

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ
ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ
ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова
праця на правах рукопису

Сторожик Денис Володимирович

УДК 681.5

ДИСЕРТАЦІЯ

УДОСКОНАЛЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ДІАГНОСТУВАННЯ ДОРОЖНЬОГО ПОКРИТТЯ З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

151 Автоматизація та комп'ютерно-інтегровані технології

15 Автоматизація та приладобудування

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Д. В. Сторожик

Науковий керівник Протасов Анатолій Георгійович, доктор педагогічних наук,
професор

АНОТАЦІЯ

Сторожик Д. В. Удосконалення автоматизованої системи діагностування дорожнього покриття з використанням штучного інтелекту. - Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 151 “ Автоматизація та комп’ютерно-інтегровані технології” - Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського» МОН України, Київ, 2024.

В дисертації отримані такі нові наукові результати:

1. удосконалено систему автоматизованого діагностування стану дорожнього покриття на основі мультимодального підходу із комплексуванням зображень шляхом впровадження підсистеми опрацювання (обробки) зображень, застосування методу аугментації теплових зображень та розробленої підсистеми класифікації дефектів дорожнього покриття;
2. вперше запропоновано та розроблено метод мультиспектрального злиття зображень з адаптивним визначенням вагових коефіцієнтів на основі вейвлет перетворення, який полягає в розкладанні початкових зображень на вейвлет коефіцієнти та визначені згортковою нейронною мережею вагових коефіцієнтів, що визначають частку початкового зображення в результуючому, і дозволяє збільшити інформативність результуючого зображення при проведенні діагностування дорожнього покриття;
3. вперше розроблено метод аугментації теплових зображень на базі злиття зображень для задач класифікації, який полягає в поєднанні вейвлет коефіцієнтів отриманих з початкових зображень одного об’єкту та зворотному перетворенню для отримання аугментованого зображення, і дозволяє підвищити ймовірність правильної класифікації дефектів при проведенні діагностування дорожнього покриття.

Практична цінність дисертаційної роботи полягає в тому, що: розроблено

алгоритмічне і програмне забезпечення реалізації підсистеми опрацювання зображень при автоматизованому діагностуванні дефектів дорожнього покриття; розроблено алгоритмічне і програмне забезпечення реалізації методу аугментації що базується на комплексуванні теплових зображень; розроблено алгоритмічне і програмне забезпечення реалізації підсистеми класифікації дефектів дорожнього покриття; сформовано навчальну вибірку розмічених для задач класифікації пар знімків у видимому та інфрачервоних спектрах дефектів дорожнього покриття; розроблено структурну схему автоматизованої системи діагностування дорожнього покриття зі злиттям зображень.

У дисертаційній роботі описано, що дорожньо-транспортна система є однією з ключових ланок критичної інфраструктури будь-якої держави. В Україні цей сектор відіграє особливо важливу роль, адже загальна протяжність автомобільних доріг складає 169,5 тисяч кілометрів.

Через постійний вплив транспортних навантажень дорожнє покриття зазнає природного зношування та руйнації. Для підвищення безпеки на дорогах та зменшення витрат на подальший ремонт проводиться регулярна діагностика стану доріг. За даними досліджень, своєчасне виявлення дефектів та проведення профілактичних ремонтних робіт здатне скоротити витрати на ремонтні заходи більш ніж утричі.

Стан дорожнього покриття безпосередньо впливає на частоту дорожньо-транспортних пригод (ДТП). Якість та параметри доріг належать до об'єктивних чинників, що можуть призводити до аварійних ситуацій. Об'єктивні фактори підлягають контролю та коригуванню, що відкриває можливості для зниження кількості ДТП.

Оскільки автомобільні дороги є однією з найбільших інфраструктурних систем, їх діагностика та обслуговування є дуже ресурсоємними процесами. Витрати на проведення таких робіт, включаючи людські ресурси та час, досягають значних обсягів. Одним з оптимальних шляхів скорочення цих витрат є автоматизація процесів моніторингу стану покриття та ухвалення рішень щодо

необхідних ремонтних робіт.

Важливим бар'єром для масового впровадження автоматизованих систем є висока вартість обладнання, що використовується для точної діагностики дефектів дорожнього полотна. Наприклад, лазерні сканери, які здатні надавати високоточні результати, мають ціну в межах сотень тисяч доларів. У той час, менш дорогі сенсори, які можливо використовувати, такі як камери видимого та інфрачервоного спектрів, не забезпечують достатньої точності при ідентифікації дефектів.

Фахівці зазначають, що зниження точності виявлення дефектів за допомогою камер часто пов'язане з низкою факторів, таких як оптичні спотворення, забруднення поверхні доріг, наявність тіней та погодні умови. У багатьох випадках ці фактори призводять до значних інформаційних перешкод.

Наявні наукові дослідження пропонують використовувати методи обробки зображень для вирішення цієї проблеми. Однак, попри всі зусилля науковців, існуючі автоматизовані системи діагностики стану доріг все ще не досягли необхідного рівня ефективності, що потребує подальшого вдосконалення або розробки нових методик і технологій.

Для вирішення існуючої проблеми запропоновано використати елементи штучного інтелекту, а саме нейромережеві технології. Проблему нестачі обсягу та різноманітності даних у навчальних вибірках нейронних мереж пропонується вирішити за рахунок запропонованого методу аугментації зображень, який заснований на їх комплексуванні з використанням вейвлет-перетворення. Метод має додаткову перевагу — підвищення інформативності результуючих зображень, що позитивно впливає на ефективність навчання нейронних мереж.

Було проведено порівняння нового методу аугментації, розробленого на основі злиття зображень, з іншими наявними підходами для задач класифікації. У результаті проведених досліджень підтверджено, що новий метод не лише покращує точність класифікації, але й демонструє меншу алгоритмічну складність у порівнянні з методом випадкової заміни вейвлет-коефіцієнтів, так, для реалізації запропонованого методу потрібно лише два вихідних зображення,

тоді як метод випадкової заміни вейвлет-коефіцієнтів вимагає шість зображень.

Проблему впливу неінформативних елементів вихідних зображень на результат діагностики, пропонується вирішити запропонованим методом комплексування зображень з адаптивним визначенням вагових коефіцієнтів, що оптимально поєднує зображення двох спектрів з підвищенням інформативності результуючого зображення.

Розроблено нейронну мережу для адаптивного визначення ваг у методі комплексування зображень. Мережа забезпечує автоматичне обчислення вагових коефіцієнтів, які використовуються для підвищення якості результатів. Було обґрунтовано структуру цієї мережі та проведено аналіз її ефективності на валідаційних наборах даних.

Порівняння нового методу комплексування з іншими підходами показало його високу ефективність у створенні більш інформативних зображень. Для оцінки інформативності використовувався метод на основі ентропії.

Було зібрано та оброблено навчальний набір даних, що включає зображення видимого та інфрачервоного спектрів, які містять дефекти дорожнього покриття. Це дало можливість емпірично розробити підсистему для обробки зображень, отриманих з камер лабораторій для діагностики дорожнього покриття.

Запропоновано структурну схему автоматизованої системи діагностики та ремонту доріг, що використовує елементи штучного інтелекту. Ця система реалізовує повний цикл втручання, від діагностики до ремонту. Так, на основі отриманої інформації про стан при обробці зображень, підсистема підтримки прийняття рішень надає рекомендації щодо організації необхідної технології ремонту. Також були обґрунтовані процеси попередньої обробки даних для обох спектрів, що дозволило знизити інформаційні завади та підвищити точність діагностики.

Було навчено модель для автоматизованого визначення типів дефектів дорожнього покриття, яка продемонструвала значно кращу точність у порівнянні з існуючими рішеннями. Удосконалення відбулося завдяки впровадженню

запропонованих методів злиття зображень та методу аугментації. Це дозволило підвищити ймовірність коректного визначення типу дефекту на 0,66%, забезпечуючи більш надійні результати навіть при обмеженому обсязі навчальних даних. Запропонований підхід також зменшив необхідну кількість навчальних зображень до 3 тисяч, що еквівалентно 13,5 тисячам у існуючих рішеннях. Така оптимізація значно спрощує навчання моделі та підвищує її ефективність у задачах діагностування стану дорожнього покриття. А також, було обрано нейронну мережу для сегментації зображень, яка дозволяє точно визначати площу дефектів, забезпечуючи високу ефективність при використанні текстових підказок.

Розроблено алгоритми для автоматизованого руху лабораторій та підсистеми обробки зображень. Реалізовано обробку зображень та прийняття рішень на базі хмарних сервісів, що скорочує час на впровадження нових версій та прискорило процес моделювання. Підсистема підтримки прийняття рішень надає рекомендації щодо проведення необхідних ремонтних робіт, оптимізуючи використання ресурсів та забезпечуючи своєчасне усунення виявлених дефектів.

За результатами імітаційного моделювання підтверджено працездатність запропонованої системи, яка демонструє високі показники точності при визначенні типів дефектів дорожнього покриття та розрахунку їх площі.

Ключові слова: діагностика дорожнього покриття, комплексування зображень, аугментація, критерій якості, обробка сигналів, вейвлет-перетворення, методи неруйнівного контролю, нейронні мережи, автоматизація, експертна система, математична модель, прийняття рішень, машинне навчання, ідентифікація.

ANNOTATION

Storozhyk D.V. Improvement of the Automated Road Surface Diagnostics System Using Artificial Intelligence. - A qualification scientific work as a manuscript.

Dissertation for the degree of Doctor of Philosophy in Specialty 151 "Automation and Computer-Integrated Technologies." - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Ministry of Education and Science of Ukraine, Kyiv, 2024.

The dissertation presents the following new scientific results:

1. The automated road surface diagnostics system based on a multimodal approach with image integration was improved through the implementation of an image processing subsystem, the use of a thermal image augmentation method, and the developed subsystem for classifying road surface defects.
2. A method for multispectral image fusion with adaptive determination of weighting coefficients based on wavelet transformation was proposed and developed for the first time. The method involves decomposing initial images into wavelet coefficients and determining the weighting coefficients using a convolutional neural network. These coefficients define the contribution of the initial image to the resulting image and allow for an increase in the informativeness of the resulting image during road surface diagnostics.
3. A thermal image augmentation method based on image fusion for classification tasks was developed for the first time. This method combines wavelet coefficients obtained from initial images of the same object and applies inverse transformation to generate augmented images, which increases the probability of correct defect classification during road surface diagnostics.

The practical value of the dissertation lies in the following developments: algorithmic and software tools for implementing an image processing subsystem in automated road defect diagnostics; algorithmic and software tools for implementing the augmentation method based on thermal image fusion; algorithmic and software tools for implementing the subsystem for classifying road surface defects; a labeled dataset of paired visible and infrared spectrum images of road surface defects for

classification tasks; a structural scheme of the automated road surface diagnostics system using image fusion.

The dissertation describes that the road transport system is one of the key components of critical infrastructure in any country. In Ukraine, this sector plays a particularly important role, with a total length of 169,500 kilometers of roads.

Due to the constant influence of transport loads, road surfaces undergo natural wear and tear and degradation. To improve road safety and reduce repair costs, regular diagnostics of road conditions are conducted. Studies show that timely detection of defects and preventive repair work can reduce repair costs by more than three times.

The condition of road surfaces directly affects the frequency of road traffic accidents (RTAs). Road quality and parameters are objective factors that can lead to accidents. These factors can be controlled and adjusted, offering opportunities to reduce the number of RTAs.

Since roads are one of the largest infrastructural systems, their diagnostics and maintenance are highly resource-intensive processes. The costs of such work, including human resources and time, are significant. One optimal way to reduce these costs is to automate the monitoring of road conditions and decision-making regarding necessary repairs.

A major barrier to the widespread implementation of automated systems is the high cost of equipment used for accurate diagnostics of road defects. For instance, laser scanners that provide highly accurate results cost hundreds of thousands of dollars. Meanwhile, less expensive sensors, such as visible and infrared cameras, do not ensure sufficient accuracy in defect identification.

Experts note that the reduced accuracy of defect detection using cameras is often related to factors such as optical distortions, road surface contamination, the presence of shadows, and weather conditions. In many cases, these factors lead to significant informational noise.

Existing scientific research suggests using image processing methods to address this issue. However, despite the efforts of researchers, existing automated road condition diagnostics systems have not yet achieved the necessary level of efficiency,

requiring further improvement or the development of new methods and technologies.

To address this issue, it is proposed to use elements of artificial intelligence, specifically neural network technologies. The problem of insufficient volume and diversity of data in the training sets of neural networks is proposed to be solved by the developed image augmentation method, which is based on image fusion using wavelet transformation. This method has an additional advantage—enhancing the informativeness of the resulting images, which positively impacts the efficiency of neural network training.

A comparison of the newly developed augmentation method, based on image fusion, with other existing approaches for classification tasks has been conducted. The results of the study confirm that the new method not only improves classification accuracy but also demonstrates lower algorithmic complexity compared to the method of random wavelet coefficient replacement. Specifically, the proposed method requires only two initial images, whereas the random coefficient replacement method requires six.

The problem of the impact of non-informative elements of the initial images on the diagnostic results is proposed to be solved by the developed image fusion method with adaptive determination of weighting coefficients, which optimally combines images from two spectra, increasing the informativeness of the resulting image.

A neural network for adaptive weighting in the image fusion method has been developed. The network provides automatic calculation of weighting coefficients, which are used to improve the quality of results. The structure of this network has been substantiated, and its effectiveness has been analyzed using validation data sets.

A comparison of the new fusion method with other approaches demonstrated its high efficiency in creating more informative images. Entropy-based methods were used to assess informativeness.

A training dataset, including visible and infrared spectrum images containing road surface defects, was collected and processed. This enabled the empirical development of a subsystem for processing images obtained from cameras in road diagnostics laboratories.

A structural scheme of an automated road diagnostics and repair system using artificial intelligence was proposed. This system implements a complete intervention cycle, from diagnostics to repair. Based on the processed image data, the decision-support subsystem provides recommendations for organizing the necessary repair technologies. Data preprocessing processes for both spectra were also substantiated, reducing informational noise and improving diagnostic accuracy.

A model for automated defect type detection was trained, demonstrating significantly better accuracy compared to existing solutions. Improvements were achieved through the implementation of the proposed image fusion and augmentation methods. This approach increased the probability of correct defect type identification by 0.66%, ensuring more reliable results even with limited training data. The proposed approach also reduced the required number of training images to 3,000, equivalent to 13,500 in existing solutions. This optimization simplifies model training and increases its efficiency in road surface diagnostics tasks.

A neural network for image segmentation was chosen, allowing accurate determination of defect areas and ensuring high efficiency using textual prompts.

Algorithms for automated laboratory movement and image processing subsystems were developed. Image processing and decision-making were implemented using cloud services, reducing the time required for new version deployment and accelerating the modeling process. The decision-support subsystem provides recommendations for necessary repair work, optimizing resource use and ensuring timely elimination of detected defects.

Simulation modeling confirmed the operability of the proposed system, which demonstrates high accuracy in identifying road surface defect types and calculating their areas.

Keywords: road surface diagnostics, image fusion, augmentation, quality criterion, signal processing, wavelet transformation, non-destructive testing methods, neural networks, automation, expert system, mathematical model, decision-making, machine learning, identification.

Список публікацій здобувача:

Статті у вітчизняних наукових фахових виданнях

1. Сторожик Д. Комплексування мультиспектральних зображень, як метод підвищення їх інформативності при бінарній сегментації / Д. Сторожик, А. Протасов, О. Муравйов, В. Баженов, Г. Богдан // Наукові вісті КПП. - 2020. - № 2. - С. 82–87. - DOI: <https://doi.org/10.20535/kpi-sn.2020.2.197955>.

2. Сторожик Д. Автоматизація процесу теплового неруйнівного контролю шляхом застосування методу комплексування термограм / Д. Сторожик, А. Протасов, О. Муравйов, В. Петрик, Д. Петренко // Технічна діагностика і неруйнівний контроль. - 2022. - № 2. - С. 20–23. - DOI: <https://doi.org/10.37434/tdnk2022.02.03>.

3. Сторожик Д., Протасов А. Технології опрацювання зображень на основі комплексування даних (Огляд) // Технічна діагностика і неруйнівний контроль. - 2022. - № 4. - С. 17–26. - DOI: <https://doi.org/10.37434/tdnk2022.04.03>.

4. Сторожик Д., Протасов А. Технології автоматизації діагностики дорожнього покриття // Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки. - 2023. - Том 34 (73) № 2. - С. 219-227. - DOI: <https://doi.org/10.32782/2663-5941/2023.2.1/35>.

5. Сторожик Д., Протасов А. Автоматизована система діагностики дорожнього покриття з комплексуванням зображень // Технічна діагностика та неруйнівний контроль. - 2024. - № 3. - С. 39-44 - DOI: <https://doi.org/10.37434/tdnk2024.03.06>.

Статті у наукових фахових виданнях, що входять до наукометричної бази Scopus

6. Storozhyk D., Protasov A., Kuts Y., Muraviov O., Lysenko Y., Mirchev Y.. Enhancing neural network efficiency in automated image analysis for thermal nondestructive testing // Journal of Theoretical and Applied Mechanics, Sofia. - 2024. - Vol. 54. - Pp. 242-252. - DOI: <https://doi.org/10.55787/jtams.24.54.2.242>.

Наукові праці, які засвідчують апробацію матеріалів дисертації

7. Сторожик Д., Муравйов О. Комплексування зображень, як спосіб

покращення якості бінарної сегментації // XV Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «Ефективність інженерних рішень у приладобудуванні» : зб. пр. конф., м. Київ, 10–11 грудня 2019. - С. 290–293.

8. Сторожик Д., Муравйов О. Застосування мультиспектрального комплексування зображень у неруйнівному контролі // XIII Науково-практична конференція студентів, аспірантів та молодих вчених «Погляд у майбутнє приладобудування», м. Київ, 13–14 травня 2020. - С. 323–326.

9. Сторожик Д. Комплексування зображень ультразвукової та комп'ютерної томографії // XIII Науково-практична конференція студентів, аспірантів та молодих вчених «Погляд у майбутнє приладобудування», м. Київ, 13–14 травня 2020. - С. 327–329.

10. Storozhyk D. An overview of solutions for the use of image fusion in nondestructive testing // Збірник праць XV Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених “Погляд у майбутнє приладобудування”, Київ. - 4-15 червня 2022. - С. 184–187.

11. Сторожик, Д. В. Комплексування зображень, як спосіб покращення точності класифікації при тепловому неруйнівному контролі // XXII Міжнародна науково-технічна конференція "Приладобудування: стан і перспективи", Київ. - 16–17 травня 2023. - С. 258.

12. Storozhyk D., Protasov A., Lysenko I. Image fusion as an augmentation method to improv the classification efficiency of thermal image //2024 IEEE 42th International Conference on Electronics and Nanotechnology IEEE ELNANO – 2024. Kyiv, Ukraine, May 13-15, 2024.

ЗМІСТ

ВСТУП	16
РОЗДІЛ 1. Об'єкт та методи діагностики.....	25
1.1. Дорожній одяг	25
1.2. Дефекти дорожнього покриття.....	26
1.3. Автоматизовані системи діагностики дорожнього покриття	28
1.4. Методи діагностики стану дорожнього покриття	31
1.4.1. Методи дистанційного зондування	32
1.4.2. Штучні нейронні мережі.....	38
1.4.3. Комбіновані методи з використанням мультисенсорних систем ..	40
ВИСНОВКИ ДО РОЗДІЛУ 1	44
Список використаних джерел до розділу 1	45
РОЗДІЛ 2. Технології діагностування стану дорожнього покриття в автоматизованих системах	50
2.1. Комплексування зображень дефектів дорожнього покриття	50
2.1.1. Методи комплексування зображень	51
2.2. Використання нейронних мереж в системах діагностики.....	58
2.2.1. Модель штучного нейрона	58
2.2.2. Згорткові нейронні мережі	61
2.2.2.1. Нейронна мережа AlexNet	62
2.2.2.2. Нейронна мережа VGGNet	63
2.2.2.3. Нейронна мережа ResNet.....	65
2.2.2.4. Нейронна мережа EfficientNet	66
2.3. Методи штучного збільшення набору даних для нейронних мереж	69

2.4. Методи оцінки нейронних моделей	72
ВИСНОВКИ ДО РОЗДІЛУ 2	74
Список використаних джерел до розділу 2	75
РОЗДІЛ 3. Розробка та оцінка методів аугментацій та комплексування зображень	80
3.1. Розробка методу аугментації на базі комплексування зображень з вейвлет перетворенням	80
3.2. Оцінка методу аугментації на базі комплексування зображень з вейвлет перетворенням	81
3.2.1. Опис даних для проведення експерименту	81
3.2.2. Реалізація експерименту	82
3.2.3. Аналіз результатів дослідження	83
3.3. Розробка методу комплексування зображень з адаптивним визначенням ваг	86
3.4. Розробка нейронної мережі для визначення ваг	88
3.4. Оцінка методу комплексування зображень з адаптивним визначенням ваг	91
ВИСНОВКИ ДО РОЗДІЛУ 3	95
Список використаних джерел до розділу 3	96
РОЗДІЛ 4. Розробка автоматизованої системи діагностики та ремонту дорожнього покриття	97
4.1. Формування навчальної вибірки зображень дефектів у видимому та інфрачервоному спектрах	97
4.2. Розробка структурної схеми автоматизованої системи діагностики та ремонту дорожнього покриття	98
4.3. Розробка підсистеми опрацювання зображень	101

4.3.1. Розробка процесу попередньої обробки зображень з комплексуванням зображень	101
4.3.2. Підбір та навчання нейронної мережі для класифікації типів дефектів.....	106
4.3.3. Обґрунтування вибору моделі для сегментації зображень.....	112
4.4. Розробка алгоритмів системи.....	113
4.4.1. Розробка алгоритму руху ходової лабораторії.....	113
4.4.1. Розробка алгоритму підсистеми обробки зображень.....	114
ВИСНОВКИ ДО РОЗДІЛУ 4	115
РОЗДІЛ 5. Експериментальна перевірка роботи системи	117
5.1. Формування тестової вибірки для імітаційного моделювання	117
5.2. Реалізація конвеєра обробки зображень та системи прийняття рішень	117
5.3. Аналіз результатів імітаційного моделювання автоматизованої системи діагностики та ремонту дорожнього покриття.....	118
5.3.1. Результати імітаційного моделювання на наборі з низьким рівнем дефектів.....	118
5.3.2. Результати імітаційного моделювання на наборі з середнім рівнем наявності дефектів.....	121
5.3.3. Результати імітаційного моделювання на наборі з високим рівнем наявності дефектів.....	123
ВИСНОВКИ ДО РОЗДІЛУ 5	128
Список використаних джерел до розділу 5	129
ВИСНОВКИ.....	130
ДОДАТКИ.....	133

ВСТУП

Дорожньо-транспортна система є одним з основних об'єктів критичної інфраструктури будь-якої країни. До прикладу в Україні це найбільший об'єкт критичної інфраструктури з протяжністю 169,5 тис. кілометрів.

Під впливом великого навантаження дорожнє покриття схильне до зносу та старіння. Задля забезпечення безпеки на дорозі, а також для зниження витрат проводиться діагностика дорожнього покриття. Вчасна діагностика та проведення ремонтних робіт здатна зменшити витрати більше ніж втричі [1].

Не слід применшувати вплив стану дорожнього покриття на кількість дорожньо-транспортних пригод (ДТП). Параметри й стан доріг є однією з об'єктивних причин виникнення ДТП. На відміну від суб'єктивних причин, об'єктивні піддаються контролю, що дозволяє зменшувати кількість ДТП [2].

Однак попри велику економію коштів при ремонті дорожнього покриття діагностика ще не проводиться повсюдно, через складність та довготривалість таких перевірок. Цей недолік може прибрати поширення автоматизованих методів збору та обробки інформації.

Актуальність теми.

Діагностика та ремонт такого великого об'єкта інфраструктури як дорожнє покриття є надзвичайно затратним. Людино ресурси та час необхідний на збір та обробку даних про стан та ремонт дорожнього покриття набуває величезних показників. Одним зі способів зменшення витрат на діагностику та ремонт є автоматизація процесів аналізу даних та прийняття рішення щодо необхідного втручання.

Більшість дослідників сходяться в думці, що головною звадою впровадження таких автоматизованих систем повсюдно є велика ціна обладнання та точність визначення дефектності дорожнього покриття. Так, ціна високоточних лазерних сканерів сягає до сотень тисяч доларів, а менш дешеві датчики (камери інфрачервоного та видимого спектра) не дають бажаної точності при виявленні дефектів.

Головними причинами поганих результатів в точності виявлення дефектів

при використанні камер інфрачервоного та видимого спектра дослідники вважають наявність інформативних завад, віднести до яких можна, оптичні ілюзії, засміченість дорожнього покриття, наявність тіней, вплив поганих погодних умов тощо.

Дослідження науковців, таких як Oliveira H., Correia P., Zalama E., намагалися розв'язати цю проблему використовуючи класичні методи обробки зображень. Значна кількість дослідників (Ouyang A., Luo C., Zhou C., Chen Z., Yusof, N. A. M.) спрямовували свої дослідження в бік використання сучасних технологій на базі нейромереж.

Деякі автори, такі як Su Y., Kang S., Chang J., Hsieh S., Bystrov A., Chen, Cheng, Sindhu Chandra, спрямували дослідження у бік поєднання інформації добутої з датчиків різної модальності.

Однак, хоча дані дослідження наблизили розв'язання цієї задачі, наявні автоматизовані системи діагностики та ремонту не досягли необхідної ефективності, а значить, потребують або їх удосконалення або розроблення нових.

Зв'язок роботи з науковими програмами, планами, темами

Робота відповідає основним напрямкам діяльності Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» і мала безпосередній зв'язок з проєктами, що виконувались на кафедрі автоматизації та систем неруйнівного контролю. Автор був співвиконавцем ініціативного проєкту «Автоматизація процедур теплового неруйнівного контролю з використанням технологій штучного інтелекту. Державний реєстраційний номер: 0123U103688, 2023 р.

Результати дисертаційної роботи автора були впроваджені у виробничий процес науково виробничої фірми «Діагностичні прилади» та навчальний процес кафедри «Автоматизації та систем неруйнівного контролю» Приладобудівного факультету.

