

18. Ciodaro T. et al. Online particle detection with neural networks based on topological calorimetry information / T. Ciodaro, D. Deva et al. // *Journal of physics: conference series*. – IOP Publishing. - 2012. – V. 368. – No 1. – P. 1-12.

19. LeCun Y. et al. Backpropagation applied to handwritten zip code recognition / Y. LeCun, B. Boser et al. // *Neural computation*. – 1989. – V. 1. – No. 4. – P. 541-551.

20. Шифр «Пожежна безпека». Метод зворотного поширення помилки для тренування багатошарових архітектур / Шифр «Пожежна безпека» // *Проблеми інформатизації: Тези доп. VI міжнародної науково-практичної конференції (14-16 листопада 2018р.)*. – Черкаський державний технічний університет, 2018 – С.104-105.

21. Бондарчук В.К. Метод інтелектуальної обробки мультиспектральних зображень / В.К. Бондарчук, шифр «Пожежна безпека» // *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доп. XXV міжнародної науково-практичної конф. MicroCAD-2017 у 4 ч. (17-19 травня 2017 р., м. Харків)*. – Харків: НТУ «ХП», 2017. - С. 65.

22. Sabour S., Frosst N., Hinton G. E. Dynamic routing between capsules / S. Sabour, N. Frosst, G. E. Hinton // *Advances in Neural Information Processing Systems*. – 2017. – P. 3856-3866.

23. Шифр «Пожежна безпека». Капсульні нейронні мережі / Шифр «Пожежна безпека» // *Системи управління навігації та зв'язку*. – Полтавський національний технічний університет імені Юрія Кондратюка. - №. 5 (51). – 2018. – С.132-135.

24. Шифр «Пожежна безпека». CapsNet versus ConvNet / Шифр «Пожежна безпека» // *Інформатика, управління та штучний інтелект: Тези доп. V міжнародної науково-технічної конференції студентів, магістрів та аспірантів (20-22 листопада 2018 року, м. Харків)*. – Харків, НТУ «ХП», 2017. – С. 22-23.

25. Mukhometzianov R., Carrillo, J. CapsNet comparative performance evaluation for image classification / R. Mukhometzianov, J. Carrillo // *Cornell University*. – 2018. – P. 1-14.