Мета і завдання дослідження.

Метою дослідження є підвищення якості та зменшення трудомісткості діагностування стану дорожнього покриття шляхом розробки та застосування інформаційного, методичного, алгоритмічного та програмного забезпечення.

Досягнення поставленої мети передбачає розв'язання таких наступних задач:

- Проаналізувати сучасний стан розвитку методів автоматизованого діагностування стану дорожнього покриття, визначити основні напрями наукових досліджень та їх актуальність, сформулювати основні проблеми у цій галузі та обґрунтувати методи для їх вирішення.
- Зібрати зображення дефектів дорожнього покриття у двох спектральних діапазонах, розробити та обґрунтувати процес обробки цих зображень з використанням нейронних мереж.
- Розробити та дослідити метод злиття зображень на основі вейвлет перетворення з адаптивним визначенням вагових коефіцієнтів.
- Розробити та дослідити метод аугментації зображень на основі злиття зображень для нейромережевих класифікаторів.
- Удосконалити систему автоматизованого діагностування стану дорожнього покриття шляхом впровадження підсистеми опрацювання (обробки) зображень та застосування методів аугментації.
- Дослідити підвищення якості та зменшення трудомісткості діагностування стану дорожнього покриття удосконаленої автоматизованої системи діагностування стану дорожнього покриття.

Об'єктом дослідження є процес автоматизованого керування діагностикою поточного стану дорожнього покриття.

Предметом дослідження є методи та засоби підвищення інформативності зображень дефектів в процесі діагностування стану дорожнього покриття.

Методи дослідження:

- методи аналізу для обґрунтування актуальності, формулювання мети дослідження, визначення основних задач;
- методи аналізу що базуються на теорії ймовірностей та математичної

статистики;

- теорії методів теплового неруйнівного контролю;
- методах обробки зображень та опрацювання експериментальних даних;
- методах злиття (комплексуювання) зображень;
- методах аугментації зображень для використання в нейромережах.

Наукова новизна отриманих результатів:

- удосконалено систему автоматизованого діагностування стану дорожнього покриття на основі мультимодального підходу із комплексуванням зображень шляхом впровадження підсистеми опрацювання (обробки) зображень, застосування методу аугментації теплових зображень та розробленої підсистеми класифікації дефектів дорожнього покриття.;
- вперше запропоновано та розроблено метод мультиспектрального злиття зображень з адаптивним визначенням вагових коефіцієнтів на основі вейвлет перетворення, який полягає в розкладанні початкових зображень на вейвлет коефіцієнти та визначені згортковою нейронною мережею вагових коефіцієнтів, що визначають частку початкового зображення в результуючому, і дозволяє збільшити інформативність результуючого зображення при проведенні діагностування дорожнього покриття;
- вперше розроблено метод аугментації теплових зображень на базі злиття зображень для задач класифікації, який полягає в поєднанні вейвлет коефіцієнтів отриманих з початкових зображень одного об'єкту та зворотному перетворенню для отримання аугментованого зображення, і дозволяє підвищити ймовірність правильної класифікації дефектів при проведенні діагностування дорожнього покриття.

Практичне значення отриманих результатів:

- розроблено алгоритмічне і програмне забезпечення реалізації підсистеми опрацювання зображень при автоматизованому діагностуванні дефектів дорожнього покриття;

- розроблено алгоритмічне і програмне забезпечення реалізації методу аугментації що базується на комплексуванні теплових зображень;
- розроблено алгоритмічне і програмне забезпечення реалізації підсистеми класифікації дефектів дорожнього покриття;
- сформовано навчальну вибірку розмічених для задач класифікації пар знімків у видимому та інфрачервоних спектрах дефектів дорожнього покриття;
- розроблено структурну схему автоматизованої системи діагностування дорожнього покриття зі злиттям зображень.

Особистий внесок здобувача.

Дисертація є завершеною науковою працею автора. Основні наукові ідеї, результати теоретичних та експериментальних досліджень розроблені і сформульовані автором особисто.

Автором було запропоновано методу комплексування зображень з адаптивним визначенням ваг. Розроблено архітектуру моделі для адаптивного визначення ваг при комплексуванні. Проведено дослідження ефективності запропонованого методу в порівнянні з наявними аналогами. Запропоновано метод аугментації зображень для збільшення навчальної вибірки теплових зображень при використанні для навчання нейронних мереж для задач класифікації. Проведено дослідження та аналіз запропонованого методу розширення навчальної вибірки. Розроблено структурну схему автоматизованої системи діагностики та ремонту дорожнього покриття з використанням елементів штучного інтелекту. Розроблено процес обробки зображень в системі, з емпіричним підбором архітектури нейронної мережі для класифікації зображень. Проведено дослідження впливу розроблених методів і процесу обробки зображень на точність виявлення типу дефекту дорожнього покриття згортковою нейронною мережею. Проведено імітаційне моделювання підсистем обробки зображень та підтримки прийняття рішень для дослідження ефективності запропонованої автоматизованої системи.

Апробація результатів дисертації.

Результати дисертаційного дослідження оприлюднено й обговорено на наукових форумах – міжнародних, всеукраїнських і регіональних конференціях, семінарах.

1. Сторожик Д., Муравйов О. Комплексування зображень, як спосіб покращення якості бінарної сегментації // XV Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «Ефективність інженерних рішень у приладобудуванні» : зб. пр. конф., м. Київ, 10–11 грудня 2019. - С. 290–293. *Здобувачем проаналізовано результати бінарної сегментації при використанні методів комплексування зображень.*
2. Сторожик Д., Муравйов О. Застосування мультиспектрального комплексування зображень у неруйнівному контролі // XIII Науково-практична конференція студентів, аспірантів та молодих вчених «Погляд у майбутнє приладобудування», м. Київ, 13–14 травня 2020. - С. 323–326. *Здобувачем проаналізовано сучасний стан використання методів комплексування зображень при діагностиці.*
3. Сторожик Д. Комплексування зображень ультразвукової та комп'ютерної томографії // XIII Науково-практична конференція студентів, аспірантів та молодих вчених «Погляд у майбутнє приладобудування», м. Київ, 13–14 травня 2020. - С. 327–329. *Здобувачем проаналізовано сучасний стан використання методів комплексування зображень при ультразвуковій діагностиці та комп'ютерній томографії.*
4. Storozhyk D. An overview of solutions for the use of image fusion in nondestructive testing // Збірник праць XV Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених “Погляд у майбутнє приладобудування”, Київ. - 4-15 червня 2022. - С. 184–187. *Здобувачем проаналізовано сучасний стан використання методів комплексування зображень при мультиспектральній*

діагностиці.

5. Сторожик, Д. В. Комплексування зображень, як спосіб покращення точності класифікації при тепловому неруйнівному контролі // XXII Міжнародна науково-технічна конференція "Приладобудування: стан і перспективи", Київ. - 16–17 травня 2023. - С. 258. *Здобувачем проаналізовано результати класифікації підповерхневих дефектів при використанні комплексування зображень.*
6. Storozhyk D., Protasov A., Lysenko I. Image fusion as an augmentation method to improv the classification efficiency of thermal image //2024 IEEE 42th International Conference on Electronics and Nanotechnology IEEE ELNANO – 2024. Kyiv, Ukraine, May 13-15, 2024. *Здобувачем проаналізовано результати ефективності використання методу аугментації на основі комплексування зображень при діагностиці з використанням тепловізорів.*

Публікації.

1. Сторожик Д. Комплексування мультиспектральних зображень, як метод підвищення їх інформативності при бінарній сегментації / Д. Сторожик, А. Протасов, О. Муравйов, В. Баженов, Г. Богдан // Наукові вісті КПП. - 2020. - № 2. - С. 82–87. - DOI: <https://doi.org/10.20535/kpi-sn.2020.2.197955>. *Здобувачем проведено дослідження точності сегментації при використанні методів комплексування зображень.*
2. Сторожик Д. Автоматизація процесу теплового неруйнівного контролю шляхом застосування методу комплексування термограм / Д. Сторожик, А. Протасов, О. Муравйов, В. Петрик, Д. Петренко // Технічна діагностика і неруйнівний контроль. - 2022. - № 2. - С. 20–23. - DOI: <https://doi.org/10.37434/tdnk2022.02.03>. *Здобувачем проведено дослідження впливу використання комплексування зображень при класифікації дефектів.*
3. Сторожик Д., Протасов А. Технології опрацювання зображень на основі комплексування даних (Огляд) // Технічна діагностика і неруйнівний

- контроль. - 2022. - № 4. - С. 17–26. - DOI: <https://doi.org/10.37434/tdnk2022.04.03>. *Здобувачем проаналізовано сучасний стан технологій опрацювання зображень на основі комплексування даних.*
4. Сторожик Д., Протасов А. Технології автоматизації діагностики дорожнього покриття // Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки. - 2023. - Том 34 (73) № 2. - С. 219-227. - DOI: <https://doi.org/10.32782/2663-5941/2023.2.1/35>. *Здобувачем проаналізовано сучасний стан технологій автоматизації діагностики дорожнього покриття.*
5. Сторожик Д., Протасов А. Автоматизована система діагностики дорожнього покриття з комплексуванням зображень // Технічна діагностика та неруйнівний контроль. - 2024. - № 3. - С. 39-44 - DOI: <https://doi.org/10.37434/tdnk2024.03.06>. *Здобувачем розроблено структурну схему автоматизована система діагностики дорожнього покриття з комплексуванням зображень.*
6. Storozhyk D., Protasov A., Kuts Y., Muraviov O., Lysenko Y., Mirchev Y.. Enhancing neural network efficiency in automated image analysis for thermal nondestructive testing // Journal of Theoretical and Applied Mechanics, Sofia. - 2024. - Vol. 54. - Pp. 242-252. - DOI: <https://doi.org/10.55787/jtams.24.54.2.242>. *Здобувачем проведено дослідження і проаналізовано результати класифікації дефектів при використанні методу аугментації на базі комплексування зображень.*

Основні результати дисертаційної роботи опубліковано у 12 наукових роботах; 6 статей у провідних фахових виданнях, з яких 2 – у виданнях , що входять до міжнародних наукометричних баз даних, 1 – у виданнях, що входять до наукометричної бази Scopus;

Структура роботи

Дисертація складається з анотації, вступу, п'яти розділів, висновків до

кожного з розділів, списку використаних джерел до кожного з розділів. Загальний обсяг дисертації становить сторінок 136, у тому числі сторінок 118 основного тексту, 43 рисунки та 6 таблиць.

РОЗДІЛ 1. Об'єкт та методи діагностики

1.1. Дорожній одяг

Дорожній одяг (рис. 1.1) — одно- або багат шарова конструкція що розподіляє навантаження спричинена автотранспортом на ґрунт. За наявності кількох шарів складається з покриття, основи та при необхідності з додаткової основи. Дорожній одяг поділяють на жорсткий та нежорсткий [3].

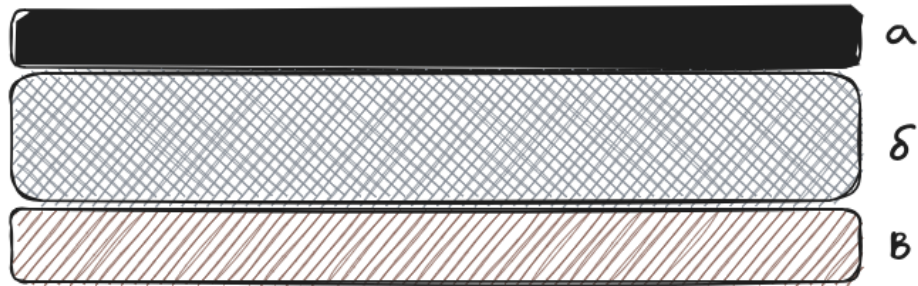


Рис. 1.1 Загальна схема дорожнього одягу: а — дорожнє покриття, б — дорожня основа, в — ґрунт земляного полотна

Дорожнє покриття (рис. 1.1, а) — верхня частина дороги. Матеріал з якого виготовляють покриття має декілька характеристик, а саме, зносостійкість, міцність, шорсткість, спротив пластичним деформаціям, зберігати суцільність при зміні температур спричинених зміною пори року.

Дорожня основа (рис. 1.1, б) — необхідний для забезпечення розподілу навантаження на ґрунт земляного полотна, щоб протистояти прогинам дорожнього одягу при великих зовнішніх навантаженнях.

Додаткова основа — виконує ті самі функції що і звичайна основа, однак також може виконувати такі функції: відведення води з верхньої частини земляного полотна, зменшення глибини промерзання ґрунту, забезпечення проїзду для будівельної техніки під час будівельних робіт. Один шар додаткової основи здатен виконувати декілька функцій.

Конструкцію дорожнього одягу та матеріали визначають виходячи з транспортно-експлуатаційних вимог, інтенсивності руху, складу транспортних засобів в потоці, ґрунтів, кліматичних факторів, вимог безпеки та комфортності руху.

Залежно від ступеня надійності дороги поділяють на категорії табл. 1.1:

Таблиця 1.1. Коефіцієнти надійності різних категорій доріг

Категорія дороги	I-a	I-б, II	III	IV	V
Коефіцієнт надійності	0,97	0,95	0,9	0,85	0,75

1.2. Дефекти дорожнього покриття

Класифікацію основних дефектів зображено на рис. 1.2.

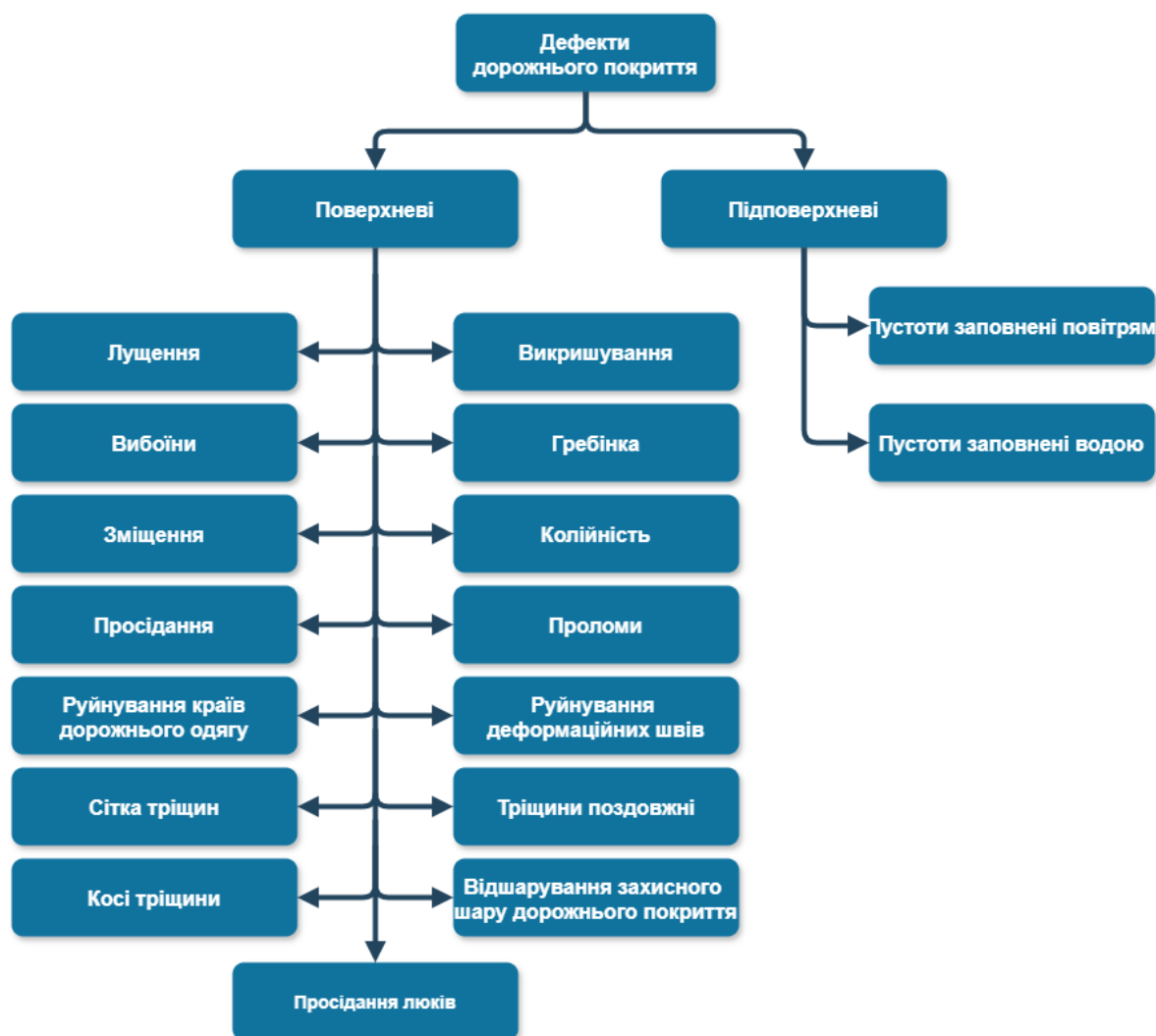


Рис. 1.2 Класифікація дефектів дорожнього покриття

Найбільш часто зустрічаються у дорожньому полотні такі дефекти: вибіони, тріщини загальні поперечні, сітка тріщин, тріщини поздовжні, косі

тріщини.

Вибойни — місцеві пошкодження верхнього шару дорожнього покриття у формі виїмок з виразними гострими та нерівними краями (рис. 1.3).



Рис. 1.3. Вибойни

Тріщини загальні поперечні — руйнування у вигляді поперечних тріщин.

Сітка тріщин — розгалуження тріщини, схоже на сітку (рис. 1.4).



Рис. 1.4. Сітка тріщин

Тріщини поздовжні — руйнування у вигляді поздовжніх тріщин, внаслідок інтенсивних втомних процесів.

Косі тріщини — тріщини які розташовані під кутом (рис. 1.5).



Рис. 1.5. Косі тріщини

1.3. Автоматизовані системи діагностики дорожнього покриття

Автоматизовані системи діагностики дорожнього покриття (рис. 1.6) зазвичай складаються з підсистем збору даних для визначення наявного стану дорожнього покриття, підсистеми підтримки прийняття рішення та підсистеми вирішення несправностей [4-5].

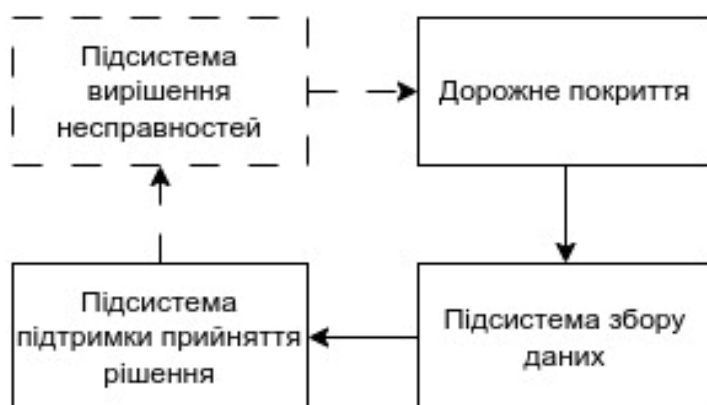


Рис. 1.6. Схема типової системи діагностики дорожнього покриття

Підсистема діагностики зазвичай складається з підсистем збору та обробки сигналів отриманих з датчиків. Для збору інформації можуть бути застосовані як автомобілі так і дрони. Обов'язковим елементом є обробка сигналів, при якій використовують, як класичні методи, так і штучні нейронні мережі. Ефективність даних підсистем можна оцінити економічними показниками [5], точністю виділення інформативних ознак [4], швидкістю збору та обробки сигналів, а також адаптивністю методів обробки, тобто здатністю якісно

аналізувати наявні сигнали в різних умовах отримання цих даних (різні датчики, різні погодні умови, різні способи збору інформації).

Експертна система — це тип системи, що імітує здатність експерта до прийняття рішень в галузі. Вона використовує базу знань, яка складається з правил і фактів, щоб оцінювати ситуації та надавати рекомендації або рішення.

Підсистема підтримки прийняття рішень (експертна система) робить висновки й рекомендації щодо необхідності втручання, часу в який дане втручання буде необхідно зробити та який саме тип втручання необхідний. Це стає можливим завдяки механізму логічного висновку та бази знань, також стають популярними рішення на базі нечітких нейронних мереж [7] та нейронних мереж для прогнозування [8-10]. Рішення в такі підсистемах формується на основі таких даних: історичних даних щодо втручань (ремонтів), нормативно технічних документах, даних про погодні умови на ділянці та оцінці наявного стану дорожнього покриття отриманих при діагностиці [6].

Дана підсистема своїми висновками допомагає експерту вирішити суперечності які виникають в процесі прийняття рішення щодо необхідності втручання. Ефективність даної підсистеми напряму залежить від якості та точності отриманих рекомендацій, які оптимізують час необхідний експертові для прийняття рішення та аналізу наявних даних.

У якості підсистеми вирішення несправностей можуть використовуватися ремонтні бригади, а також, ремонтні роботи, однак, останні мають багато технічних обмежень та високу ціну. Так, існують роботизовані системи здатні робити ремонт вибоїн [11], нанесення маркування [12] та герметизацію тріщин [13]. Дані роботи оснащені високоточними 3D сканерами, завдяки яким вони будуть модель ділянки й обраховують послідовність своїх дій.

Система інтелектуального моніторингу дорожнього покриття (рис. 1.7), яку показано на схемі рис., складається з кількох ключових компонентів, що забезпечують ефективний збір та аналіз даних про стан доріг. Основою системи є обладнання для збору інформації, яке включає різноманітні сенсори та супутникові навігаційні системи (GNSS, GPS, BDS), що встановлені на

спеціалізованих транспортних засобах. Ці сенсори дозволяють фіксувати як інженерну, так і екологічну інформацію в режимі реального часу. Зібрані дані передаються до системи обробки для їхнього подальшого аналізу [52].

Дані, отримані з сенсорів, обробляються за допомогою технологій візуалізації та аналізу зображень. Це дозволяє виявляти дефекти дорожнього покриття, такі як тріщини, нерівності чи ями, та оцінювати їхній масштаб і вплив на загальний стан дороги. Результати аналізу фіксуються у базі даних, яка містить детальну інформацію про стан окремих ділянок дороги. Це створює умови для довгострокового моніторингу та накопичення історичних даних про дорожню інфраструктуру.

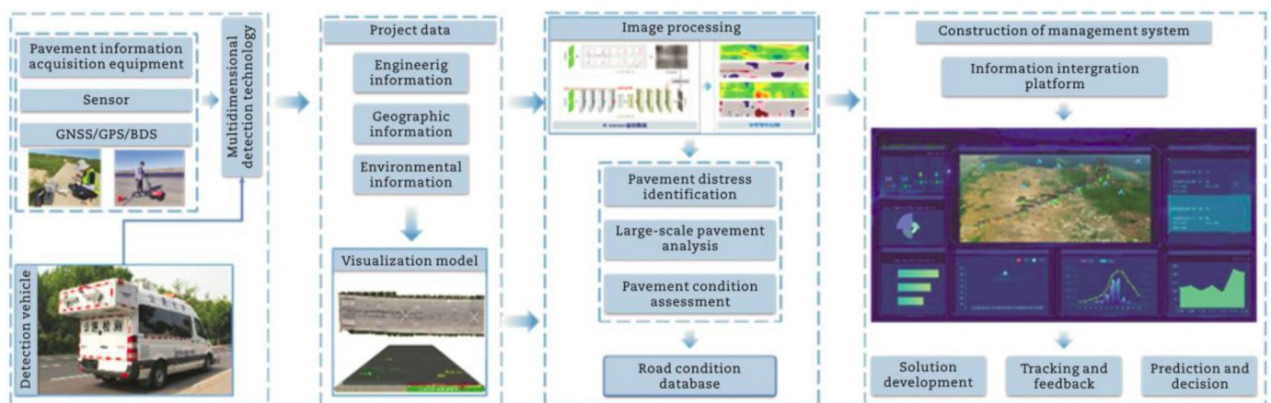


Рис.1.7. Система інтелектуального моніторингу дорожнього покриття [52]

Інформація, що зберігається у базі даних, інтегрується у платформу управління, яка дозволяє розробляти рішення для ремонту, а також прогнозувати та відстежувати стан покриття. Платформа об'єднує різні типи даних і забезпечує користувачам доступ до аналітичних інструментів, що сприяють підвищенню ефективності управління дорожньою мережею. Моніторинг здійснюється як у режимі реального часу, так і на основі накопичених історичних даних, що дозволяє передбачати майбутні зміни у стані доріг.

Основними викликами для цієї системи є необхідність складного програмного забезпечення, дорогих сенсорів та супутникових технологій, складність інтеграції великого обсягу даних і потреба в стабільному доступі до інтернету для передачі інформації. Окрім того, система залежить від точності зібраних даних, тому збої у роботі сенсорів можуть негативно вплинути на

результати моніторингу, а також, необхідні значні ресурси для зберігання та аналізу великих масивів інформації, що може сповільнювати роботу системи. І головним недоліком даної системи є відсутність включеного в систему механізму виправлення несправностей, що відокремлює теоретичну діагностику від практичного застосування отриманої інформації.

У дисертаційній роботі [4] було запропоновано автоматизовану систему діагностики (моніторингу) дорожнього покриття. Дана система складається з підсистем: збору даних, обробки даних, підтримки прийняття рішень. При зборі даних використовувались камери видимого випромінення (Nikon D5200, GoPro Hero 3), лазерні сканери (FARO Focus3D 120) та GPS. У підсистемі обробки даних відбувалась попередня обробка сигналів, інтегрувались геопросторові дані для побудови мапи пошкоджень, завдяки використанню моделей глибокого навчання відбувалась ідентифікація типу дефектів та на основі даних з лазерних сканерів будувалась 3D модель пошкодження. Підсистема підтримки прийняття рішення на основі історичних даних про ремонтні роботи, трафік і метеорологічні умови, використовуючи алгоритми прогнозування визначались пріоритетні зони ремонту та рекомендовані дії. Результати виводились в зручному для користувача інтерфейсі. Така система багато переваг, наприклад, висока точність визначення прощі та глибини пошкодження завдяки лазерним сканерам, однак, загальний відсоток правильно визначених дефектів не перевищив 87.2%, а також, лазерні сканери є значно дорожчими за інші сенсори, що використовуються при діагностиці.

1.4. Методи діагностики стану дорожнього покриття

Існує 3 види методів діагностики (рис. 1.8): експертно-візуальний, візуально-інструментальний та автоматизовані методи.

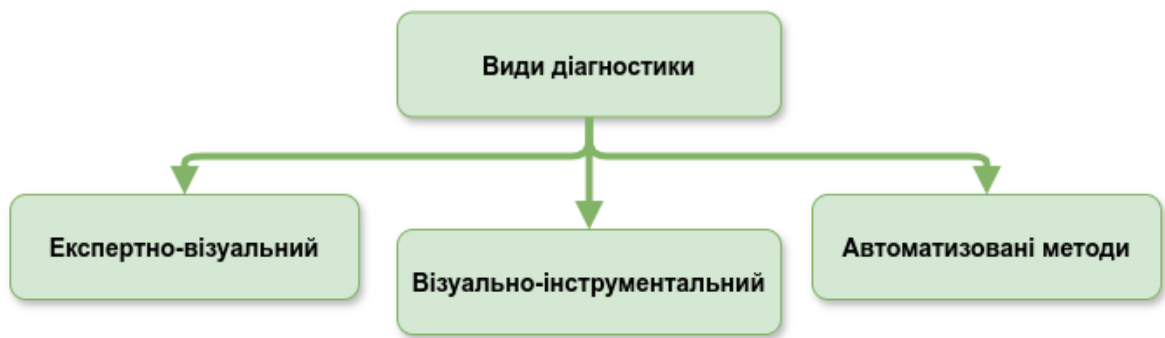


Рис. 1.8. Види методів діагностики

Перші два базуються на ручній перевірці стану дорожнього покриття оператором (експертом), відрізняються вони лише використанням спеціальних вимірювальних пристроїв у другому. Після проведення діагностики оператор вносить результати в централізовану базу даних. На базі цих результатів приймається рішення щодо ремонтних робіт на проаналізованій ділянці.

При автоматизованій діагностиці зазвичай дані отримуються з датчиків які розташовані на борту ходових лабораторій, дронів або супутників. Реєстраторами є прилади для фіксування видимого та інфрачервоного випромінювання, лазерні сканери та георадари, тобто, використовуються класичні методи неруйнівного контролю.

1.4.1. Методи дистанційного зондування

Видиме випромінювання.

Для більшості рішень автоматизації діагностики використовуються класичні алгоритми розпізнавання образів. Одними з таких алгоритмів є аналіз бінарних зображень [14, 15]. Прикладом використання є робота [16] де завдяки обробці бінарних зображень домоглися повного автоматичного визначення типу, серйозності тріщин не жорстких покриттів дорожніх тротуарів.

Частим розв'язанням задач виявлення дефектів дорожніх покриттів є метод виявлення країв об'єктів на зображенні. Одним з методів виявлення країв є використання фільтрів Габора, автори [17] використали його для виявлення поздовжніх та поперечних тріщин покриття. Інша робота також ставить за мету автоматизації виявлення тріщин [18]. Пропонують використовувати фільтри

Кенні (рис. 1.9) та швидке перетворення Хаара FHT, що використовуються для крайового аналізу зображень. В роботі використана статична процедура аналізу робочих характеристик приймача для того, щоб знайти оптимальний набір параметрів. Фільтр Кенні також був використаний у статті [19] де розглядалися зображення тротуару в градаціях сірого. Для зменшення шуму використали метод “розмивання” зображення Гаусовим фільтром. Зменшення шуму привело до здатності виявляти дефекти при наявності шуму та знаходити слабкі краї. Проте, при застосуванні “розмивання” відбувається згладжування меж, які є високочастотною ознакою.

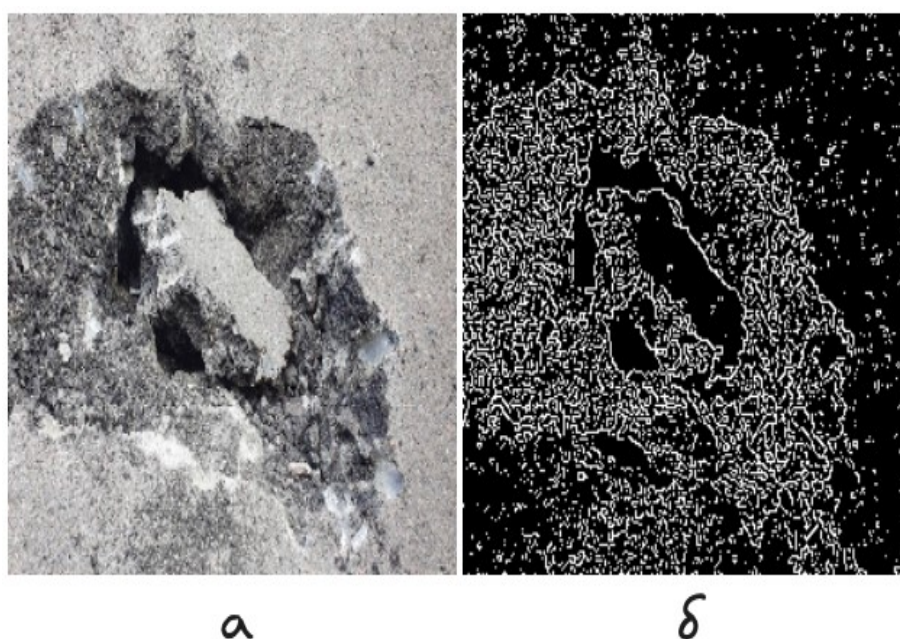


Рис. 1.9. Результат роботи фільтра Кенні: а — початкове зображення, б — результат обробки.

Дані датчики дозволяють виявляти видимі поверхневі дефекти в умовах хороших погодних умов, гарного освітлення та відсутності забруднення покриття. Зазвичай мають високу роздільну здатність і частоту кадрів, недорогі відносно інших приладів реєстрації.

Класичні методи, такі як порогова сегментація, виділення контурів або використання фільтрів для виділення текстур, мають суттєві обмеження в порівнянні з сучасними методами на основі нейронних мереж у завданнях діагностики дорожнього покриття. Вони часто виявляються недостатньо точними при роботі зі складними та варіативними умовами зйомки (освітлення,

погодні умови, тіні), що робить їх використання обмеженим в системах діагностики дорожнього покриття.

Інфрачервоне випромінювання

З появою тепловізійних камер з високою роздільною здатністю та термочутливістю розширився спектр можливого використання інфрачервоної термографії (ІЧТ), в тому числі для оцінки дорожнього покриття. Однак як і будь-яка технологія ІЧТ має низку не розв'язаних проблем. А саме, складність автоматизації обробки й аналізу термограм [20]. Низка факторів негативно впливають на якість ІЧТ: коливання температури навколишнього середовища, неоднорідність коефіцієнта випромінювання матеріалу дорожнього покриття, наявність тіней від придорожніх об'єктів, недостатня частота розгортки та інше [21, 22].

Вченні Чеської республіки у своїй праці [23] досліджували можливість використання ІЧТ для діагностики дорожнього покриття. Аналіз видимих та теплових знімків показав можливість виявлення дефектів на термограмі, а також, де що в деяких випадках зображення інфрачервоного спектра значно краще показує наявність дефектів. Також було виявлено деякі обмеження використання тепловізійних камер: швидкість машини має бути меншою за 40 км/г та кут нахилу камери має бути не менше 45°.

Автори роботи [24] використали коефіцієнт випромінювання асфальтного покриття як основний інформаційний параметр. Запропонований метод базується на тому, що зі старінням асфальтного покриття зменшується вміст маслянистих складових, що призводить до збільшення частки вапняку. Використовуючи високе поглинання вапняком ІЧ випромінювання на довжині хвилі 11,2 мкм, автори змогли виявити ділянки зношеності.

Дослідження використання ІЧТ для виявлення порожнин під дорожнім покриттям було описано в роботі [25]. ІЧТ надає можливість виявляти дефекти, які знаходяться під поверхнею, проте вона не завжди ефективна, коли використовуються традиційні методи обробки зображень, такі як, наприклад,

похідні зображення. Тому проведено дослідження щодо застосування сучасних методів аналізу ІЧ-зображень, таких як аналіз коефіцієнта детермінації (R^2), термографія основних компонентів (РСТ) і термографія розріджених основних компонентів (S-РСТ).

Для оцінки ефективності цих методів та їх точності виявлення порожнин під дорожнім покриттям проводили експеримент на лабораторній моделі проїжджої частини. Поверхню моделі дороги нагрівали протягом 6 годин для імітації сонячного нагріву, а потім, поки модель остигала, отримували ІЧ-зображення. Зокрема, зображення фіксувалися з інтервалами: 2 хвилини, 5 хвилин, 10 хвилин та 15 хвилин.

Для використання методу РСТ було використано послідовність ІЧ-зображень, знятих у моменти часу $t = t_1, t_2, \dots, t_n$. Перед застосуванням методу РСТ, кожне інфрачервоне зображення перетворювалося у вектор температури T_t . РСТ є методом лінійної декомпозиції, який чутливий до шуму на зображенні та викидів. Для боротьби з нелінійністю та прихованими змінними була створена модифікована версія РСТ, відома як термографія розріджених основних компонентів (S-РСТ). Найзначущою відмінністю від РСТ є те, що головні компоненти витягуються за допомогою завдання оптимізації, заснованої на загальній кількості ненульових елементів у векторі температури та сумі всіх абсолютних значень у векторі [26]. S-РСТ став більш надійним методом, за РСТ, що підтвердили результати виявлення порожнин.

Вчені корейського університету [27] дослідили можливість виявлення дерев'яних брусків, якими вони намагались моделювати внутрішній дефект, шляхом природного нагріву та тепловізора. Запропонований авторами метод аналізу відносної поверхневої температури дозволив виявляти дефекти на глибині до 30 мм та розмірами 20x20 мм.

Представлені автори сходяться в думці, що без попередньої обробки виявити підповерхневі дефекти дорожнього покриття важко та ідеального рішення досі не має. В наведених дослідженнях не було проаналізовано можливість використання зазначених методів для автоматизованої діагностики

дорожнього полотна. При розробці автоматизованих рішень буде важко спромогтися робити хоча б кілька знімків однієї ділянки для аналізу змін температури. А також існує проблема правильної класифікації дефектів типу порожнина та вибоїна, які при невеликій роздільній здатності на термограмі виглядають майже однаково.

Marcello Stimolo [28] провів ґрунтовний огляд пошкоджень, які можна виявити за допомогою ІЧТ в дорожньому та мостовому покритті. Особлива увага приділяється виявленню бульбашок у гідроізоляційних мембранах та настилах мостів. Однак виявлено, що цей метод не є ефективним при контролі доріг із наявним ямковим ремонтом. При аналізі термограм, яма після ремонту виглядає як дефект покриття через її схожість з порушенням суцільності.

Сучасні тепловізійні камери мають гарну роздільну здатність та термочутливість. На відміну від камер видимого спектру менше залежать від якості освітлення та погодних умов, повністю відсутній вплив короткострокових тіней (від автомобілів та пішоходів). Дають змогу аналізувати теплові профілі поверхні дорожнього покриття для прийняття якісних рішень щодо ремонту та здатність виявляти підповерхневі дефекти. До недоліків даного методу можна віднести чутливість камер даного типу до коливань температури навколишнього середовища, неоднорідності коефіцієнта випромінювання матеріалу дорожнього покриття, недостатня частота розгортки.

Георадари

Точним та популярним методом контролю порожнин вважають використання георадарів. Георадар випромінює короткохвильові імпульси в поверхню та реєструє відбите електромагнітне випромінювання, що забезпечує отримання зображень будь-якого підземного об'єкта або розриву. Недоліками такого підходу є складна інтерпретація для невідготовленого персоналу, помилки спричинені невизначеним складом матеріалу та розсіюванням сигналу через неоднорідність ґрунту, важко проводити контроль на ділянках з постійним рухом та великі часові витрати [29].

Автори роботи [30] використали георадар для виявлення неглибокого розшарування у покриттях із суцільного армованого бетону. В праці метод був оцінений як потенційно перспективний, але отримане зображення важко інтерпретувати, оскільки важко відрізнити область з дефектом від бездефектної.

При використанні георадара також виявленні проблеми при обробці та інтерпретації даних. Цій проблемі присвячена робота [31]. Автори прийшли до висновку, що автоматизація процесу буде вимагати використання багатоканальних радіолокаційних систем та багатоканальної обробки георадарних даних. Вони підкреслюють, що важливим аспектом застосування георадару у дослідженнях дорожніх покриттів є розробка зручних програмних систем для інтеграції результатів георадара з іншими методами.

Робота [32] присвячена ґрунтовному аналізу стану використання георадарів для контролю доріг. Зазначається що перспективним є напрямок використання інших методів контролю разом з георадаром. Також в роботі виділяють декілька головних проблем використання георадарів: складність інтерпретації результатів обстеження; ґрунт, який обстежується, може мати невідомі діелектричні властивості, які важко враховувати під час геолокації, і його неоднорідність може призводити до шуму на зображеннях, що робить процес аналізу та інтерпретації отриманих даних складним та вимагає додаткових досліджень та розробки алгоритмів для обробки результатів.

Лазерні сканери

Методи дистанційного зондування на основі лазерних сканерів або так званих лідарів (LiDAR – Light Detection and Ranging), набувають популярності.

Технології на основі лідарів дозволяють створювати просторово точні цифрові моделі фізичних об'єктів з високою роздільною здатністю, які відомі як "хмари точок" [33].

В роботі [34] використано технології 3D лазерного сканування для виявлення дефектів дорожнього покриття, а саме, маленькі вибоїни, вибоїни великих площ та поверхневі латки.

3D-моделі поверхні використали для визначення низинних ділянок,

схильних до дренажу та проблем із затопленням. Точність вимірювань таких нерівностей може досягати до 2 см [35].

Роздільна здатність висока LiDAR, як і точність вимірювання кута висока. Однак його робота сильно порушується при поганих погодних умовах, таких як дощ і туман, вітер і пісок тощо. Вимоги до робочого середовища високі і вартість висока [36-38].

1.4.2. Штучні нейронні мережі

Через великі структурні та природні відмінності між дефектами дорожнього покриття доцільно використати методи розпізнавання образів на основі нейронних мереж які ефективніше здатні розпізнавати об'єкти складної форми, вони здатні підсилити класичні методи неруйнівного контролю.

Так авторами роботи [39] була розроблена система контролю дорожнього покриття на базі нейронних мереж «Road Atlas». Що розв'язує задачі класифікації та сегментації дорожніх дефектів. Система розроблена за сучасною архітектурою web-додатків, має web-інтерфейс для взаємодії з системою, система має свій сервер, що виконує всі розрахунки, написаний на мікро фреймворку (micro framework) для web-додатків Flask. Сервер забезпечує обмін даними з інтерфейсом та безпечне зберігання інформації в документо-орієнтованій базі даних MongoDB, яка є різновидом NoSQL баз даних.

Іншим прикладом використання технологій нейронних мереж є робота [40], де для виявлення та класифікації поверхневих тріщин тротуарної плитки використали архітектуру глибокої згорткової нейронної мережі (DCNN). Оптичні зображення автори ділили на сегменти та з отриманих зображень сформували навчальну вибірку тріщин. Автори отримали показники точності на рівні 99%, однак, вони не змогли розв'язувати проблему наявності об'єктів-шумів, таких як, тіні, масляні плями, водяні плями на поверхні дороги. Наявність цих об'єктів на зображенні значно ускладнює класифікацію та зменшує точність DCNN.

Робота [41] виявила ще одну перешкоду до використання нейромережових

технологій, а саме погане освітлення покриття при фільмуванні. Для задачі виявлення дефектів тротуарів була використана нейронна мережа YOLO. Отримана точність виявлення дорівнювала 73,64%. Швидкість обробки зображень в 9 разів швидше, ніж у мережі швидких областей Faster R-CNN і тільки на 70%, ніж у SSD – мережі, що здатна виявляти об'єкти в режимі реального часу.

У роботі [42] пропонується економічно ефективний підхід на основі Kinect для тривимірної реконструкції поверхні дорожнього покриття та розпізнавання тріщин. Результати на рівні 89.09% точності класифікації показують, що цей метод є ефективним для отримання аналітичної інформації про вимірювання тріщин дорожнього покриття.

Робота [43] показує результати дослідження використання лазерних сканерів для класифікації тріщин які виникають в дорожньому покритті. Дослідникам вдалося досягти 99.6% точності класифікації в середньому.

Ще одна робота [44] показує результати класифікації дефектів дорожнього покриття при використанні датчика мобільного лазерного сканування. Точність класифікації тріщин склала 97.2%. А точність середня точність визначення ширини та довжини тріщин перевищила 81.2% та 89.7%.

Дослідження показують що при спільному використанні датчиків видимого та інфрачервоного діапазонів спектра можливо наблизитися до результатів точності георадарів та лазерних сканерів.

Головними завадами для використання нейромереж при автоматизованій діагностиці дорожнього покриття з використанням цифрових камер є наявність інформаційних завад на зображенні при реєстрації інформативних ознак стану дорожнього покриття (рис. 1.10). Такими завадами є наявність забруднення на поверхні, оптичні ілюзії, наявність тіней, погане освітлення, дорожня розмітка, які несуть в собі не інформативний характер при діагностиці.



Рис. 1.10. Інформативні ознаки при діагностиці дорожнього покриття

1.4.3. Комбіновані методи з використанням мультисенсорних систем

При використанні двох чи більше датчиків дозволяють отримати більше інформації про стан об'єкта і компенсувати недоліки різних методів реєстрації.

Дослідження [45] пропонує рішення для зменшення хибних спрацювань при попаданні в поле огляду плям розливу оливи, нафти, тіні або дорожньої розмітки — метод двосвітлової інспекції (DLI dual-light inspection). Інші методи часто помилково сприймають такі об'єкти за дефекти. При DLI використовується два зображення одного об'єкту з різною освітленістю, сам метод складається з чотирьох етапів:

- Захоплення зображень об'єкта
- Отримання по піксельної різниці двох зображень
- Цифрова обробка зображення
- Класифікація

Автори порівняли точність запропонованого методу з традиційними і виявили значне підвищення точності визначення тріщин та масляних розливів.

У роботі [46] аналіз застосування оптичних, лазерних, ультразвукових та мікрохвильових датчиків для класифікації поверхонь, виявив, що використання даних з декількох датчиків здатно дозволити отримати точніші й вірогідніші результати. Однак, це значно ускладнює автоматизацію процесу контролю.

Робота [47] представила метод сегментації тріщин на основі злиття даних видимого та інфрачервоного спектрів. В дослідженні порівнювались результати на 3 наборах даних видимих зображеннях, інфрачервоних та комплексованих. Результати показали, що для зображень видимих та злитих, точність роботи нейронної мережі майже однакова, однак значно краща за інфрачервоні зображення. У разі схожості тріщини з фоном найкращий результат показує злите зображення.

Комбінація використання георадару та інфрачервоної термографії визначення характеристик тріщин у дорожньому покритті описана в роботі [48]. Завдяки отриманню даних з інфрачервоного датчика стає можливим виявляти поверхневі та під поверхневі ушкодження. Однак важливо зауважити, що цей метод не надає інформації про глибину дефектів.

Автори пропонують комбінувати цей метод з даними з георадара. Дані, отримані двома різними методами під час виявлення тріщин, можуть взаємодоповнювати один одного. Георадар надає інформацію про глибину та розташування дефектів під поверхнею, тоді як інфрачервона термографія допомагає виявити та класифікувати тріщини на поверхні. Ця комбінація методів може забезпечити повну та точну інформацію про стан дорожнього покриття.

В оглядовій статті [49] наводяться актуальні роботи з поєднання даних георадару та інших датчиків для моніторингу стану дорожнього покриття. Використання різних методів неруйнівного контролю дозволяє отримати докладні карти стану дорожньої поверхні за різних погодних умов та освітлення. Як висновок зазначається, що застосування різних методів НК до одного об'єкту зводить до мінімуму обмеження кожного окремого методу, а загальна інформативність зростає.

Однак при використанні багатьох методів контролю росте складність та коштовність проведення таких досліджень, а також спільна обробка даних потребує більших зусиль через різну модальність даних. Через що постає питання щодо підвищення інформативності кожного окремого методу НК.

Автори роботи [50] провели аналіз робіт пов'язаних з моніторингом та

оцінкою мостів та цивільної інфраструктури загалом. Де ними було виокремлено 3 основних напрями майбутніх досліджень: розробка роботизованих платформ, використання багатьох датчиків для контролю, розробка алгоритмів опрацювання інформації отриманої від безлічі датчиків.

Робота [51] представила результати дослідження класифікації дефектів при використанні двох датчиків видимого та інфрачервоного спектрів. За результатами дослідження виявлено що таке поєднання датчиків здатно значно підвищити точність класифікації дефектів дорожнього покриття, так при однакових конфігураціях, точність класифікації при використанні тільки видимого спектра склала 96.57%, при використанні інфрачервоного спектра 93.83%, а при використанні обох спектрів попередньо поєднаних комплексуюванням зображень отримано точність на рівні 98.34%. Автори наголошують що такий підхід дає можливість більш дешевим датчикам змагатися в точності з дорогими лазерними сканерами. Також автори відмітили велику частку в результаті аргументації даних при навчанні нейромережевого класифікатора, так при додаванні розширення даних точність класифікації значно виросла.

У роботі [52] автори так само вивчали вплив використання даних інфрачервоного та видимого спектра для класифікації дефектів дорожнього покриття. Однак в цій роботі інфрачервоний спектр не був поєднаний з видимим у єдине зображення, замість цього автори пішли іншим шляхом, вони додали інфрачервоне зображення як ще один канал до RGB видимого перед подачею у нейромережевий класифікатор.

Як видно з представлених робіт поєднання інформації з декількох спектрів дає позитивний вплив на точність автоматизованої діагностики дорожнього покриття.

На основі опрацьованих даних можна запропонувати порівняльну таблицю датчиків для реєстрації поверхні дорожнього покриття табл. 1.2.

Таблиця. 1.2. Порівняльна таблиця методів зондування

	Георадар	Лідар	Камера	Тепловізійна камера
Ціна	від \$20,000	від \$50,000	від \$1,000	від \$5,000
Роздільна здатність	Висока	Висока	Висока	Низька
Складність обробки	Висока	Висока	Помірна	Помірна
Чутливість поганих погодних умов	Висока	Висока	Помірна	Помірна

Георадари та лазерні сканери попри свою точність і роздільну здатність в десятки разів дорожчі за цифрові камери, що робить напрями досліджень спрямовані на підвищення точності автоматизованого діагностування при використанні датчиків видимого та інфрачервоного діапазонів спектра актуальними.

Дослідження показують що при спільному використанні датчиків видимого та інфрачервоного діапазонів спектра можливо наблизитися до результатів точності георадарів та лазерних сканерів.

ВИСНОВКИ ДО РОЗДІЛУ 1

1. Дорожнє покриття є одним з найбільших об'єктів критичної інфраструктури. Своєчасна й точна діагностика запобігає надмірним витратам на ремонт, а також, збільшує безпеку на дорозі. Найпоширенішими дефектами дорожнього покриття є вибоїни та тріщини. Найнебезпечнішими є підповерхневі пустоти, бо в будь-який час вони можуть спричинити провал дорожнього покриття.

2. Жоден з наявних методів реєстрації не здатен повною мірою охопити всі аспекти діагностики дорожнього покриття. Методи мають свої переваги й недоліки, що приводить до думки одночасного використання декількох датчиків різної модальності, однак це значно ускладнює автоматизацію обробки отриманих даних. Дослідження показують що при спільному використанні датчиків видимого та інфрачервоного діапазонів спектра можливо наблизитися до результатів точності георадарів та лазерних сканерів.

3. Класичні методи комп'ютерного зору значно відстають в продуктивності й точності розв'язання задач в порівнянні з сучасними методами на базі нейронних мереж. З використанням нейронних мереж в автоматизованих системах діагностики стану дорожнього покриття зросла точність та швидкодія діагностики. Однак, існують фактори що негативно впливають на точність, а саме, інформаційні завади у вигляді тіней, забруднення дорожнього покриття, поганого освітлення, тощо.

4. Поєднання методів діагностики з нейронними мережами, що дає можливість використання сучасних методів машинного навчання для автоматизації обробки та оцифрування даних та використання цих даних для створення надійних моделей, що навчаються. Використання комбінованих методів зі штучним інтелектом дозволяє звузити певні обмеження, підвищити їх ефективність та проводити діагностику дорожнього покриття, що може стати основою для створення надійних систем автоматизованої діагностики дорожніми покриттями.

Список використаних джерел до розділу 1

- [1] Road Information Program. (2009). Rough roads ahead: Fix them now or pay for it later. AASHTO.
- [2] України, М. В. С. (2021). АНАЛІЗ ФАКТОРІВ, ЩО ВПЛИВАЮТЬ НА ДОРОЖНЬО-ТРАНСПОРТНІ ПРИГОДИ. Транспортна безпека: правові та організаційні аспекти: матеріали XVI Міжнародної науково-практичної конференції (в авторській редакції), (м. Кривий Ріг, 19 листопада 2021 року). Кривий Ріг, 2021. 238 с., 193.
- [3] ДСТУ 8954:2019. (2019). Автомобільні дороги. Оцінювання рівня дефектності дорожнього одягу.
- [4] Roberts, R. A. (2021). A new approach to Road Pavement Management Systems by exploiting Data Analytics, Image Analysis, and Deep Learning (Doctoral dissertation, Doctoral dissertation, University of Palermo).
- [5] Yang, X., Zhang, J., Liu, W., Jing, J., Zheng, H., & Xu, W. (2024). Automation in road distress detection, diagnosis and treatment. *Journal of Road Engineering*.
- [6] Pirayonesi, S. M., & El-Diraby, T. (2018). Using Data Analytics for Cost-Effective Prediction of Road Conditions: Case of The Pavement Condition Index:[summary report] (No. FHWA-HRT-18-065). United States. Federal Highway Administration. Office of Research, Development, and Technology.
- [7] Singh, A. P., Sharma, A., Mishra, R., Wagle, M., & Sarkar, A. K. (2018). Pavement condition assessment using soft computing techniques. *International Journal of Pavement Research and Technology*, 11(6), 564-581.
- [8] Sundin, S., & Braban-Ledoux, C. (2001). Artificial intelligence-based decision support technologies in pavement management. *Computer-Aided Civil and Infrastructure Engineering*, 16(2), 143-157.
- [9] Domitrović, J., Dragovan, H., Rukavina, T., & Dimter, S. (2018). Application of an artificial neural network in pavement management system. *Tehnički vjesnik*, 25(Supplement 2), 466-473.
- [10] Attoh-Okine, N. O. (1999). Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance. *Advances in engineering software*, 30(4), 291-302. Katsamenis, Iason, et al. "Robotic maintenance of road infrastructures: The heron project." *Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments*. 2022.
- [11] Kim, Y. S., & Haas, C. T. (2002). A man machine balanced rapid object model for automation of pavement crack sealing and maintenance. *Canadian Journal of Civil Engineering*, 29(3), 459-474.
- [12] Erdem, E., Haspalamutgil, K., Palaz, C., Patoglu, V., & Uras, T. (2011, May). Combining high-level causal reasoning with low-level geometric

- reasoning and motion planning for robotic manipulation. In 2011 IEEE International Conference on Robotics and Automation (pp. 4575-4581). IEEE.
- [13] Kulkarni, N. N., Dabetwar, S., Benoit, J., Yu, T., & Sabato, A. (2022). Comparative analysis of infrared thermography processing techniques for roadways' sub-pavement voids detection. *NDT & E International*, 129, 102652.
 - [14] Georgopoulos, A., Loizos, A., & Flouda, A. (1995). Digital image processing as a tool for pavement distress evaluation. *ISPRS Journal of Photogrammetry and Remote sensing*, 50(1), 23-33.
 - [15] Koutsopoulos, H. N., & El Sanhoury, I. (1991). Methods and algorithms for automated analysis of pavement images (No. 1311).
 - [16] Oliveira, H., & Correia, P. L. (2012). Automatic road crack detection and characterization. *IEEE Transactions on Intelligent Transportation Systems*, 14(1), 155-168.
 - [17] Zalama, E., Gómez-García-Bermejo, J., Medina, R., & Llamas, J. (2014). Road crack detection using visual features extracted by Gabor filters. *Computer-Aided Civil and Infrastructure Engineering*, 29(5), 342-358.
 - [18] Hutchinson, T. C., & Chen, Z. (2006). Improved image analysis for evaluating concrete damage. *Journal of Computing in Civil Engineering*, 20(3), 210-216.
 - [19] Du, Y., Pan, N., Xu, Z., Deng, F., Shen, Y., & Kang, H. (2021). Pavement distress detection and classification based on YOLO network. *International Journal of Pavement Engineering*, 22(13), 1659-1672.
 - [20] Сторожик, Д. В., Протасов, А. Г., Муравйов, О. В., Петрик, В. Ф., & Петренко, Д. В. (2022). Автоматизація процесу теплового неруйнівного контролю шляхом застосування методу комплексування термограм. *Технічна діагностика та неруйнівний контроль*, (2), 20-23.
 - [21] Sham, J. F., Memon, S. A., & Lo, Y. (2013). Application of continuous surface temperature monitoring technique for investigation of nocturnal sensible heat release characteristics by building fabrics in Hong Kong. *Energy and buildings*, 58, 1-10.
 - [22] Solla, M., Lagüela, S., González-Jorge, H., & Arias, P. (2014). Approach to identify cracking in asphalt pavement using GPR and infrared thermographic methods: Preliminary findings. *Ndt & E International*, 62, 55-65.
 - [23] Janků, M., & Stryk, J. (2017, September). Application of infrared camera to bituminous concrete pavements: measuring vehicle. In *IOP Conference Series: Materials Science and Engineering* (Vol. 236, No. 1, p. 012104). IOP Publishing.
 - [24] Kulkarni, N. N., Dabetwar, S., Benoit, J., Yu, T., & Sabato, A. (2022). Comparative analysis of infrared thermography processing techniques for

- roadways' sub-pavement voids detection. *NDT & E International*, 129, 102652.
- [25] Wang, Q., Xia, R., Liu, Q., Zhou, H., Qiu, J., & Zhao, B. (2020). Pixel-based thermal sequence processing algorithm based on R2 fractile threshold of non-linear fitting in active infrared thermography. *Infrared Physics & Technology*, 109, 103422.
- [26] Sim, J. G., & Kim, K. H. (2015). Detecting of the defects of pavement of a road by using infrared thermography. *Journal of the Korean Society for Advanced Composite Structures*, 6(3), 69-76.
- [27] Stimolo, M. (2002, March). Practical utilization of thermography in road construction and in waterproofing systems. In *Thermosense XXIV* (Vol. 4710, pp. 299-306). SPIE.
- [28] Amran, T. S. T., Amin, M. S. M., Ahmad, M. R., Azreen, N. M., Sani, S., Adnan, M. A. K., ... & Sayuti, S. (2021, March). NDT methods in inspecting road and highway structures. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1106, No. 1, p. 012034). IOP Publishing.
- [29] Liu, J., Zollinger, D. G., & Lytton, R. L. (2008). Detection of delamination in concrete pavements using ground-coupled ground-penetrating radar technique. *Transportation research record*, 2087(1), 68-77.
- [30] Bianchini Ciampoli, L., Tosti, F., Economou, N., & Benedetto, F. (2019). Signal processing of GPR data for road surveys. *Geosciences*, 9(2), 96.
- [31] Solla, M., Pérez-Gracia, V., & Fontul, S. (2021). A review of GPR application on transport infrastructures: Troubleshooting and best practices. *Remote Sensing*, 13(4), 672.
- [32] Lato, M. J., Diederichs, M. S., Hutchinson, D. J., & Harrap, R. (2012). Evaluating roadside rockmasses for rockfall hazards using LiDAR data: optimizing data collection and processing protocols. *Natural Hazards*, 60, 831-864.
- [33] Chang, K. T., Chang, J. R., & Liu, J. K. (2005). Detection of pavement distresses using 3D laser scanning technology. In *Computing in civil engineering* (2005) (pp. 1-11).
- [34] Uddin, W. (2011). Remote sensing laser and imagery data for inventory and condition assessment of road and airport infrastructure and GIS visualization. *International Journal of Roads and Airports*, 1(1), 53-67.
- [35] Coenen, T. B., & Golroo, A. (2017). A review on automated pavement distress detection methods. *Cogent Engineering*, 4(1), 1374822.
- [36] Tsai, Y. C. J., & Li, F. (2012). Critical assessment of detecting asphalt pavement cracks under different lighting and low intensity contrast conditions using emerging 3D laser technology. *Journal of Transportation Engineering*, 138(5), 649-656.

- [37] Zhang, D., Xu, X., Lin, H., Gui, R., Cao, M., & He, L. (2019). Automatic road-marking detection and measurement from laser-scanning 3D profile data. *Automation in Construction*, 108, 102957.
- [38] Ouyang, A., Luo, C., & Zhou, C. (2011). Surface distresses detection of pavement based on digital image processing. In *Computer and Computing Technologies in Agriculture IV: 4th IFIP TC 12 Conference, CCTA 2010, Nanchang, China, October 22-25, 2010, Selected Papers, Part IV 4* (pp. 368-375). Springer Berlin Heidelberg.
- [39] Chen, Z., Zhang, Y., Luo, Y., Wang, Z., Zhong, J., & Southon, A. (2021, December). Roadatlas: Intelligent platform for automated road defect detection and asset management. In *Proceedings of the 3rd ACM International Conference on Multimedia in Asia* (pp. 1-3).
- [40] Yusof, N. A. M., Ibrahim, A., Noor, M. H. M., Tahir, N. M., Yusof, N. M., Abidin, N. Z., & Osman, M. K. (2019, November). Deep convolution neural network for crack detection on asphalt pavement. In *Journal of Physics: Conference Series* (Vol. 1349, No. 1, p. 012020). IOP Publishing.
- [41] Zhang, Y., Chen, C., Wu, Q., Lu, Q., Zhang, S., Zhang, G., & Yang, Y. (2018). A kinect-based approach for 3D pavement surface reconstruction and cracking recognition. *IEEE Transactions on Intelligent Transportation Systems*, 19(12), 3935-3946.
- [42] Zhou, S., & Song, W. (2020). Deep learning-based roadway crack classification using laser-scanned range images: A comparative study on hyperparameter selection. *Automation in Construction*, 114, 103171.
- [43] Zhong, M., Sui, L., Wang, Z., & Hu, D. (2020). Pavement crack detection from mobile laser scanning point clouds using a time grid. *Sensors*, 20(15), 4198.
- [44] Su, Y., Kang, S., Chang, J., & Hsieh, S. (2010). Using dual lights for robotic pavement inspection. In *Proceedings of SICE Annual Conference* (Vol. 1).
- [45] Bystrov, A., Hoare, E., Tran, T. Y., Clarke, N., Gashinova, M., & Cherniakov, M. (2018, September). Sensors for automotive remote road surface classification. In *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)* (pp. 1-6). IEEE.
- [46] Liu, F., Liu, J., & Wang, L. (2022). Asphalt pavement crack detection based on convolutional neural network and infrared thermography. *IEEE Transactions on Intelligent Transportation Systems*, 23(11), 22145-22155.
- [47] Solla, M., Lagüela, S., González-Jorge, H., & Arias, P. (2014). Approach to identify cracking in asphalt pavement using GPR and infrared thermographic methods: Preliminary findings. *Ndt & E International*, 62, 55-65.

- [48] Elseicy, A., Alonso-Díaz, A., Solla, M., Rasol, M., & Santos-Assunção, S. (2022). Combined use of GPR and other NDTs for road pavement assessment: An overview. *Remote Sensing*, 14(17), 4336.
- [49] Ahmed, H., La, H. M., & Gucunski, N. (2020). Review of non-destructive civil infrastructure evaluation for bridges: State-of-the-art robotic platforms, sensors and algorithms. *Sensors*, 20(14), 3954.
- [50] Chen, C., Chandra, S., Han, Y., & Seo, H. (2021). Deep learning-based thermal image analysis for pavement defect detection and classification considering complex pavement conditions. *Remote Sensing*, 14(1), 106.
- [51] Chen, C., Chandra, S., & Seo, H. (2022). Automatic pavement defect detection and classification using RGB-thermal images based on hierarchical residual attention network. *Sensors*, 22(15), 5781.
- [52] Liu, P., Wang, C., Lu, W., Moharekpour, M., Oeser, M., & Wang, D. (2022). Development of an FEM-DEM model to investigate preliminary compaction of asphalt pavements. *Buildings*, 12(7), 932.

РОЗДІЛ 2. Технології діагностування стану дорожнього покриття в автоматизованих системах

2.1. Комплексування зображень дефектів дорожнього покриття

Комплексування зображень (злиття зображень з англ. image fusion) - поєднання двох або більше зображень одного об'єкту з метою збільшення інформативності результуючого зображення та підвищення придатності для подальшої комп'ютерної обробки. Цей процес також здатен зменшити вплив різного роду шумів та завад. [1-3] або написати якого побічного ефекту.

Залежно від природи зображень та цілей, комплексування поділяють на [4]:

Багаторакурсне злиття зображень (англ. multiview image fusion) — зображення однієї модальності, наприклад, зображення одного діапазону хвиль електромагнітного спектра, але з різних точок спостереження;

Мультимодальне злиття зображень (англ. multimodal image fusion) — поєднання зображень різної природи, наприклад, видимого та інфрачервоного діапазонів спектра;

Різночасове злиття зображень (англ. multitemporal image fusion) — комплексування зображень отриманих в різний час;

Багатофокусне злиття зображень (англ. multifocus image fusion) — поєднання зображень зроблених на об'єктиви з різними фокусними відстанями;

Комплексування зображень з метою зменшення шумів (англ. fusion for image restoration);

Комплексування зображень для збільшення роздільної здатності (англ. fusion for image restoration).

Методи комплексування по типах поділяють на такі категорії (рис.2.1): комплексування на рівні пікселів, комплексування на рівні ознак, комплексування на рівні рішень.

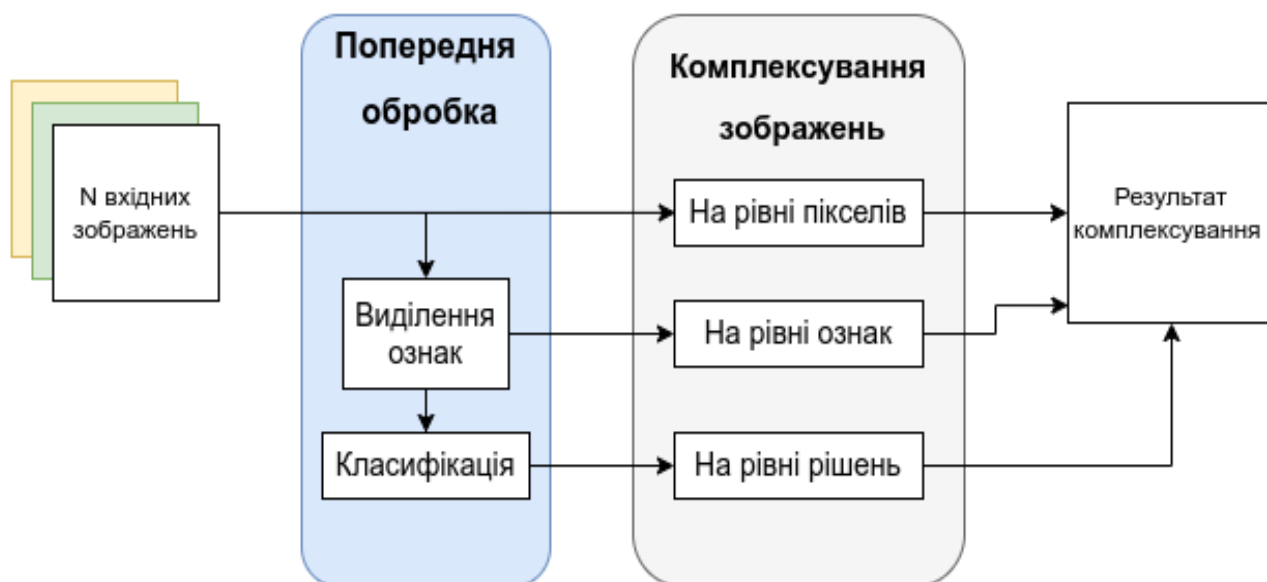


Рис. 2.1. Схема процесу комплексування залежно від типу методу.

Комплексування на рівні пікселів є одночасно і найпростішим типом і найнадійнішим, що виконується на найнижчому рівні [5]. Таке комплексування відбувається шляхом по піксельного перетворення інтенсивності пікселів вхідних зображень в піксель нового зображення.

Комплексування на рівні ознак або рівні функцій ґрунтуються на вилученні та поєднанні відповідних функцій з вхідних зображень, тобто, виділяє об'єктами для операцій вже певні характерні області. До таких функцій (ознак) можуть бути віднесені: інтенсивність пікселів, текстури, краї зображення або цілі об'єкти [6, 7]. Недоліком цього рівня є ресурсовитратні алгоритми та відсутність універсального методу сегментації та розпізнавання ознак.

Найменш вивченою групою є комплексування на рівні рішень, що спричинено абстрактністю результату. Результатом є вже не комбіноване зображення, а деякий інформативний опис реальної картини. При використанні цього методу зменшується надмірність і невизначеність інформації.

2.1.1. Методи комплексування зображень

Методи злиття можливо використовувати як і в просторовій, так і в частотній області. При просторовому методі операції проводяться над пікселями, а при частотній, зображення розкладається на масштабовані коефіцієнти вже на рівні яких відбувається злиття [8]. Частотні методи здатні усувати просторові

спотворення.

До просторових методів комплексування зображень відносять: метод максимуму, метод мінімуму, міжрядковий метод, просте усереднення, виважене усереднення, пороговий метод, степеневе перетворення, підсилення відмінностей, метод маски, метод вагової функції, виділення низькочастотної складової [9], перетворення інтенсивності-насиченості (Hue-Intensity-Saturation (HIS)), перетворення Брові (Brovey), аналіз головних компонентів (Principal Component Analysis (PCA)).

Метод максимуму ґрунтується на послідовному виборі максимального значення інтенсивності пікселя з двох або більше зображень. Результат комплексування можна побачити на рис. 2.2., а математично метод описується виразом (2.1) [10].

$$F(x, y) = \max\{A(x, y), B(x, y)\}, \quad (2.1)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення.

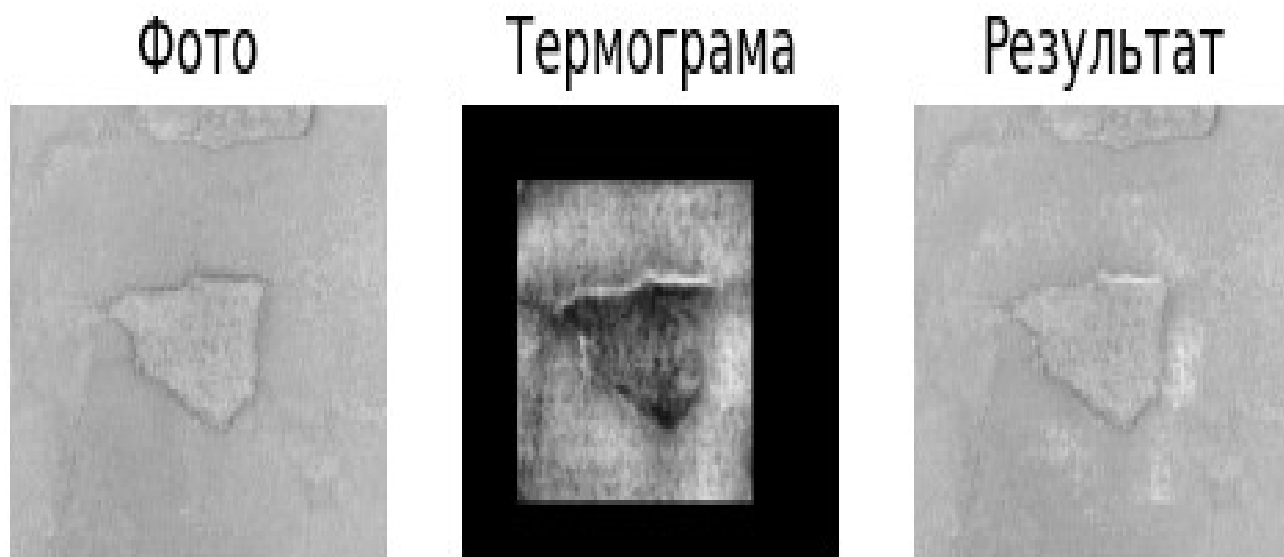


Рис. 2.2. Результат комплексування зображення дорожніх дефектів методом максимуму.

За таким же принципом працює **метод мінімумів**, однак, результат формується з мінімальних значень інтенсивності [11]:

$$F(x, y) = \min\{A(x, y), B(x, y)\}, \quad (2.2)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення.

Метод черзрядкового комплексування полягає у послідовному чергуванні рядків зображень. Цей метод реалізується шляхом складання зображень $F1(x, y)$ і $F2(x, y)$ з чергуванням непарних і парних рядків з номером N [12]:

$$F(x, y) = \{A(x, y), y \in 1, 3, 5, \dots, N - 1 \cup B(x, y) y \in 2, 4, 6, \dots, N \}, \quad (2.3)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення, N – кількість рядків в матриці пікселів.

Метод усереднення описується виразом (2.4), результат якого наведено на рис. 2.3 [13].

$$F(x, y) = \frac{A(x, y) + B(x, y)}{2}, \quad (2.4)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення.

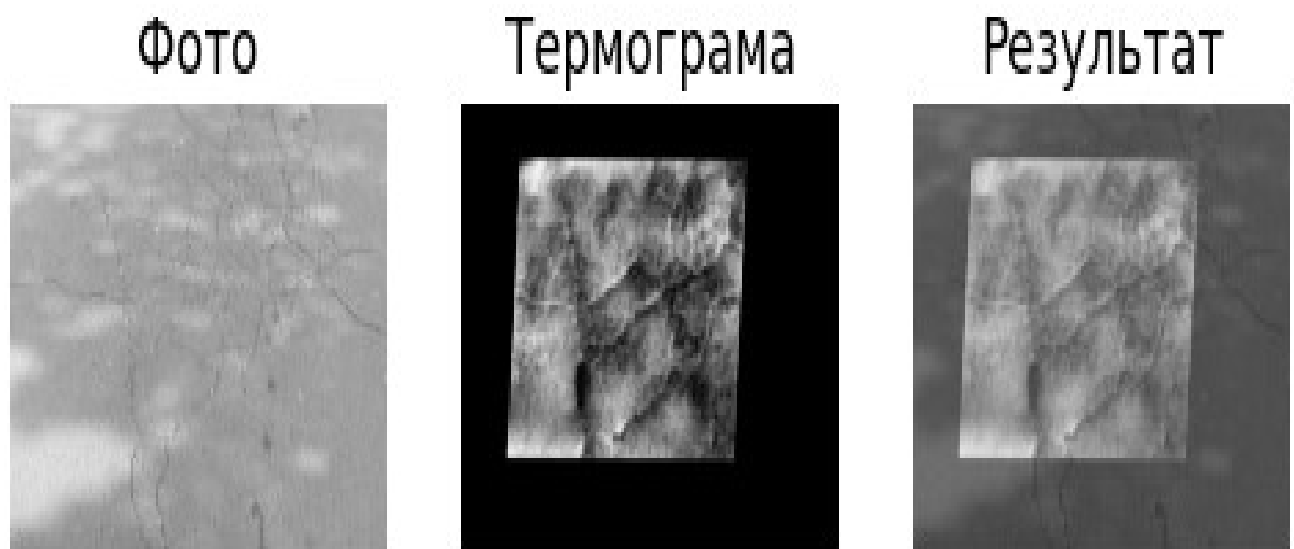


Рис. 2.3. Результат комплексування зображення дорожніх дефектів методом усереднення [13]

При **виваженому усередненні** кожному зображенню надають ваговий коефіцієнт. Результатом комплексування є зображення з інтенсивністю пікселя розрахованою як сума зважених інтенсивностей пікселів вхідних зображень.

Вираз що описує цю операцію (2.5) [14]. Популярним методом для автоматичного визначення вагів є аналіз головних компонент (PCA) [15].

$$F(x, y) = w_1 A(x, y) + w_2 B(x, y), \quad (2.5)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення, w_1 та w_2 – вагові коефіцієнти.

Метод порогового комплексування обирає з якого пікселя буде формуватися результат порівнюючи інтенсивність з деяким пороговим рівнем (2.6) [16].

$$F(x, y) = \{A(x, y) < T \mid A(x, y) > T\}, \quad (2.6)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення, T - порог.

Метод степеневого перетворення ґрунтується на виразі (2.7), де одне зображення використовується в якості степеневого коефіцієнта для іншого [17]. Зазвичай використовується для отримання зображення зі збільшеним контрастом.

$$F(x, y) = A(x, y)^{1-B(x, y)/2^n}, \quad (2.7)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення, n – коефіцієнт для регуляції сили перетворення.

Методу **підсилення відмінностей** [18] на основі свого алгоритму підсилює відмінності між початковими зображеннями.

Метод маски передбачає виконання такого алгоритму [19]:

1. Нормалізація інтенсивностей пікселів зображень
2. Визначення порогового значення k інтенсивності пікселів
3. Комплексування згідно виразу (2.8):

$$F(x, y) = \{2A(x, y)B(x, y) + A(x, y)^2(1 - 2B(x, y)), A(x, y) < T \mid 2A(x, y)(1 - 2B(x, y)) + \sqrt{A(x, y)}(2B(x, y) - 1), A(x, y) \geq T\}, \quad (2.8)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення, T – пороговий коефіцієнт з маски.

Метод вагової функції описується виразом [20]:

$$F(x, y) = kA(x, y) + (1 - k)B(x, y), \quad (2.9)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення, k – ваговий коефіцієнт.

Комплексування шляхом **виділення низькочастотної складової** проводиться згідно виразу [9]:

$$F(x, y) = \frac{1}{3} \left[(1 - A(x, y)) + (1 - B(x, y)) + \frac{1}{2} (A(x, y) + B(x, y)) \right], \quad (2.10)$$

де x та y координати пікселів зображення, A та B початкові зображення, F результуюче зображення.

Часто даний метод використовується для отримання в результаті більш чітких контурів в зображенні.

Метод **перетворення інтенсивності-насиченості** (Hue-Intensity-Saturation (HIS)) техніка злиття кольорів при якій зображення перетворюється на мультиспектральні компоненти HIS та панхроматичні зображення [10, 21]. Спектральна компонента складається з трьох матриць, по одній на кожен колір, малого обсягу та панхроматичну матрицю високої роздільної здатності. Для зворотного процесу використовують зворотне перетворення. Даний метод може викликати спотворення кольору при великих відмінностях між зображеннями у відтінках сірого [22].

Комплексування зображень на базі **перетворення Брові**. Дане перетворення є методом об'єднання зображень, що здатне зливати зображення різної роздільної здатності. При цьому методі добуток панхроматичного та мультиспектральних знімків ділиться суму яскравостей мультиспектрального. Фактично це є операцією нормалізації спектральних смуг перед їх перемноженням з панхроматичним зображенням. Перевагами цього методу є хороша передача кольору та стабільність результатів [23]. Такий метод використовують для комплексування зображень з різних датчиків [24].

Аналіз основних компонентів (PCA) дуже схожий з перетворення інтенсивності-насиченості, однак здатен використовувати будь-яку кількість спектральних зображень. При використанні цього методу визначають один із головних компонентів, який містить загальну інформацію всіх діапазонів. Цей компонент має високу дисперсію і надає більше інформації про панхроматичне

зображення. Компонент панхроматичного зображення високої роздільної здатності розтягується, щоб мати ту ж дисперсію, що й головний компонент, і замінює цей головний компонент. Потім застосовується зворотне перетворення РСА для отримання комплексованого зображення високої роздільної здатності. Недоліком методу РСА є спектральна деградація та спотворення кольору [25,26].

До частотних методів комплексування зображень відносять: методи пірамід (Лапласа та Гауса), дискретне вейвлет-перетворення, дискретне косинусне перетворення.

Метод пірамід ґрунтується на перетворенні початкового зображення на масив зображень різної роздільної здатності (рівня дискретизації), що є описом бінокулярної зорової системи людини. Кожне наступне зображення виходить з попереднього, але зі зменшеною роздільною здатністю. Для комплексування кожен рівень піраміди першого зображення зливають з відповідним рівнем піраміди другого зображення, після чого з новоутвореної піраміди генерується комплексоване зображення зворотнім перетворенням.

При формуванні **піраміди Гауса** кожен рівень перед дискретизацією фільтрується фільтром Гауса (низькочастотна фільтрація) для утворення більш гладкого зображення, тим самим прибираються шуми. Через нелінійність підвищення та зниження дискретизації, після отримання результуючого зображення спостерігається втрата частини інформації та зображення буде розмите [27].

Метод злиття **піраміди Лапласа** схожий на піраміду Гауса, однак, на кожному рівні, окрім найменшого, зберігається різниця розмитих зображень. При перетворенні зображення розкладається на різні частотні діапазони. Об'єднання відбувається окремо на кожному рівні, що дає можливість виділяти деталі на конкретному діапазоні частот [28].

Дискретне вейвлет-перетворення (Discrete Wavelet Transform (DWT)) є альтернативою короткочасовим перетворенням Фур'є. Перевагою DWT є забезпечення бажаної роздільної здатності як в просторовій, так і в частотних

областях, на відміну від Фур'є що дає високу роздільну здатність лише на частотній області. Дискретне вейвлет-перетворення розкладає вхідний сигнал на масштабовані та зсунуті форми вихідного вейвлета. При комплексуванні зображення спочатку розкладають на коефіцієнти DWT, після чого використовуючи деяке правило злиття (наприклад, середнє значення), результуюче зображення формується з отриманих коефіцієнтів зворотнім перетворенням [29].

Комплексування на основі DWT (рис. 2.4) мінімізує спектральні спотворення в результуючих об'єднаних зображеннях з меншою роздільною здатністю через краще співвідношення сигнал/шум порівняно з піксельними методами. Тобто, мінімізує колірні спотворення які спостерігаються в просторових методах злиття [30]. У роботах [11, 31] наголошують на, тому що DWT значно краще справляється з комплексуванням в порівнянні з пірамідами Лапласа та Гауса.

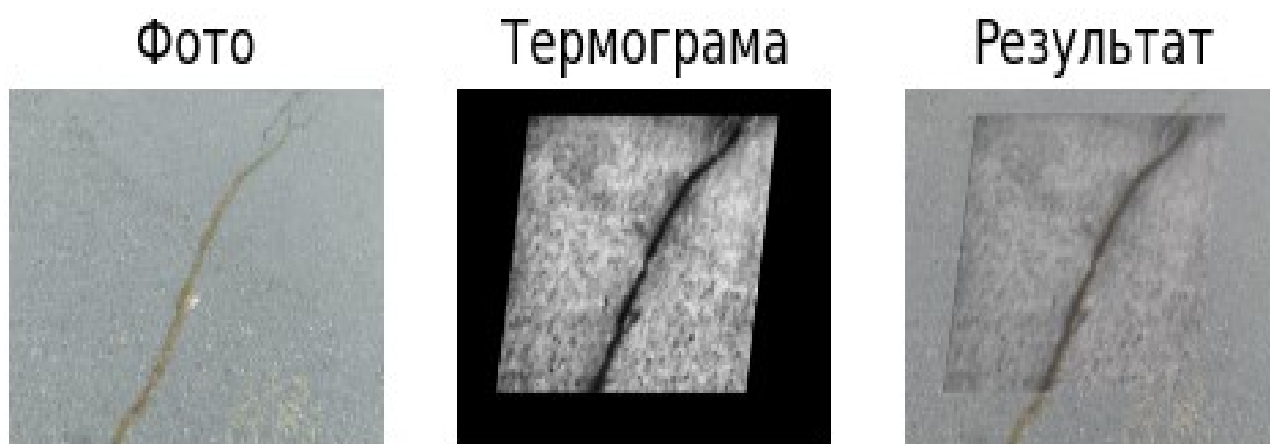


Рис. 2.4. Результат комплексування зображення дорожніх дефектів з використанням DWT

Комплексування з використанням дискретного косинусного перетворення (Discrete Cosine Transform (DCT)) злиття виконується на рівні спектральних коефіцієнтів, які характеризують амплітуди просторових частот. Дане перетворення застосовується для стиснення зображень, а саме, таких форматів MPEG, JVT, JPEG [32, 33].

Головна задача комплексування зображень це збільшення

інформативності, тому стає задача вимірювати інформативність зображень. Найпоширеніший та авторитетний критерій якості для цього вважається ентропія Шенона [34, 35].

Ентропія це міра невизначеності в повідомленні, яка часто використовується як критерій якості інформативності зображень. Для зображень інформативність (ентропію Шенона) отримують на основі розподілу інтенсивностей пікселів.

Ентропія Шенона розраховується за формулою:

$$H(X) = - \sum_{i=1}^n P(x_i) \log_2 P(x_i), \quad (2.10)$$

де $P(x)$ - ймовірність набуття значення x , випадковою величезною X . у контексті зображень це означає, що оцінюється ймовірність різних рівнів інтенсивностей пікселів для обчислення ентропії.

Більший рівень ентропії означає що зображення має більшу інформативність, тим самим воно стає більш непередбачуваним. Тому низька ентропія вказує на наявність великої кількості однакових елементів або високого рівня повторюваності.

2.2. Використання нейронних мереж в системах діагностики

2.2.1. Модель штучного нейронна

Першу модель штучного нейрона було запропоновано в 50-х роках минулого сторіччя нейробіологом Френком Розенблатом, яку він назвав перцептрон [36] - адитивна нейронна мережа. У своїй роботі вчений ґрунтувався на знаних отриманих при дослідженні біологічних нейронів, тому його модель мала такі ж стадії обробки інформації: приймання інформації, агрегація (зважена сума) інформації і якщо отримане значення вище порогу передача його іншим нейронам (рис. 2.5).

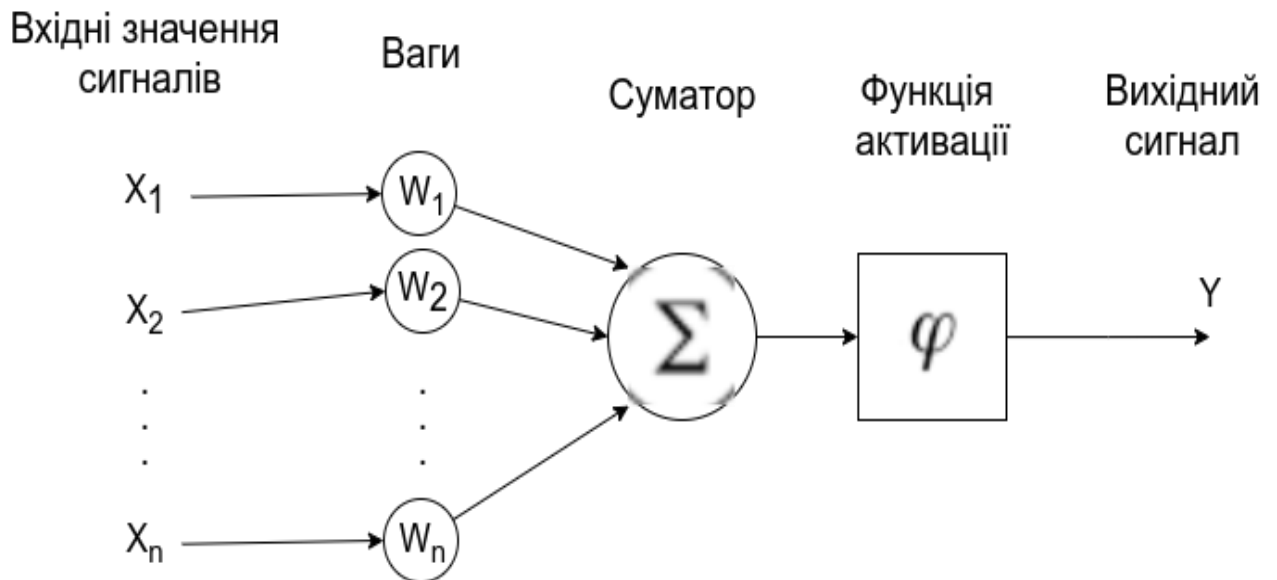


Рис. 2.5. Модель штучного нейрона (перцептрон)

Для математичного опису моделі штучного нейрона вводиться поняття *параметрів* нейронна в перелік яких входять *ваги* та *зміщення* (bias). Ваги нейрона необхідні для отримання зваженої суми отриманої інформації від інших нейронів, зміщення це негативне порогове значення.

Після вводу цих визначень математична модель перцептрона набуває вигляду:

$$f(x) = \{1, w * x + b > 0 \quad 0, \quad (2.11)$$

де w - масив всіх ваг входів нейронна, b - зміщення.

На сьогоднішній день така модель нейронна вже рахується застарілою. В першу чергу через бінарність входів та виходів нейрона. Через що стає неможливим робити прогнозування (вихід завжди 0 чи 1), а також це ускладнює машинне навчання.

Для розв'язання цих проблем було запропоновано іншу модель нейронна з *сигмоїдною функцією активації* (рис. 2.6). Що описується функцією [37]:

$$\sigma(z) = \frac{1}{1+e^{-z}}, \quad (2.12)$$

де z – це еквівалент $w x + b$.

Результатом будь-якої функції активації називають *активацію*. А будь-який нейрон що використовує сигмоїдну функцію активації називають *сигмоїдним нейроном*.

Як видно з рис. 2.6 невеликі зміни параметрів w та b викликають невеликі зміни активації, що і є головною перевагою в порівнянні з перцептроном. При великих значеннях по модулю значення z активація як і в випадку з перцептроном буде мати значення 0 або 1, що призводить до малого впливу змін коефіцієнтів на результат, такий ефект називають *насиченням нейрона*.

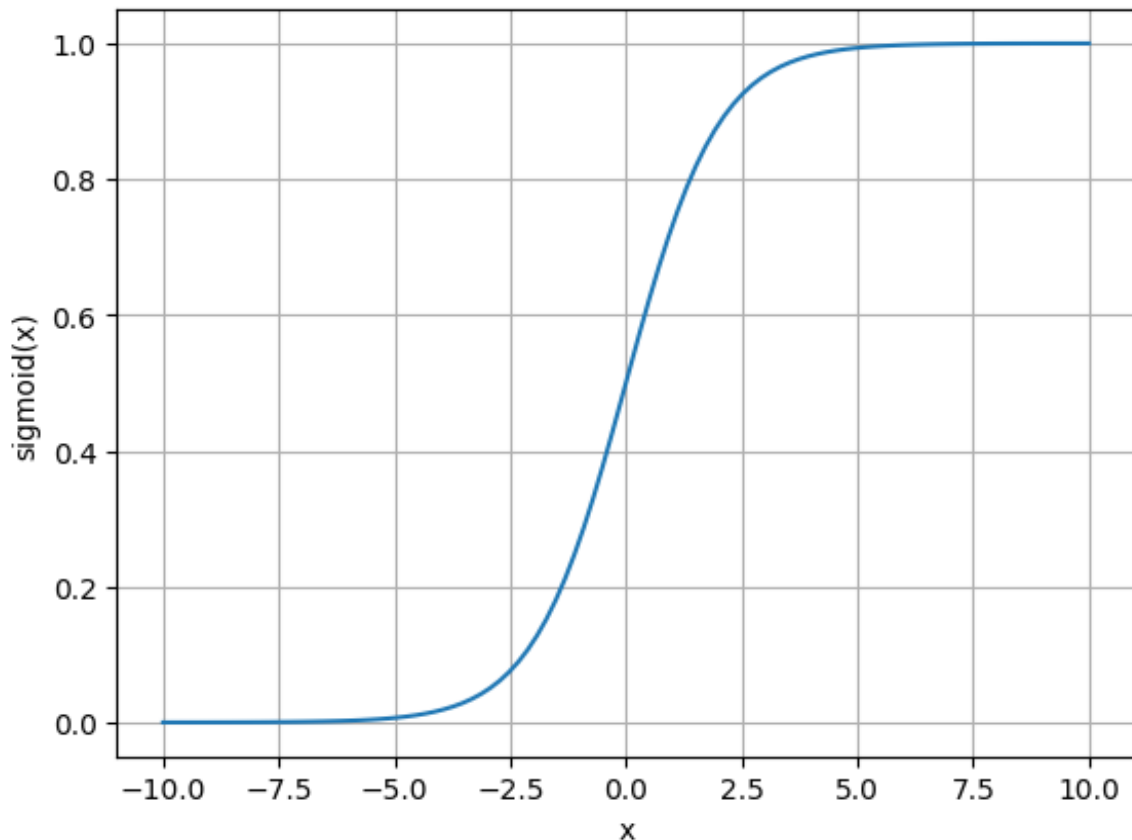


Рис. 2.6. Сигмоїдна функція активації

Прикладом нелінійної функції активації є блок лінійної ректифікації (Rectified Linear Unit, ReLU), який зображено на рис 2.7. Робота цієї функції заснована на принципах роботи біологічного нейрона, так потенціал запуску нейронів має лише позитивний режим [38]. Математично дану функцію можна описати як:

$$\sigma(z) = \begin{cases} z, & z > 0 \\ 0, & z \leq 0 \end{cases}, \quad (2.13)$$

де z – це еквівалент $w x + b$, а — активація нейрона.

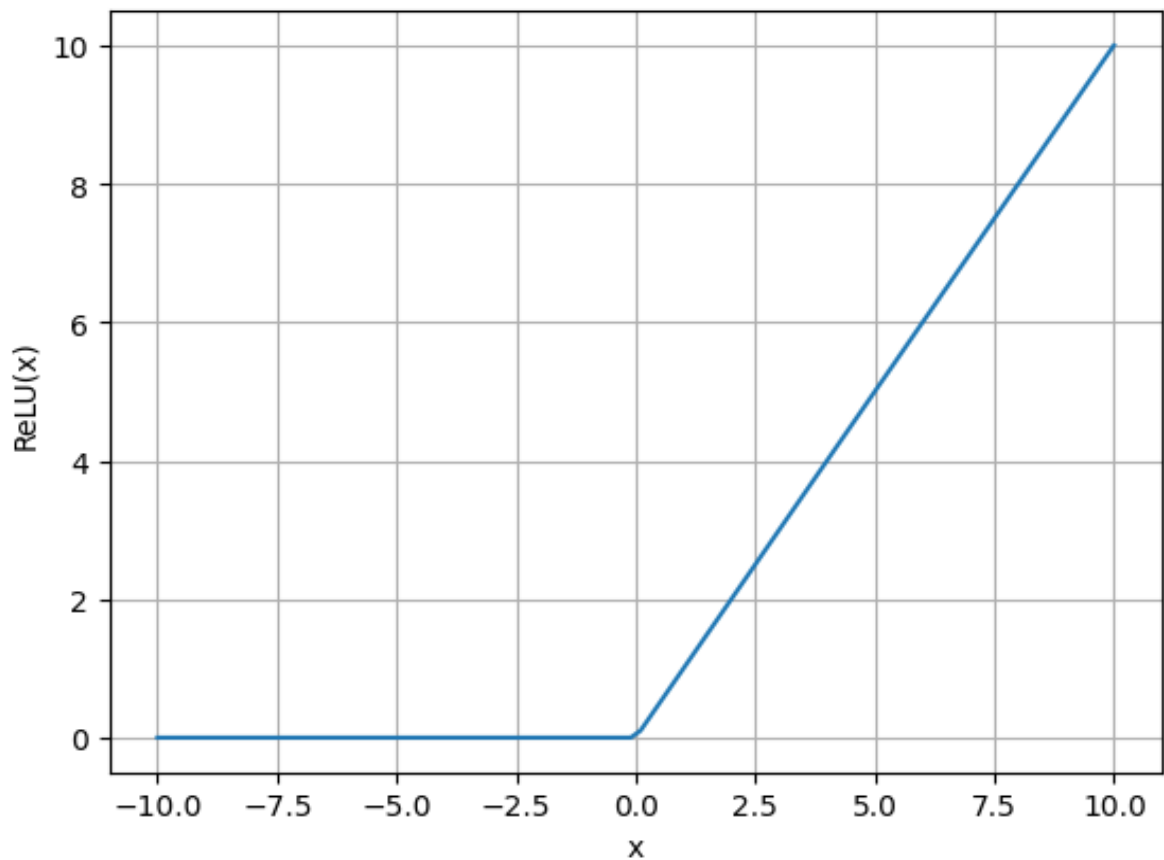


Рис. 2.7. Функція активації ReLU

Головна перевага цієї функції активації її нелінійність. Що дає можливість апроксимувати будь-яку неперервну функцію нейронним моделям з цим типом нейронів. А також, завдяки своїй простій формі, як видно з (1.3) ця нелінійна функція утворена з двох лінійних, для визначення найкращих параметрів використовуються часткові похідні які дуже ефективно рахуються на лінійних частинах функції активації [39].

2.2.2. Згорткові нейронні мережі

Згорткова нейронна мережа (Convolutional Neural Network, ConvNet, CNN) - це математична модель, що має один або декілька *згорткових шарів*. Ці шари дозволяють ефективно обробляти дані з просторовою структурою [40].

Згортковий шар виконує *операції згортки* над вхідними просторовим даними. Для згортки використовуються набір *ядер*, які називають *фільтрами*. Ці фільтри складаються з ваг, які змінюються в ході *оберненого розповсюдження помилки*. Зазвичай використовуються ядра розміру 3×3 , однак вони можуть мати

різний розмір який може визначатись конкретною задачею. Для RGB-зображень використовується фільтр розміром $3 \times 3 \times 3$, що дає 27 ваг і 28 параметрів.

У ході послідовного переміщення ядра по зображенню (зліва на право, зверху вниз) вираховується z , яку потім можна передати в будь-яку функцію активації (рис. 2.8). Результатом згортки є *карта активації* що передається на наступний шар нейронної моделі.

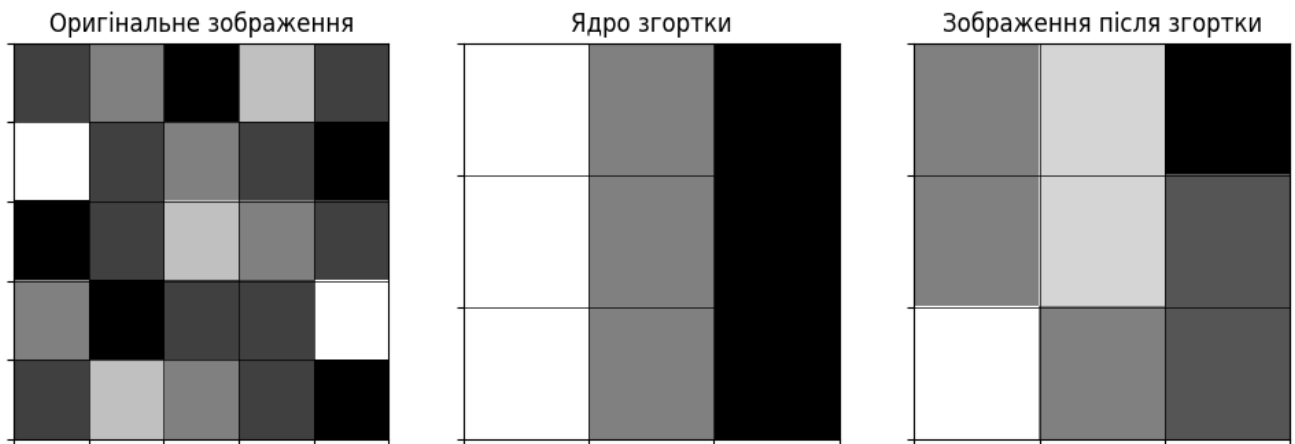


Рис. 2.8. Процес згортки

Зазвичай в одному згортковому шарі використовується декілька фільтрів. Це обумовлено принципом роботи біологічного ока, де прості нейрони реєструють лише прості форми, а подальша обробка передається складним нейронам [41]. Так і в згортковій нейронній мережі кожен фільтр здатен реєструвати лише одну просту форму, наприклад, вертикальну лінію, а інший, горизонтальну.

Кількість фільтрів в згортковому шарі є *гіперпараметром*, тобто таким параметром, що задається безпосередньо дослідником.

2.2.2.1. Нейронна мережа AlexNet

Глибока нейронна мережа що стала популярною внаслідок перемоги конкурсі ImageNet Large Scale Visual Recognition Challenge (ILSVRC) у 2012 році. Ця мережа проклала шлях для подальших досягнень глибокому навчанні, а саме, в таких сферах, як розпізнавання образів, класифікація і автономні системи [42].

На відміну від попередніх підходів, де використовувалися неглибокі штучні нейронні мережі, AlexNet мала декілька згорткових шарів (рис. 2.9),

нормалізацію та функцію активації ReLU. Мережа також використовувала метод дроп-аут (dropout) для запобігання перенавчання, що було нововведенням на той час.

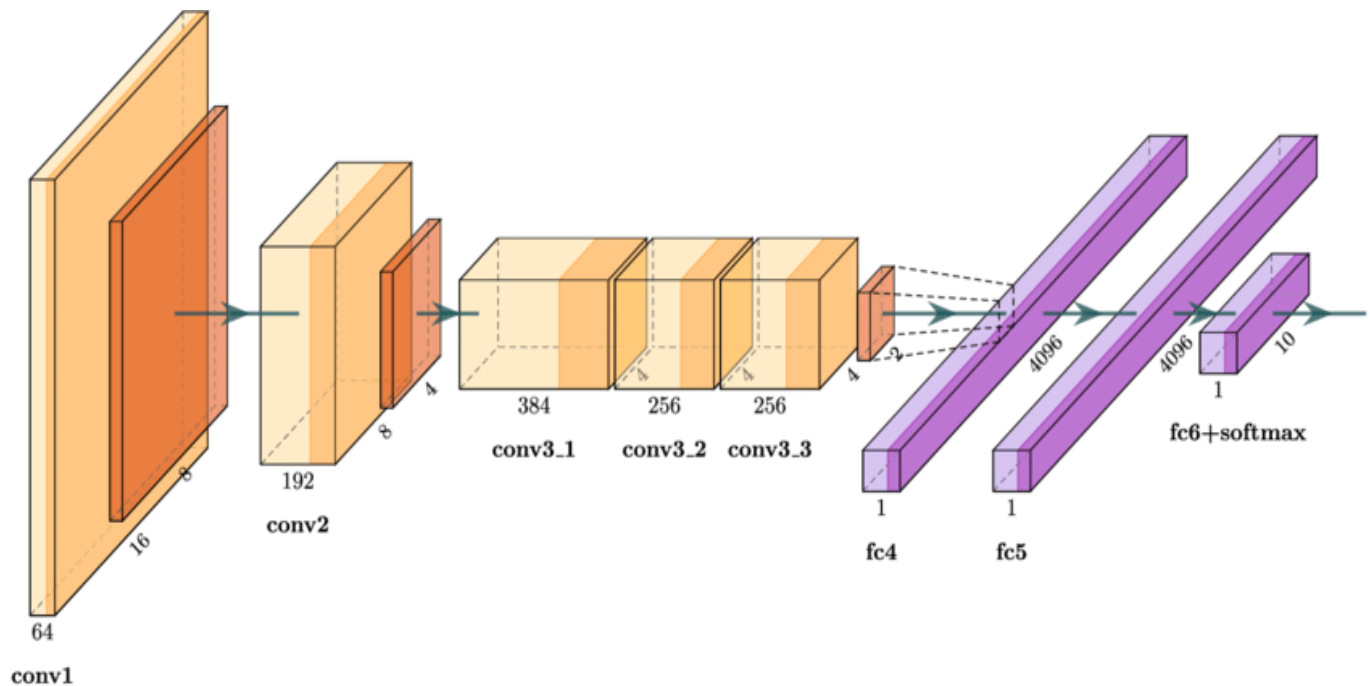


Рис. 2.9. Структура AlexNet

Структура AlexNet складається з п'яти згорткових шарів для виявлення ознак на зображенні та трьох повнозв'язних шарів для виконання класифікації. Ваговий внесок у результати мережі внісла функція активації ReLU, яка прискорює процес навчання шляхом сприяння уникненню проблеми зникання градієнта. А також до складу архітектури входить max-pooling для зменшення розміру просторових характеристик.

AlexNet стала проривом у світі глибокого навчання завдяки своїй архітектурі та здатності розв'язувати завдання класифікації з високою точністю. Однак, на сьогодні вона вважається менш ефективною порівняно з сучасними архітектурами через обмеженість у можливості масштабування.

2.2.2.2. Нейронна мережа VGGNet

Розроблена вченими з Оксфордського університету в 2014 році, VGGNet продемонструвала позитивний вплив збільшення шарів глибокої нейронної

мережі на її якість. Головною особливістю мережі є використання малих фільтрів для згорток (розміром 3×3) для всіх згорткових шарів, що як виявилось дозволяє визначати більш детальні особливості на зображеннях [43].

Відрізняючись поміж попередніх архітектур, VGGNet спирається на використання простих згорток і подвоєні глибини фільтрів після кожного шару згортки. Це дозволяє мережі бути глибшою та водночас більш ефективною у навчанні.

Головним недоліком мережі є необхідність в великих обчислювальних потужностях при навчанні і промисловому використанні. Це може стати серйозною перешкодою для впровадження цієї архітектури у системах з обмеженими ресурсами.

В структурі VGGNet (рис. 2.10) кожен блок складається з двох чи трьох згорткових шарів, після яких виконується max-pooling. Ця дозволяє мережі ефективно виявляти просторові залежності на різних рівнях абстракції. Повнозв'язні шари використовуються для інтеграції інформації з попередніх шарів в цілях класифікації зображень.

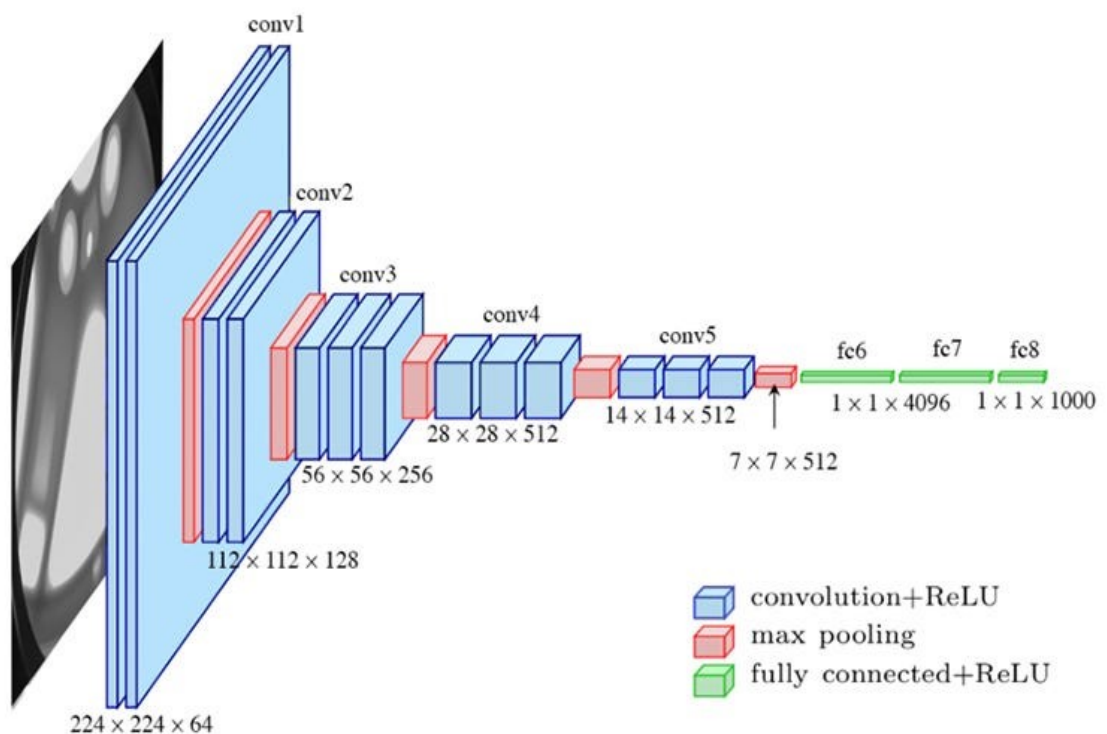


Рис. 2.10. Структура VGGNet

2.2.2.3. Нейронна мережа ResNet

ResNet було запропоновано о 2015 році дослідниками з компанії Microsoft, ключовою перевагою цієї архітектури стала здатність долати проблему зникання градієнтів, що є проблемою в глибоких мережах [44].

Головна новаторська ідея, що лежить в основі мережі, це використання залишкових блоків (residual blocks). Ці блоки дозволяють передавати інформацію оминаючи один або кілька шарів мережі, за рахунок цього, дослідники змогли будувати надзвичайно глибокі моделі, що сприяє досягненню кращих результатів у завданнях комп'ютерного зору. Завдяки залишковим зв'язкам, що дозволяють передавати градієнти безпосередньо через шари, ResNet мінімізує ризик втрати інформації.

Крім перелічених переваг, ResNet демонструє високу продуктивність у багатьох завданнях, включаючи класифікацію зображень, детекцію об'єктів та сегментацію.

Структура ResNet (рис. 2.11) складається з послідовності залишкових блоків, в кожному такому блоці декілька згорткових шарів та нелінійних активацій. Завершальним блоком є повнозв'язний блок що агрегує інформацію з попередніх шарів для розв'язання поставленої задачі комп'ютерного зору.

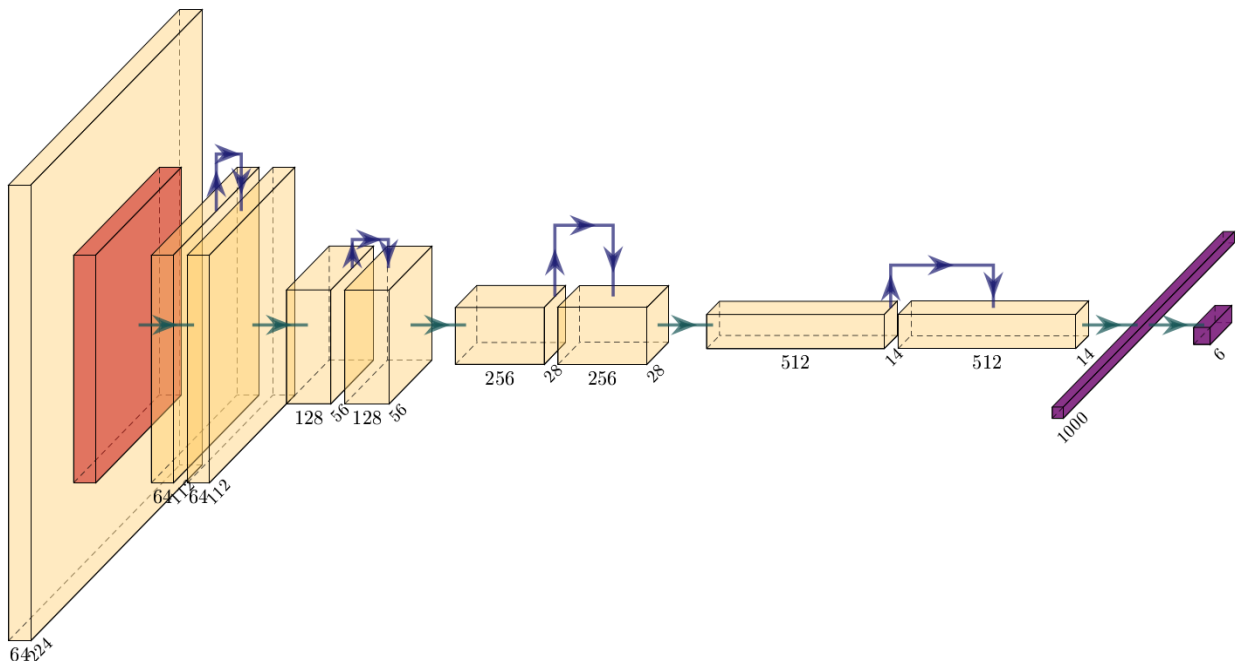


Рис. 2. 11. Структура ResNet

ResNet внесла вагомий внесок в розвиток навчання глибоких нейронних мереж. Мережа демонструє, як інновації в архітектурі можуть сприяти подоланню обмежень наявних підходів, відкриваючи нові можливості для розв'язання складних задач у сфері комп'ютерного зору.. ResNet залишається популярною архітектурою у сучасних дослідженнях та додатках.

2.2.2.4. Нейронна мережа EfficientNet

EfficientNet — це новітня архітектура глибоких нейронних мереж, яка поєднує високу ефективність із чудовою точністю. Розроблена дослідниками Google у 2019 році, вона швидко завоювала популярність серед фахівців у сфері комп'ютерного зору завдяки унікальному підходу до масштабування моделі. Основна мета EfficientNet — створити збалансовану архітектуру, яка максимально використовує обчислювальні ресурси, забезпечуючи при цьому високу продуктивність [61].

Сьогодні, коли обсяги даних зростають із величезною швидкістю, а завдання нейронних мереж стають дедалі складнішими, потрібні моделі, здатні ефективно справлятися з обмеженнями ресурсів. EfficientNet якраз і є такою архітектурою: вона успішно поєднує низьку обчислювальну складність із високою точністю, що робить її ідеальним інструментом для задач від класифікації зображень до розпізнавання об'єктів.

EfficientNet вирізняється з-поміж інших архітектур завдяки своєму інноваційному підходу до масштабування моделі, який називається Compound Scaling. У класичних методах збільшення продуктивності нейронних мереж зазвичай фокусувалися на одному із трьох напрямів:

- Збільшення глибини мережі, додаючи більше шарів.
- Збільшення ширини шарів, додаючи більше нейронів.
- Підвищення роздільної здатності вхідних даних.

Ці методи, хоча й підвищували точність, часто призводили до неефективного використання обчислювальних ресурсів. EfficientNet підходить

до проблеми інакше: вона пропонує комбіноване масштабування, яке враховує всі три аспекти одночасно. Завдяки цьому вдається досягти збалансованого зростання моделі, яке не призводить до зайвого споживання ресурсів.

Ще одним важливим аспектом є використання автоматизованого пошуку архітектур (Neural Architecture Search, NAS). Цей метод допоміг розробникам знайти базову архітектуру EfficientNet-B0, яка стала основою для всіх інших версій (від B1 до B7). Завдяки NAS вдалося забезпечити оптимальну структуру моделі, зважаючи на обмеження обчислювальної потужності.

EfficientNet заснована на базовій архітектурі MobileNetV2, але має низку вдосконалень, які значно підвищують її ефективність. Основні структурні компоненти (рис.) цієї мережі включають [62]:

1. Інверсні резидентні блоки (Inverted Residual Blocks). Ці блоки дозволяють зменшити обчислювальну складність, зберігаючи високу точність. Їхня робота базується на концепції вузьких шарів, що передають найважливіші характеристики даних.
2. Механізм Squeeze-and-Excitation (SE). Цей компонент додає в модель механізм уваги, який допомагає підвищити точність прогнозів. SE-блоки акцентують увагу мережі на найважливіших частинах вхідних даних, зменшуючи вплив менш значущих елементів.
3. Збалансоване масштабування. Головною особливістю EfficientNet є її здатність збільшувати продуктивність завдяки одночасному масштабуванню ширини, глибини та роздільної здатності. Це дозволяє створювати версії моделей, які адаптуються до конкретних вимог проєкту, не перевантажуючи систему.

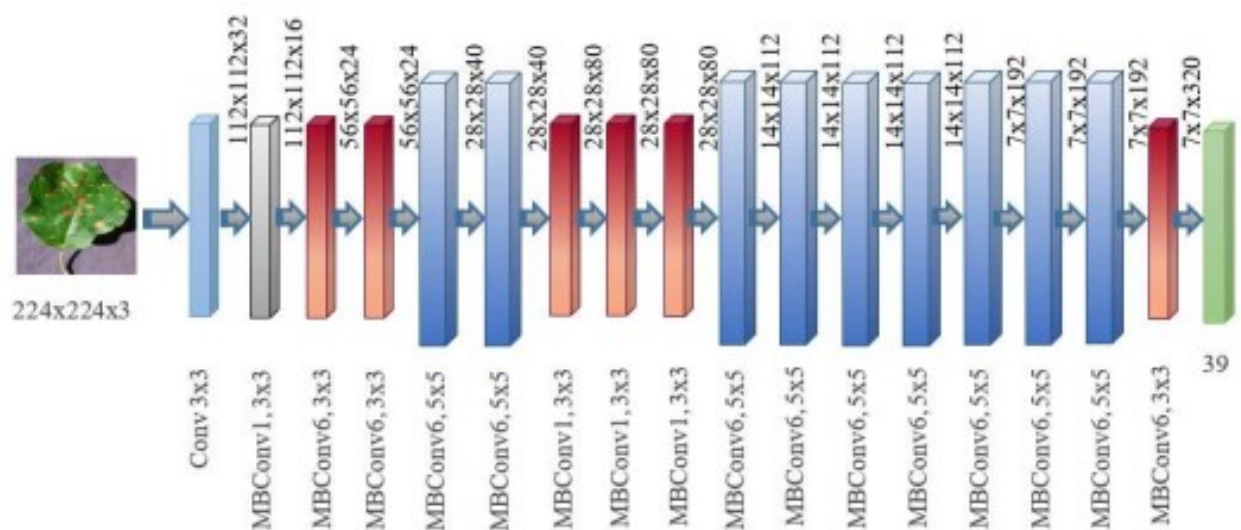


Рис. 2. 12. Структура EfficientNet [62]

EfficientNet має безліч переваг, які роблять її однією з найкращих архітектур для роботи з даними:

1. Висока точність. Завдяки унікальному підходу до масштабування, EfficientNet показує чудові результати на відомих наборах даних. Наприклад, модель EfficientNet-B7 досягла точності понад 84% на ImageNet, перевершивши багато популярних архітектур, як-от ResNet та Inception.
2. Ефективне використання ресурсів. Ця архітектура забезпечує високий рівень продуктивності, водночас використовуючи менше обчислювальної потужності. Це робить її ідеальною для застосувань у мобільних пристроях і середовищах із обмеженими ресурсами.
3. Гнучкість і масштабованість. EfficientNet пропонує моделі різного масштабу — від EfficientNet-B0 до EfficientNet-B7, що дозволяє обирати оптимальний варіант залежно від доступних ресурсів і задач.
4. Автоматизоване проектування. Використання Neural Architecture Search скоротило час і зусилля на розробку, дозволивши створити модель, яка працює максимально ефективно для заданого набору даних.

Попри всі свої переваги, EfficientNet має і певні обмеження:

- Використання EfficientNet вимагає знань сучасних бібліотек для глибокого навчання, таких як TensorFlow або PyTorch. Це може бути

викликом для новачків у сфері машинного навчання.

- Хоча EfficientNet є ефективною під час використання, процес її навчання, особливо для великих моделей (як-от B7), потребує значних обчислювальних ресурсів.
- Комбіноване масштабування, на якому базується EfficientNet, працює добре в межах визначених сценаріїв. Проте для специфічних задач або дуже великих зображень модель може вимагати додаткової оптимізації.

EfficientNet — це передова архітектура, яка встановлює нові стандарти у сфері комп'ютерного зору. Її унікальний підхід до масштабування дозволяє досягти високої точності при мінімальних обчислювальних витратах. Завдяки своїй гнучкості та ефективності, ця модель знаходить широке застосування в різних галузях, від мобільних додатків до складних наукових досліджень.

Хоча реалізація EfficientNet може вимагати значних ресурсів і знань, її переваги значно переважають ці труднощі. Вона є чудовим прикладом того, як інновації в дизайні нейронних мереж можуть сприяти створенню потужних і водночас ефективних рішень для обробки великих обсягів даних.

2.3. Методи штучного збільшення набору даних для нейронних мереж

Однією з важливих проблем при розробці нейронної мережі є недостатня кількість даних [45] та нерівномірність їх розподілу між класами у навчальній вибірці. Брак даних суттєво зменшує ефективність роботи систем штучного інтелекту і може призвести до перенавчання, коли модель добре працює з навчальною вибіркою, але не може ефективно обробляти нові дані. Нерівномірність між класами може виникати в таких сферах як, наприклад, неруйнівний контроль (НК) та діагностика. Це пов'язано з тим, що в реальних умовах виробництва і експлуатації технічних засобів дефекти матеріалів та виробів мають різну природу виникнення, різні геометричні розміри та форму, а також різні частоти появи [46].

Збільшення навчальної вибірки можливе за рахунок наявності необхідних даних у відкритих джерелах інформації чи шляхом проведення додаткових

емпіричних досліджень. Однак зазвичай потрібних навчальних вибірок у відкритому доступі просто не існує, а бюджет та час на подібні дослідження завжди обмежений [47]. Ще однією можливістю розширення набору даних, які можуть бути використані для навчання нейронної мережі, є застосування методів штучного збільшення їх кількості. Одним із таких методів є аугментація даних [48,49].

Аугментація даних – методи підвищення ефективності нейронних мереж шляхом штучного збільшення набору даних, що використовується для навчання. Дані методи широко застосовуються в глибокому навчанні, особливо в галузі комп'ютерного зору, де використовуються з метою підвищення точності нейронних мереж. Крім того застосування аугментації дає змогу розширити обмежені набори даних, збільшити різноманітність даних і уникнути перенавчання систем штучного інтелекту.

Відомо три найбільш поширені шляхи реалізації аугментації: розширення набору даних, доповнення даних на льоту, поєднання попередніх підходів.

Розширення набору даних. Результатом такої аугментації є створення нового розширеного згенерованими випадковими перетвореннями набору даних. Проблема такого підходу в тому, що він дає ілюзію великої кількості даних, однак генеруючи великий навчальний набір з малого не вдається збільшити здатність моделі до узагальнення. Крім того реалізації такого підходу потребує значно більших обчислюваних потужностей.

Доповнення даних на льоту. При такому підході дані випадково замінюються на згенеровані одним з методів аугментації не збільшуючи загальний обсяг навчальної вибірки. Тобто аугментовані дані частково замінюють початкові.

Методи аугментації даних розділяють на традиційні та методи «чорної скриньки» (GAN – Generative Adversarial Networks). Традиційні, в свою чергу, поділяються на групи: методи афінних перетворень зображень (такі, як поворот, масштабування, зсув), кольорової модифікації (підвищення контрастності, зміна балансу білого, підвищення різкості зображення), додавання шуму,

використання методу випадкового стирання та змішування зображень [47]. Аугментацію даних, що заснована на використанні GAN [48,50], зазвичай використовують для генерації нових зображень, які дуже схожі на оригінальні, перенесення текстури (стилю) з одного зображення на інше [49] та змішування зображень [51]. Широка номенклатура наявних процедур аугментації даних обумовлена тим, що застосування кожного методу може по-різному впливати на результат в залежності від задачі та архітектури нейронної мережі [52].

Кращих результатів досягають шляхом підбору комбінацій методів аугментації для конкретної задачі. Актуальним напрямом досліджень є удосконалення алгоритми автоматизованого підбору комбінації методів аугментації [53].

У роботі [54] представлено всебічний огляд методів збільшення кількості зображень для глибокого навчання. Автори поділяють їх на три категорії — без моделей, на основі моделей та на основі політики оптимізації. У категорії без моделей використовуються методи обробки зображень, а в методі на основі моделей використовуються моделі генерації зображень, що навчаються. Підхід, що ґрунтується на політиці оптимізації, спрямований на пошук оптимальних операцій або їх комбінацій. На нашу думку, найбільш перспективними є методи на основі моделей генерації зображень. Метод «SamplePairing», що пропонується в [55], заснований на поєднанні двох випадково обраних зображень у навчальній вибірці, при чому, ці зображення можуть належати до різних класів. Згенероване в такий спосіб зображення відносять до класу першого обраного зображення. Застосування цього методу дало змогу знизити частоту помилок першого рівня на 4,5%. Метод «Міхур», описаний в роботі [56] - має такий самий принцип як і «SamplePairing», однак приналежність до класу визначається в залежності від долі початкових зображень у результуючому. Такі методи добре себе зарекомендували за малих вибірок, бо дають змогу збільшити обсяг навчальної вибірки до N^2 , де N – кількість елементів початкової вибірки. Недоліком таких методів є внесення додаткової похибки, через маніпуляції з приналежністю до класу. До прикладу, авторам доводиться виключати аугментовані дані з вибірки

на останніх епохах навчання.

Автори роботи [57] запропонували два методи збільшення даних: один на основі DWT, а інший на перетворенні постійної добротності CQT. Максимальну продуктивність показав метод збільшення даних заснований на DWT лише в одному з наборів даних. Алгоритм являв собою випадкову заміну коефіцієнтів вейвлет перетворення коефіцієнтами одного з п'яти випадково обраних зображень того ж класу. Запропоновані методи були перевірені на чотирьох наборах даних. Результати досліджень показали необхідність використання додаткових підходів збільшення даних, у тому числі заснованих на глибокому навчанні, таких як GANS, і на набагато більшій кількості наборів даних.

2.4. Методи оцінки нейронних моделей

При розв'язанні будь-якої задачі інженер завжди хоче обрати найкраще рішення. Часто не існує теоретичного чи емпіричного аргументу надавати перевагу певному алгоритму при розв'язанні конкретної задачі. А сучасне розуміння машинного навчання не дозволяє з повною певністю заявити про домінування над іншими. Тому для більшості складних задач необхідно перевіряти роботу декількох конкуруючих моделей, щоб виявити найкращі результати.

Ще у 30-х роках Ларсон зауважив, що машинне навчання та оцінка моделі на однакових даних дає надто оптимістичні прогнози [58]. А Mosteller і Tukey зауважили що тестування алгоритмів на “невідомих” йому даних дає можливість зробити хорошу оцінку продуктивності [59].

При вирішенні реальних задач завжди існує лише деяка обмежена кількість даних, тому було вирішено розбивати наявні дані на дві частини навчальну та валідаційну (тестову) вибірку. Навчальна вибірка використовується при навчанні, валідаційна лише при перевірці готової моделі.

Під час навчання нейронної моделі одним із факторів, які чинять вплив на точність моделі, виступає випадковість у формуванні початкових ваг та розділенні даних на навчальну та валідаційну вибірки [60]. Перша випадковість

долається шляхом жорсткого налаштування алгоритму генератора псевдовипадкових чисел. Для подолання другої використовують принцип кросвалідації (k-fold валідації), який базується на випадковому розбитті всієї вибірки на k частин (рис. 2.13). Послідовно кожна частина відіграє роль валідаційної вибірки, а інші виступають в ролі навчальної. Тому, в результаті одного такого циклу буде отримано k моделей, що дає можливість подолати ефект “щасливої вибірки”, а також дає змогу використовувати методи статистичної обробки сукупностей. Повторні цикли кросвалідації підвищують точність оцінки нейронної мережі. В рамках експерименту було проведено по 3 повторних циклів з розділенням даних на 5 частин, що в результаті принесло по 15 моделей для кожної з порівняльної категорії.



Рис. 2.13. Візуалізація принципу кросвалідації нейронних мереж. Круги представляють дані, прямокутники частини на які розбита вибірка. На кожній ітерації змінюється валідаційна вибірка – сині прямокутники та навчальна вибірка – білі прямокутники.

ВИСНОВКИ ДО РОЗДІЛУ 2

1. Методи комплексування зображень допомагають поєднати зображення різних модальностей в єдину картину, що дає можливість використовувати наявні методи обробки зображень для задач діагностики дорожнього покриття. Головною перевагою використання саме комплексування зображень для цієї цілі, це збільшення інформативності результуючого зображення, що отримується після виконання операції комплексування.

2. У жодному методі комплексування зображень не було використано механізм адаптивного визначення частки оригінальних зображень в результуючому, що дало б змогу підвищити інформативність результуючого зображення, зменшуючи частку не інформативного оригінального зображення.

3. Основною метрикою для визначення інформативності зображень є ентропія Шенона. Ця метрика дозволяє з високою точністю об'єктивно оцінювати інформативність зображень, альтернативою до цієї метрики є опитування, але воно має суб'єктивну похибку і високі витрати часу.

4. У задачах автоматизації діагностики домінують методи на базі глибоких нейронних мереж з згортковими блоками. Це обумовлено високою ефективністю операції згортки при обробці та аналізі зображень, за рахунок масової обробки пакету пікселів замість попиксельної обробки.

5. Основним методом боротьби з перенавчанням та браком даних для навчання нейронної мережі, що не потребує розробки нової архітектури, є аугментація даних. Ці методи здатні створювати нові дані за рахунок перетворення наявних, що дозволяє збільшити різноманіття навчальної вибірки. Дослідження показують що використання комбінацій різних методів аугментації підвищує ефективність навчання та зменшує перенавчання нейронних мереж. Наразі більшість існуючих методів аугментації розроблялись для зображень видимого спектру, тому спеціалізованих методів під, наприклад, інфрачервоний спектр зображень невелика кількість, що ускладнює розробку комбінацій аугментації, які забезпечують найбільший ефект для таких зображень.

Список використаних джерел до розділу 2

- [1] Сторожик, Д. В., & Муравйов, О. В. (2020). Застосування мультиспектрального комплексування зображень у неруйнівному контролі.
- [2] Kaur, H., Koundal, D., & Kadyan, V. (2021). Image fusion techniques: a survey. *Archives of computational methods in Engineering*, 28(7), 4425-4447.
- [3] Sharma, M. (2016). A review: Image fusion techniques and applications. *International Journal of Computer Science and Information Technologies*, 7(3), 1082–1085.
- [4] Гривачевський, А. П., & Прудис, І. Н. (2018). Підвищення інформативності мультиспектральних систем моніторингу шляхом комплексування зображень видимого та інфрачервоного діапазонів.
- [5] Li, S., Kang, X., Fang, L., Hu, J., & Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. *information Fusion*, 33, 100-112.
- [6] Maruthi, R., & Lakshmi, I. (2017). Multi-focus image fusion methods—a survey. *Comput Eng*, 19(4), 9-25.
- [7] Meher, B., Agrawal, S., Panda, R., & Abraham, A. (2019). A survey on region based image fusion methods. *Information Fusion*, 48, 119-132.
- [8] Yang, J., Ma, Y., Yao, W., & Lu, W. T. (2008). A spatial domain and frequency domain integrated approach to fusion multifocus images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(PART B7).
- [9] Таршин, В. А., Танцюра, О. Б., & Куравський, М. В. (2021). Шляхи підвищення імовірності виявлення та розпізнавання об'єктів на зображеннях різноспектральних оптико-електронних систем розвідки. *Наука і техніка Повітряних Сил Збройних Сил України*, (3 (44)), 141-146.
- [10] Morris, C., & Rajesh, R. S. (2014). Survey of spatial domain image fusion techniques. *Int J Adv Res Comput Sci Eng Inf Technol*, 2(3), 249-254.
- [11] Jasiunas, M. D., Kearney, D. A., Hopf, J., & Wigley, G. B. (2002, December). Image fusion for uninhabited airborne vehicles. In *2002 IEEE International Conference on Field-Programmable Technology, 2002.(FPT). Proceedings.* (pp. 348-351). IEEE.
- [12] Сторожик, Д. В. (2020). Тепловізійна система з мультиспектральним комплексуванням зображення.
- [13] Banu, R. S. (2011). Medical Image Fusion by the analysis of Pixel Level Multi-sensor Using Discrete Wavelet Transform. In *Proceedings of the national conference on emerging trends in computing science* (pp. 291-297).
- [14] Song, L., Lin, Y., Feng, W., & Zhao, M. (2009, May). A novel automatic weighted image fusion algorithm. In *2009 International Workshop on Intelligent*

Systems and Applications (pp. 1-4). IEEE.

- [15] Yonghong, J. (2012). Fusion of landsat TM and SAR images based on principal component analysis. *Remote Sensing Technology and Application*, 13(1), 46-49.
- [16] Rajangam, V., Sangeetha, N., & Gnanasivam, P. (2021, May). Multimodal medical image fusion in DWT domain using Otsu's Thresholding. In 2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP) (pp. 109-114). IEEE.
- [17] Sujatha, M., Srilekha, G., Tina, K., Tulasi, T. S., & Harish, K. (2020). Image enhancement using wavelet based image fusion and power law transform. *Journal of Computational and Theoretical Nanoscience*, 17(5), 2405-2408.
- [18] Saleem, A., Beghdadi, A., & Boashash, B. (2012). Image fusion-based contrast enhancement. *EURASIP Journal on Image and Video Processing*, 2012, 1-17.
- [19] Vekkot, S. (2010). Wavelet based medical image fusion using filter masks. In *Trends in Intelligent Robotics: 13th FIRA Robot World Congress, FIRA 2010, Bangalore, India, September 15-17, 2010. Proceedings 13* (pp. 298-305). Springer Berlin Heidelberg.
- [20] Wang, Z., & Bovik, A. C. (2002). A universal image quality index. *IEEE signal processing letters*, 9(3), 81-84.
- [21] Wang, Z., Ziou, D., Armenakis, C., Li, D., & Li, Q. (2005). A comparative analysis of image fusion methods. *IEEE transactions on geoscience and remote sensing*, 43(6), 1391-1402.
- [22] Cetin, M., & Tepecik, A. (2016). Intensity–hue–saturation-based image fusion using iterative linear regression. *Journal of Applied Remote Sensing*, 10(4), 045019-045019.
- [23] Mishra, D., & Palkar, B. (2015). Image fusion techniques: a review. *International Journal of Computer Applications*, 130(9), 7-13.
- [24] Mandhare, R. A., Upadhyay, P., & Gupta, S. (2013). Pixel-level image fusion using brovey transform and wavelet transform. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(6), 2690-2695.
- [25] Smith, L. I. (2002). A tutorial on principal components analysis.
- [26] Patil, U., & Mudengudi, U. (2011, November). Image fusion using hierarchical PCA. In 2011 international conference on image information processing (pp. 1-6). IEEE.
- [27] Olkkonen, H., & Pesola, P. (1996). Gaussian pyramid wavelet transform for multiresolution analysis of images. *Graphical Models and Image Processing*, 58(4), 394-398.

- [28] Shen, J., Zhao, Y., Yan, S., & Li, X. (2014). Exposure fusion using boosting Laplacian pyramid. *IEEE Trans. Cybern.*, 44(9), 1579-1590.
- [29] Chandrasekhar, C., Viswanath, A., & Narayana Reddy, S. (2013). Implementation of image fusion technique using DWT for micro air vehicle applications. *Field-Programmable Gate Array (FPGA)*, 4(8), 307-315.
- [30] Jiang, D., Zhuang, D., Huang, Y., & Fu, J. (2011). Survey of multispectral image fusion techniques in remote sensing applications. In *Image fusion and its applications* (Vol. 1, pp. 1-22). IntechOpen.
- [31] Wu, D., Yang, A., Zhu, L., & Zhang, C. (2014). Survey of multi-sensor image fusion. In *Life System Modeling and Simulation: International Conference on Life System Modeling and Simulation, LSMS 2014, and International Conference on Intelligent Computing for Sustainable Energy and Environment, ICSEE 2014, Shanghai, China, September 20-23, 2014, Proceedings, Part I* (pp. 358-367). Springer Berlin Heidelberg.
- [32] Naidu, V. P. S. (2012). Discrete cosine transform based image fusion techniques. *Journal of Communication, Navigation and Signal Processing*, 1(1), 35-45.
- [33] Desale, R. P., & Verma, S. V. (2013, February). Study and analysis of PCA, DCT & DWT based image fusion techniques. In *2013 international conference on signal processing, image processing & pattern recognition* (pp. 66-69). IEEE.
- [34] Chang, C. I., Du, Y., Wang, J., Guo, S. M., & Thouin, P. D. (2006). Survey and comparative analysis of entropy and relative entropy thresholding techniques. *IEE Proceedings-Vision, Image and Signal Processing*, 153(6), 837-850.
- [35] Tsai, D. Y., Lee, Y., & Matsuyama, E. (2008). Information entropy measure for evaluation of image quality. *Journal of digital imaging*, 21, 338-347.
- [36] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- [37] Han, J., & Moraga, C. (1995, June). The influence of the sigmoid function parameters on the speed of backpropagation learning. In *International workshop on artificial neural networks* (pp. 195-201). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [38] Hahnloser, R., & Seung, H. S. (2000). Permitted and forbidden sets in symmetric threshold-linear networks. *Advances in neural information processing systems*, 13.

- [39] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (pp. 807-814).
- [40] Jon, K., Grant, B., & Aglae, B. (2019). Deep Learning Illustrated.
- [41] Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J physiol*, 148(3), 574-591.
- [42] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [43] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [44] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [45] Cupiał, P., & Łacny, Ł. (2015). Neural network control design considerations for the active damping of a smart beam. *Journal of Theoretical and Applied Mechanics*, 53(4), 767-774.
- [46] Galagan, R. M., & Momot, A. S. (2017, May). The use of ART-2 neural network for processing information signals of non-destructive testing. In 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON) (pp. 981-985). IEEE.
- [47] Khalifa, N. E., Loey, M., & Mirjalili, S. (2022). A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artificial Intelligence Review*, 55(3), 2351-2377.
- [48] Chaikovskiy, O., Volokyta, A., Kyrianov, A., & Loutskii, H. (2021). Data augmentation method using generative adversarial networks. *Technical Sciences and Technology*, (2(24)), 83–91. [https://doi.org/10.25140/2411-5363-2021-2\(24\)-83-91](https://doi.org/10.25140/2411-5363-2021-2(24)-83-91)
- [49] Jiang, Y., Zhu, B., & Xie, B. (2020, October). Remote Sensing Images Data Augmentation Based on Style Transfer under the Condition of Few Samples. In *Journal of Physics: Conference Series* (Vol. 1653, No. 1, p. 012039). IOP Publishing.
- [50] Giles, M. (2018). The GANfather: The man who's given machines the gift of imagination. *MIT Technology Review*, 121(2), 48-53.
- [51] Wu, H., Zheng, S., Zhang, J., & Huang, K. (2019, October). Gp-gan: Towards realistic high-resolution image blending. In Proceedings of the 27th ACM international conference on multimedia (pp. 2487-2495).

- [52] Mikołajczyk, A., & Grochowski, M. (2018, May). Data augmentation for improving deep learning in image classification problem. In 2018 international interdisciplinary PhD workshop (IIPhDW) (pp. 117-122). IEEE.
- [53] Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. (2018). Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv:1805.09501.
- [54] Xu, M., Yoon, S., Fuentes, A., & Park, D. S. (2023). A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognition*, 137, 109347.
- [55] Inoue, H. (2018). Data augmentation by pairing samples for images classification. arXiv preprint arXiv:1801.02929.
- [56] Zhang, H. (2017). mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412.
- [57] Nanni, L., Paci, M., Brahnam, S., & Lumini, A. (2021). Comparison of different image data augmentation approaches. *Journal of imaging*, 7(12), 254.
- [58] Larson, S. C. (1931). The shrinkage of the coefficient of multiple correlation. *Journal of Educational Psychology*, 22(1), 45.
- [59] Mosteller, F., & Tukey, J. W. (1968). Data analysis, including statistics. *Handbook of social psychology*, 2, 80-203.
- [60] Wong, S. C., Gatt, A., Stamatescu, V., & McDonnell, M. D. (2016, November). Understanding data augmentation for classification: when to warp?. In 2016 international conference on digital image computing: techniques and applications (DICTA) (pp. 1-6). IEEE.
- [61] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.
- [62] Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, 61, 101182.

РОЗДІЛ 3. Розробка та оцінка методів аугментацій та комплексування зображень

3.1. Розробка методу аугментації на базі комплексування зображень з вейвлет перетворенням

Як було показано в першому та другому розділах існує проблема високої ціни збору високо якісних наборів даних для навчання нейронних мереж. З метою розв'язання даної проблеми було розроблено методику аугментації даних для збільшення їх кількості та різноманіття. На цей час існує невелика кількість методів аугментації спеціально розроблених для тепловізійних знімків і, як було висвітлено вище, комбінацій різних методів здатні значно підвищити ефективність навчання. Тому, пропонується розглянути використання комплексування зображень на базі вейвлет перетворення, як одного з найкращих наявних методів, для трансформації теплових знімків з ціллю реалізації методів аугментації. Запропонований метод можна віднести до методів змішування зображень, однак, на відміну від описаних вище процедур змішування зображень, комплексування не допускає змішування зображень різних об'єктів одного класу та зміни природних ознак одного об'єкта у змішаному зображенні. Прикладом таких змін може бути використання зображень різної модальності, з різним фокусом, зображення зроблені з часовою затримкою. В проведених дослідженнях використані зміни зображення, які обумовлені властивістю нагрітого тіла змінювати свою температуру в часі та тремтінням рук оператора [1]. Однією з головних переваг таких методів є можливість збагачення даних завдяки збільшенню їх інформативності [2].

Процес перетворення зображення для запропонованого методу аугментації складається з наступних етапів:

1. Два зображення одного і того самого об'єкта, з невеликими природними змінами, розкладаються на вейвлет коефіцієнти:

$$W_A = \text{wavelet}(A); W_B = \text{wavelet}(B), \quad (3.1)$$

де A і B - початкові зображення, *wavelet* - вейвлет перетворення, W_A та W_B - отримані вейвлет коефіцієнти зображень.

2. Виконується зворотне вейвлет перетворення для коефіцієнтів, отриманих в результаті усереднення вейвлет коефіцієнтів вхідних зображень:

$$F_{aug} = inverse_{wavelet} \left(\frac{(W_A + W_B)}{2} \right), \quad (3.2)$$

де $inverse_{wavelet}$ - зворотне вейвлет перетворення, W_A та W_B - вейвлет коефіцієнти вхідних зображень, F_{aug} - результат аугментації.

Схематично процес перетворення зображення показано на рис. 3.1.

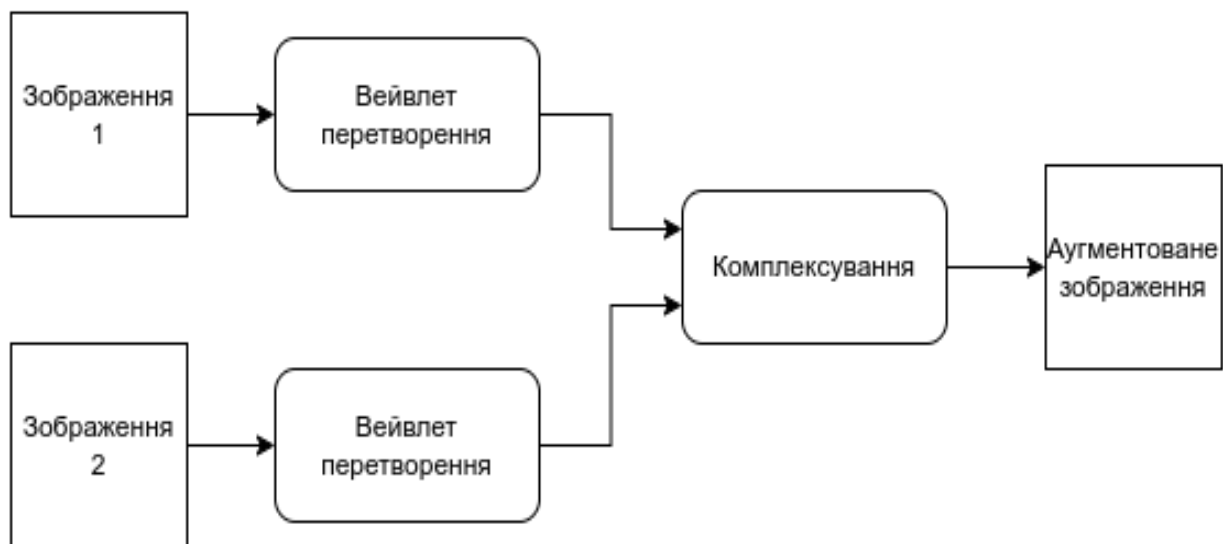


Рис. 3.1. Процес аугментації на базі комплексування зображень з вейвлет перетворенням

3.2. Оцінка методу аугментації на базі комплексування зображень з вейвлет перетворенням

3.2.1. Опис даних для проведення експерименту

Для реалізації експерименту з дослідження ефективності комплексування термограм об'єктів для аугментації навчальних даних було отримано 400 інфрачервоних (ІЧ) зображень чотирьох категорій від тест-об'єктів з імітацією дефектів типу порожнин у формі – круг, квадрат, трикутник, та без дефекту. Тестові зразки було виготовлено з гетинаксу, глибина залягання дефектів від поверхні становить 2 мм. Приклади отриманих термограм тестових зразків наведено на рис. 3.2.

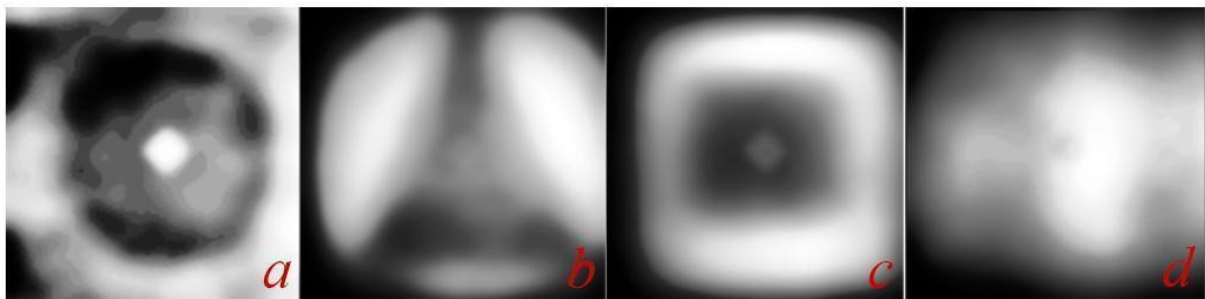


Рис. 3.2. Приклади теплових зображень різних категорій: *a, b, c* – тестові зразки з імітацією форми дефектів типу круг, трикутник та квадрат відповідно; *d* – бездефектний тестовий зразок

Кожна категорія даних містить 100 зображень, які були отримані серіями за наступною процедурою. Зразки нагрівали ІЧ електрообігрівачем, після чого із затримкою (декілька десятків секунд) та без штатива за допомогою тепловізора робили знімки, що обумовило наявність невеликих відмінностей між зображеннями однієї серії.

3.2.2. Реалізація експерименту

Експеримент полягав у порівняльному аналізі декількох методів аугментації, а саме, методу на основі додавання шуму, SamplePairing, методу випадкової заміни коефіцієнтів DWT та методу комплексування зображень на основі DWT [2]. Для контрольного зразка експеримент був повторений без застосування аугментації.

Як методику аналізу оцінок точності нейронних мереж було обрано крос валідацію, через невелику кількість даних вона здатна збільшити репрезентативність експерименту.

В якості архітектури нейронної мережі обрано структуру AlexNet [3], машинне навчання проходить 30 епох. AlexNet – класична архітектура нейронної мережі для систем комп'ютерного зору, яка забезпечує достатню точність для розв'язання простих задач (наприклад, визначення форми об'єктів). Разом з цим ця архітектура має порівняно невелику кількість параметрів (60 мільйонів), щоб бути ефективною при малих навчальних вибірках, і не потребує великих обчислювальних потужностей. Мета навчання – класифікація чотирьох типів

тепловізійних зображень відповідно до чотирьох категорій дефектів на ІЧ знімках. Метрикою якості було обрано точність класифікації (3.3) зображень нейронними мережами. Під час навчання після кожної епохи відбувалась перевірка точності варіаційної вибірки для збереження ваг мережі з найкращою точністю.

$$accuracy = \frac{N_{true}}{N}, \quad (3.3)$$

де N – загальна кількість зображень, N_{true} – загальна кількість правильно класифікованих зображень.

3.2.3. Аналіз результатів дослідження

Результати експерименту та оцінки точності класифікації нейронними мережами валідаційної вибірки в залежності від методу аугментації, а саме, додавання шуму, SamplePairing, методу випадкової заміни коефіцієнтів DWT та методу комплексування зображень на основі DWT, наведені на рис. 3.3.

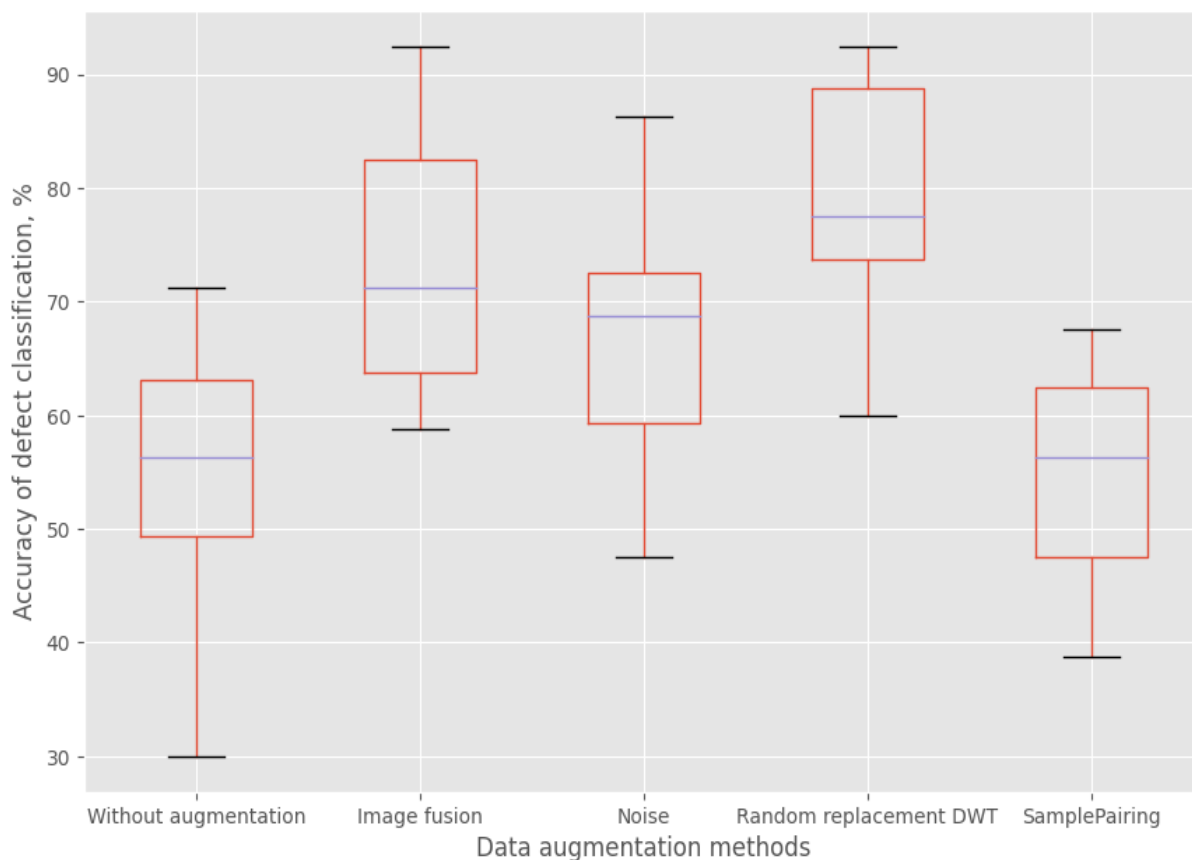


Рис. 3.3. Коробкові діаграми точності класифікації валідаційної вибірки в залежності від типу аугментації навчальної вибірки. Нижня межа замкненої прямокутної області

відповідає першому квантилю, верхня межа – третьому квантилю, центральна лінія – медіані.

Як видно з діаграм, характер розподілу точності класифікації тестової вибірки є негауссовим. Через ненормальний розподіл вибірок і потенційну наявність викидів, як метод порівняння цих вибірок, був обраний критерій Kruskal-Wallis [4], який є непараметричним критерієм перевірки гіпотези, про рівність медіан декількох вибірок. Для цього критерію нульовою гіпотезою H_0 є відсутність значних відмінностей між сукупностями точностей класифікації..

Статистика критерію визначалась за формулою:

$$H = \frac{12}{N(N-1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1), \quad (3.4)$$

де k – кількість груп використаних при порівнянні, N – загальний розмір сукупності, n_i – розмір сукупності для i -ї групи, R_i – загальна кількість рангів i -ї групи.

Для чотирьох ступенів свободи ($df = 4$) значення цієї статистики склало $H = 36.17$; статистична значущість (Sig.) отримана близькою до 0.

Отриманий рівень значущості менший за промисловий стандарт p -значення на рівні 0,05, тому нульова гіпотеза H_0 про відсутність значних відмінностей між вибірками відхилена. Отже, відмінності між вибірками значущі. Тобто, залежність між методом аугментації та точністю класифікації статистично доведена.

Розглянемо результати для кожної категорії в порядку зростання медіани точності.

Еталонна вибірка навчалась на даних без використання аугментації. Її медіанна точність становить 56%.

Метод SamplePairing, отримав значення точності класифікації за медіаною 56%. Можемо зробити висновок що цей метод майже не покращив точність класифікації, хоча її мінімальне значення значно вище ніж в еталонному варіанті.

Для методу зашумлення медіана склала 68.7% точності. Цей метод аугментації був обраний для отримання еталонних показників точності при використанні стандартних методів аугментації.

При використанні комплексування зображень отримано медіану 71.2%. В

порівнянні з методом зашумлення медіана точності виросла на 2.5%, а у порівнянні з випадком класифікації без аугментації – на 15,2%.

Для методу випадкової заміни коефіцієнтів DWT отримано медіанний показник точності 77.5%. В цьому випадку отримано приріст медіани точності на 5.3% в порівнянні з комплексуванням, однак якщо переглянути рис. 2, побачимо, що показники точності двох категорій майже повністю перетинаються, тому доцільно перевірити чи є між ними статистично значуща відмінність. Для цього скористаємося критерієм Kruskal-Wallis, але тепер лише для двох вибірок. Н-статистика склала 2.24, статистична значущість (Sig.) на рівні 0.14. р-значення значно більше 0.05, тому нульову гіпотезу про відсутність значних відмінностей між вибірками не можна відкинути. Отже, статистично значущої відмінності між вибірками точностей двох категорій немає.

Метод комплексування завдяки таким своїм властивостям, як здатність об'єднувати мультимодальні дані та можливість підвищення інформативності для результуючих зображень [5], а також, завдяки меншій алгоритмічній складності в порівнянні з методом випадкової заміни коефіцієнтів, може використовуватись в якості більш ефективного інструменту для аугментації даних та їх збагачення.

Для більш детального аналізу можливостей застосування комплексування зображень в якості метода аугментації даних доцільно дослідити залежність точності класифікації від кількості аугментованих зображень в навчальній вибірці. З цією метою було навчено нейронні мережі за методикою кросвалідації для вибірок зі змінним відсотком аугментованих зображень, по 30 мереж для кожного відсотка. Результати класифікації зображенні в залежності від кількості аугментованих зображень навчальної вибірки, що змінювалась з кроком 20%, наведено на рис. 3.4.

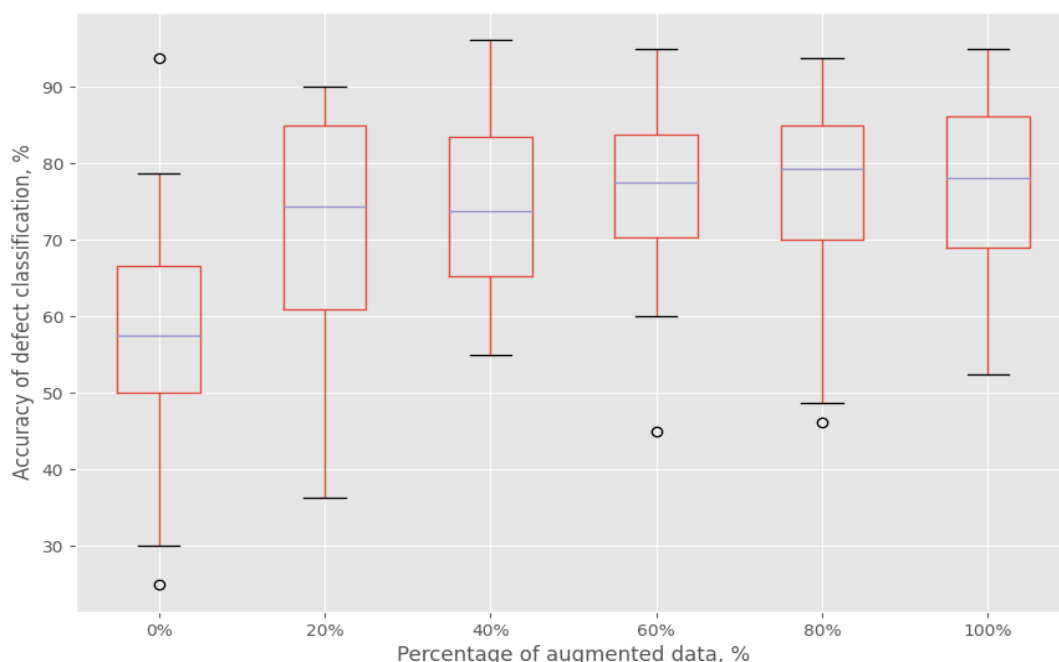


Рис. 3.4. Коробкові діаграми точності класифікації валідаційної вибірки в залежності від кількості аугментованих зображень навчальної вибірки.

При 20% аугментованих зображень від загальної кількості відбувається різкий ріст медіани, приріст склав 17% в порівнянні з вибіркою без використання аугментації. Подальші збільшення не призвели до значного росту точності. Найкращий результат отримала вибірка з 80% аугментованих зображень (приріст 22%).

Отримані результати частково вирішують наявний недолік методу комплексування, а саме, для отримання значного приросту точності нема потреби мати декілька зображень для кожного об'єкту в навчальній вибірці. В деяких областях застосування, наприклад, при автоматизованому контролі дорожнього покриття, це є досить ресурсовитратним.

3.3. Розробка методу комплексування зображень з адаптивним визначенням ваг

Аналізуючи існуючі методи комплексування та отримані дані з датчиків постає питання відсутності визначення частки кожного зображення в результуючому. Так коли частка завад на видимому зображенні більша логічніше за основу брати теплове зображення і навпаки. Як показано в розділі 2 існують методи які здатні враховувати вагові коефіцієнти при комплексуванні

зображень, однак, проблему ефективного визначення вагових коефіцієнтів не вирішено. Тому пропонується розглянути метод комплексування зображень з адаптивним підбором вагових коефіцієнтів. Який поєднує в собі адаптивні визначання ваг використовуючи згорткову нейронну мережу, принцип вагової регуляції та комплексування на базі вейвлет перетворення.

Перевагами даного підходу є:

Покращена здатність збереження інформації. Запропонований метод використовує адаптивне керування вагами на основі вейвлет-коефіцієнтів, яке дозволяє оптимально зберігати важливу інформацію з обох вхідних зображень. Нейронна мережа може визначити, які особливості в кожному зображенні є найважливішими, та відповідно керувати ваговими коефіцієнтами, що збільшує інформативність злитого зображення.

Гнучкість та адаптивність. Завдяки використанню нейронних мереж, метод може адаптуватися до різних типів даних та сценаріїв застосування без потреби в ручному налаштуванні параметрів для кожного окремого випадку. Це робить метод більш гнучким і широко застосовним у порівнянні з традиційними методами, які часто вимагають специфічних налаштувань для кожного нового типу зображень або зміни умов.

Покращення якості зображення. Використання вейвлет-перетворення допомагає ефективно відокремлювати корисну інформацію від шуму. Нейронна мережа може навчитися ідентифікувати та виключати шумові компоненти при обчисленні ваг для коефіцієнтів, що сприяє створенню чистішого і виразнішого злитого зображення.

Підтримка складних залежностей. Нейронні мережі здатні моделювати складні нелінійні залежності між вхідними даними, що є корисним при обробці зображень з високим ступенем складності та різноманітністю особливостей.

Математично даний метод можна описати таким чином:

1. Спочатку виконуємо вейвлет-перетворення для кожного зображення яке ми хочемо комплексувати:

$$W_A = \text{wavelet}(A); W_B = \text{wavelet}(B), \quad (3.5)$$

де A і B - початкові зображення, *wavelet* - вейвлет перетворення,

2. Нейронна мережа NN використовується для визначення вагових коефіцієнтів на основі вейвлет-коефіцієнтів:

$$\alpha, \beta = NN(W_A, W_B; \theta), \quad (3.6)$$

де α та β є ваговими коефіцієнтами для кожного вейвлет-коефіцієнта W_A та W_B відповідно, θ позначає параметри нейронної мережі, які оптимізуються для максимізації інформативності результуючого зображення F .

Цільова функція для оптимізації:

$$\theta \star = \operatorname{argmax}_{\theta} H(F), \quad (3.7)$$

де $H(F)$ — ентропія Шенона зображення F .

3. Злиття зображень здійснюється з використанням вагових коефіцієнтів:

$$F = \operatorname{inverse}_{\text{wavelet}}(\alpha \cdot W_A + \beta \cdot W_B), \quad (3.8)$$

де $\operatorname{inverse}_{\text{wavelet}}$ - зворотне вейвлет перетворення, яке перетворює вейвлет-коефіцієнти назад у просторовий домен для формування результуючого зображення F , W_A та W_B - вейвлет коефіцієнти вхідних зображень.

3.4. Розробка нейронної мережі для визначення ваг

Через те що вхідний сигнал має мультимодальний характер пропонується використати мультимодальний підхід з окремими згортковими блоками для кожного масиву вейвлет-коефіцієнтів для визначення базових ознак. Після цього ці ознаки передаються на спільний повнозв'язний шар для обчислення коефіцієнта, який використовується при комплексуванні зображень з метою збільшення ентропії Шеннона результуючого зображення.

Як згортковий блок було обрано згорткову частину моделі ResNet, попередньо навчену на великих об'ємах даних виявляти базові ознаки з зображення. Ця мережа відома своєю ефективністю на невеликих об'ємах даних коштом блоків з обхідними зв'язками.

Схему детальної структури запропонованої нейромережевої моделі наведено на рис 3.5.

Структура мережі ResNet розглядалась в попередньому розділі, тому

висвітливо більш детально блок повнозв'язних нейронів. Він складається з п'яти шарів, в кожному наступному на половину менше нейронів ніж попередньому для узагальнення, кількість нейронів в кожному шарі відповідно складала 1024, 512, 206, 103 та 1. Для всіх шарів, окрім останнього використовувалась функція активації ReLU, для останнього шару була використана сигмоїдна функція для гарантування вихідного сигналу в діапазоні від 0 до 1. Для всіх шарів, крім останнього, була застосована процедура пакетної нормалізації для зменшення впливу затухання градієнтів та процедура Dropout, яка випадково обнуляє деякі виходи для зменшення перенавчання та збільшення узагальнення при навчанні.

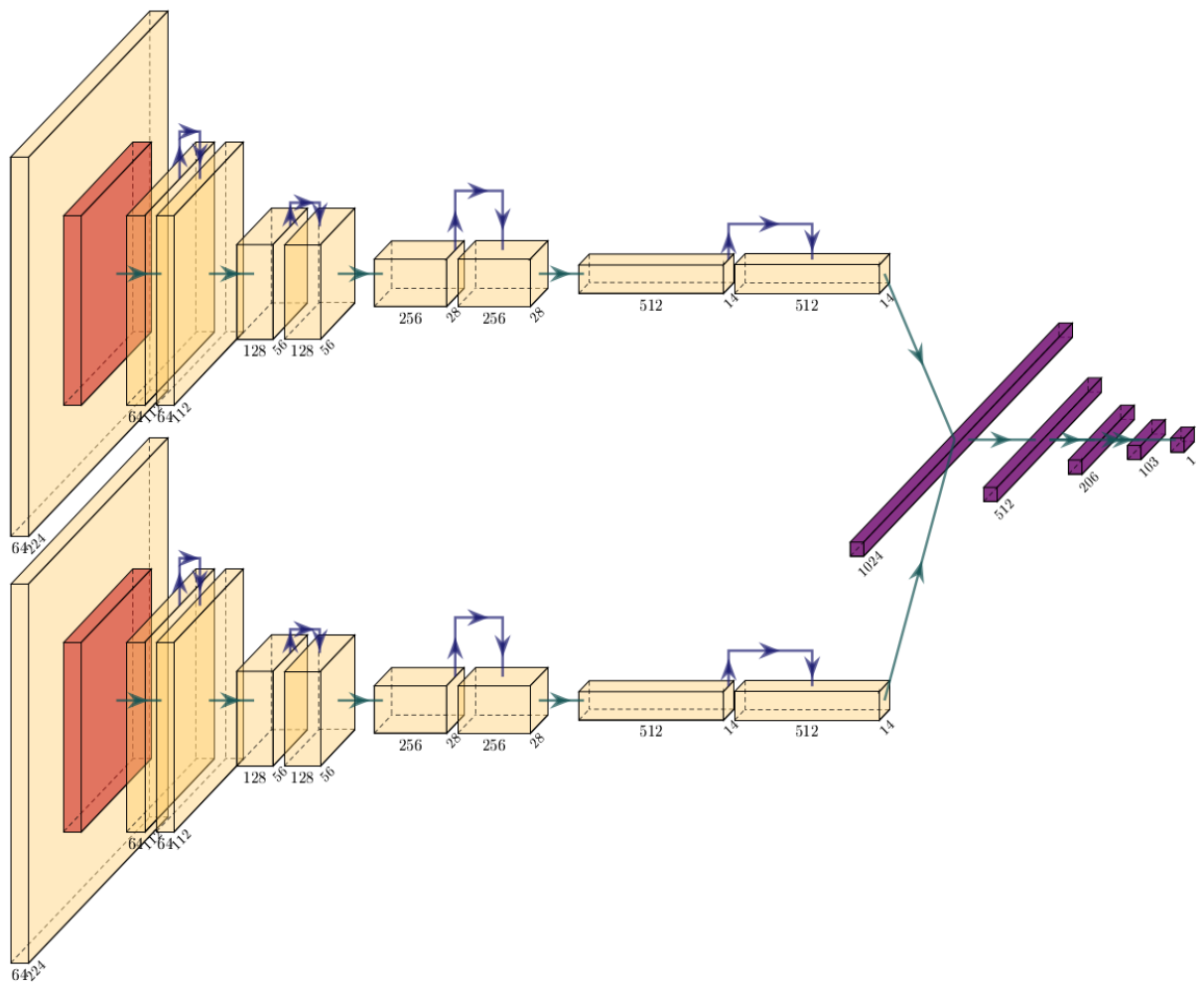


Рис 3.5. Схема архітектури запропонованої нейронної мережі для визначення ваг при комплексуванні зображень

Загальна кількість параметрів склала 24,064,489, а загальний розмір 16516.88 Мегабайтів. Така архітектура експериментально показала свою найбільшу ефективність. Згорткових шарів достатньо для виділення ознак з двох наборів вейвлет-коефіцієнтів, а кількості нейронів в повнозв'язному блоці

достатньо для визначення найкращого вагового коефіцієнта для комплексування запропонованим методом.

Як функцію втрат (loss) було обрано доповнення до одиниці нормалізовано ентропії Шенона для зображень у відтінках сірого:

$$loss = 1 - H \frac{(F)}{lo} g_2 256, \quad (3.9)$$

де $H(F)$ — ентропія Шенона комплексованого зображення, $\log_2 256$ — максимально можлива ентропія Шенона для зображень у відтінках сірого.

Так само експериментально було відібрано швидкість навчання (learning rate) на рівні 0.0001, обсяг партії (batch size) при навчанні 25 та кількість епох (epochs) навчання 5. При інших конфігураціях гіперпараметрів мережа не збігалась і швидко перенавчалась.

Зміну функції втрат для кожної епохи для навчальної та варіаційної вибірки показано на рис. 3.6, а також зміну ентропії Шенона на рис. 3.7.

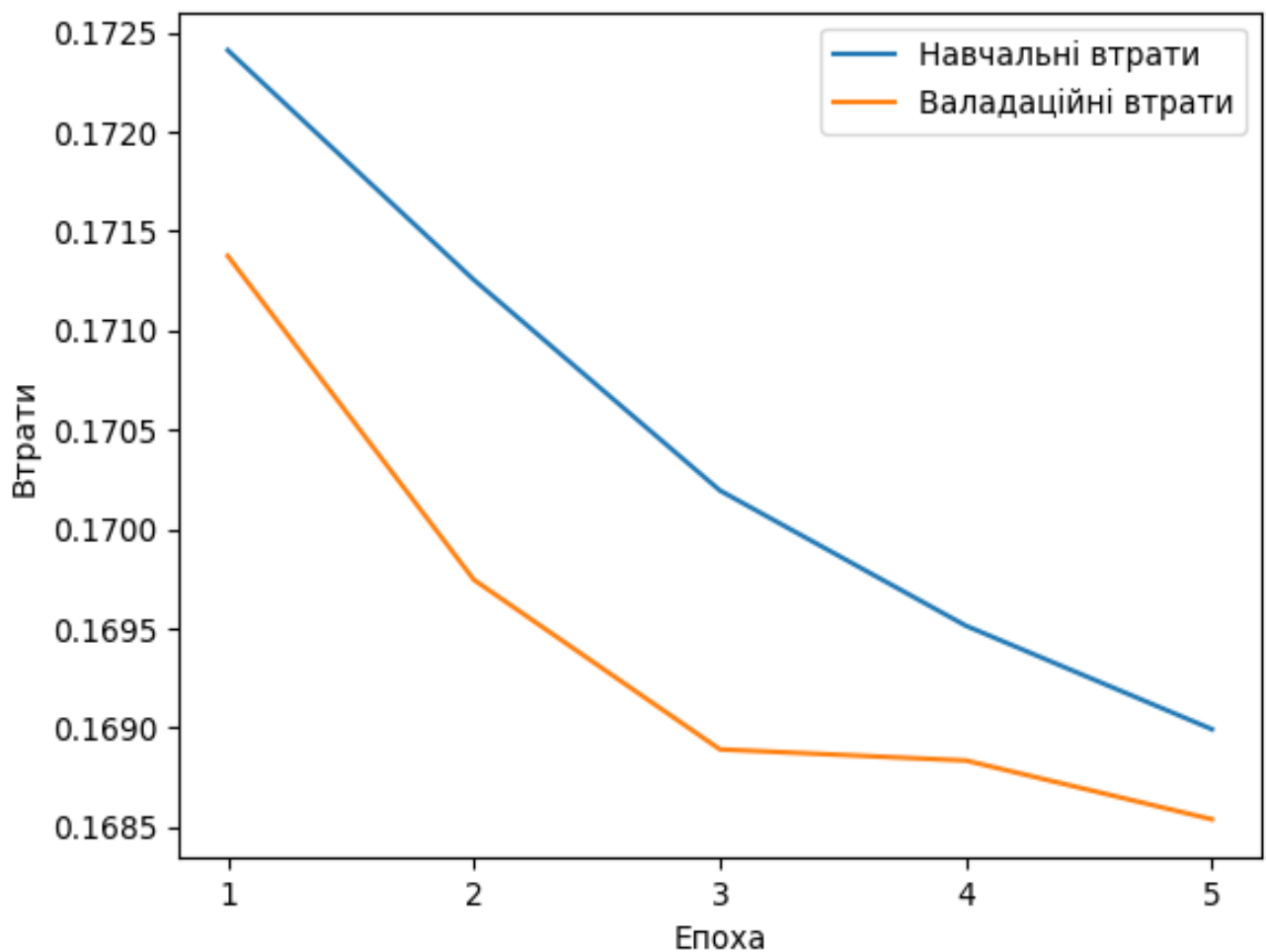


Рис. 3.6. Графік зміни втрат при навчанні для навчальної та варіаційної вибірки

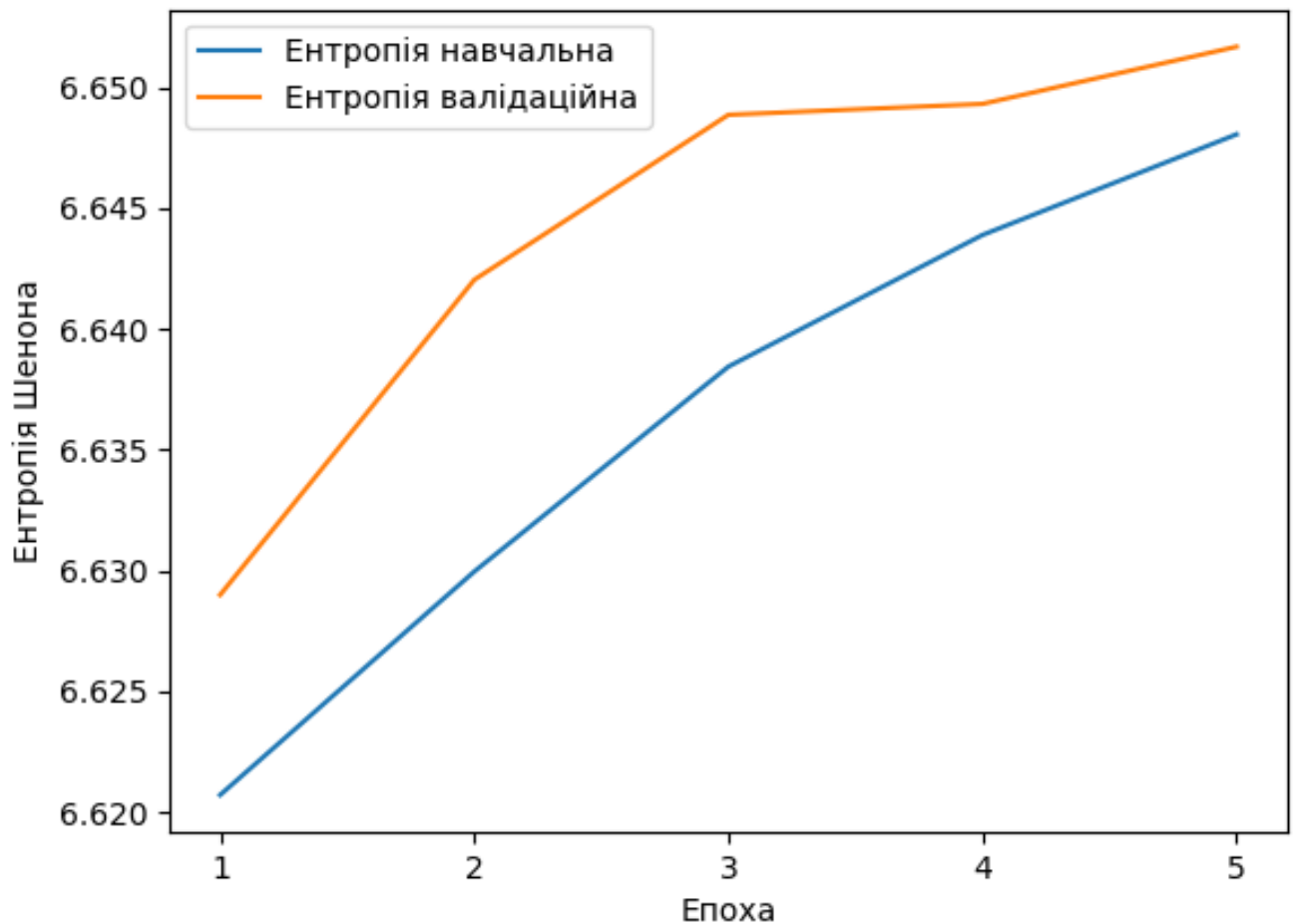


Рис. 3.7. Графік зміни ентропії Шенона при навчанні для навчальної та валідаційної вибірки

3.4. Оцінка методу комплексування зображень з адаптивним визначенням ваг

Для порівняння результатів комплексування було реалізовано дев'ять інших методів з використанням мови програмування Python. Зокрема, це метод максимумів, метод підсилення відмінностей, черезрядковий метод, метод усереднення, пороговий метод, степеневий метод, метод вагової функції, метод на основі піраміди Лапласа та метод на основі вейвлет-перетворення. Вибір цих методів був обумовлений необхідністю покрити більшість існуючих на даний час підходів для порівняння.

Приклад результатів комплексування всіма цими порівнювальними методами зображено на рис. 3.8.

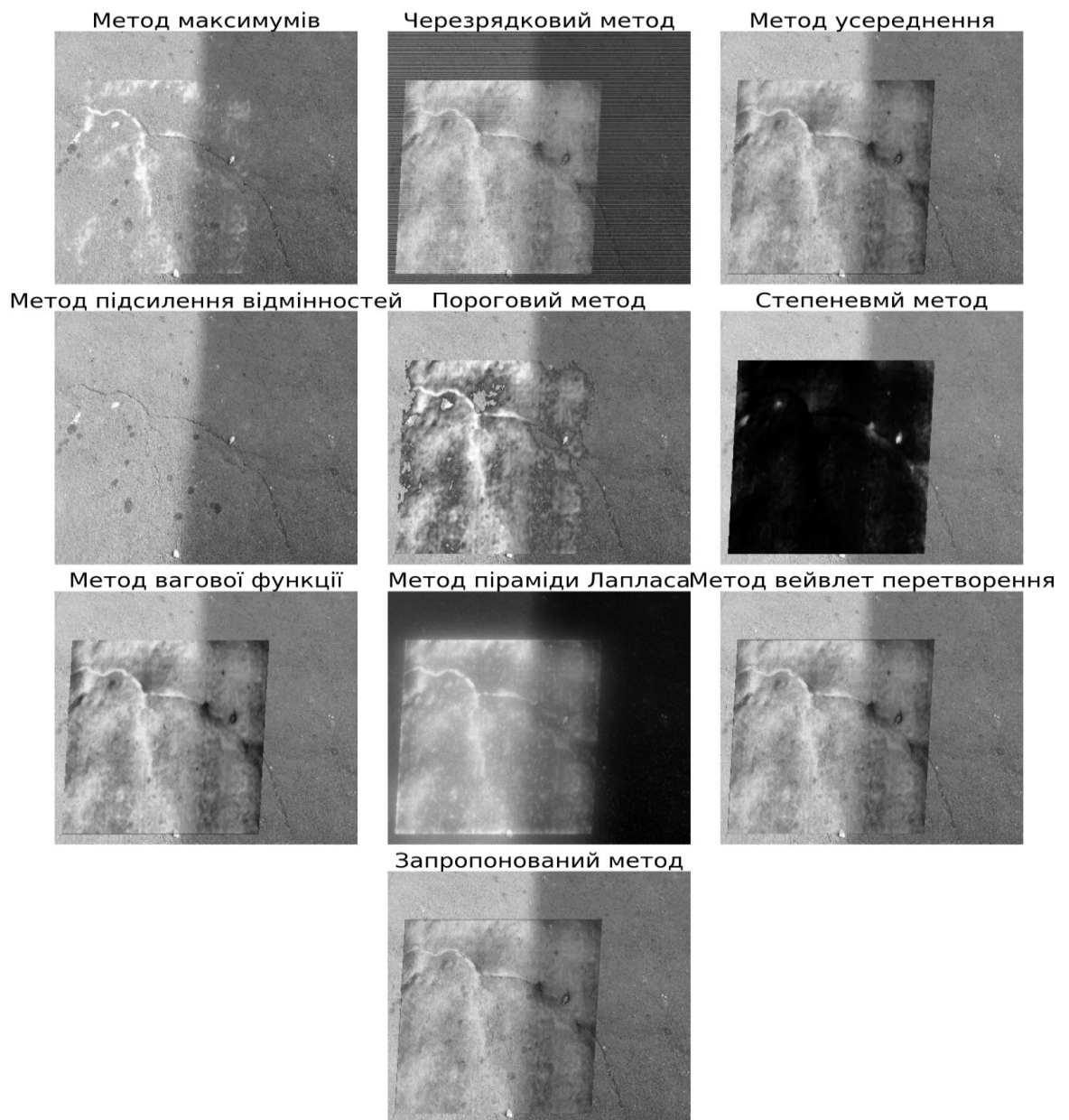


Рис. 3.8. Результати комплексування зображень всіма методами що порівнюються

Як критерій для оцінки інформативності, було обрано ентропію Шенона, що слугує показником інформативності зображень. Для оцінювання застосовувалася окрема вибірка тестових зображень у кількості 300 штук, які не використовувалися при навчанні нейронних мереж запропонованого методу, щоб уникнути будь-яких упереджень у результатах через запам'ятовування мережею даних.

Для аналізу отриманих вибірок ентропії на наявність викидів та попередніх висновків ідеально підходить коробкова діаграма (див. рис. 3.9). Як видно з діаграми, присутні можливі викиди, позначені кружечками на діаграмі. Тому для подальшого аналізу пропонується використовувати медіану та статистичну

оцінку за методом Kruskal-Wallis.

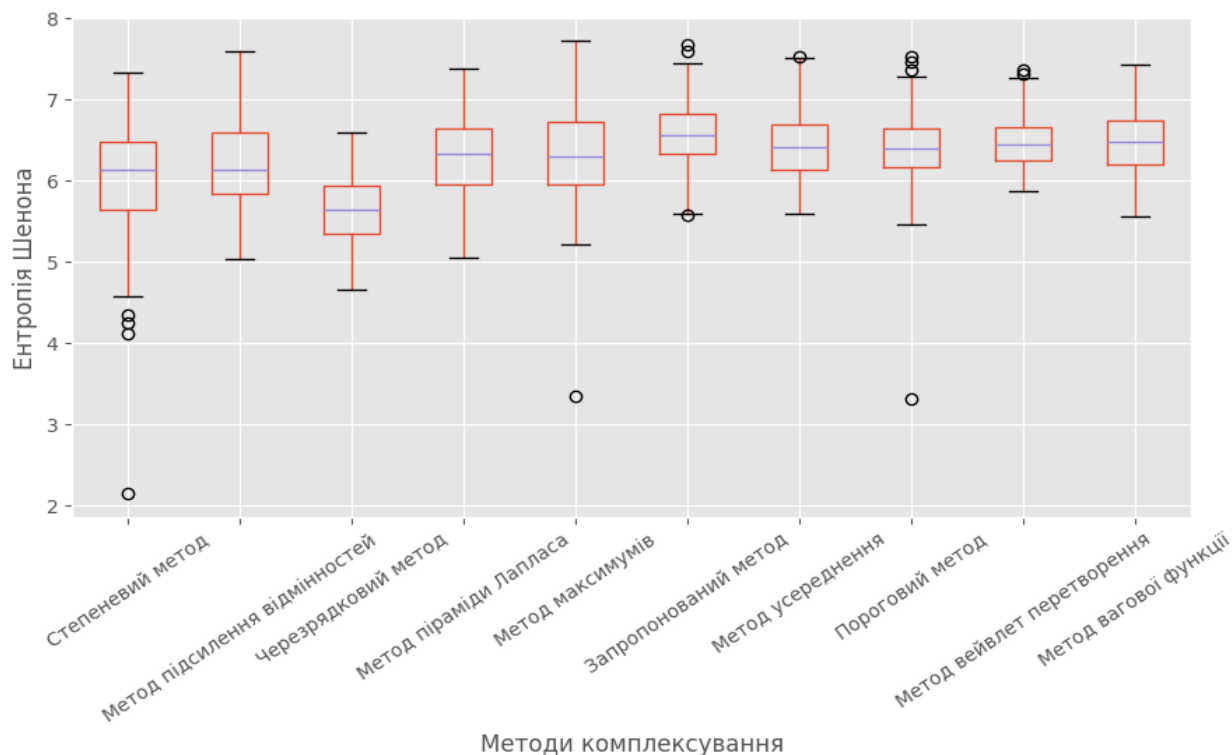


Рис. 3.9. Коробкова діаграма отриманих вибірок ентропії Шенона при комплексуванні різними методами

Таким чином, використання різних методів комплексування та застосування коробкової діаграми для аналізу результатів дозволяє отримати більш об'єктивну картину та забезпечує достовірність висновків на основі отриманих даних.

Для початку необхідно впевнитись в статистичній значущості відмінностей результатів різних методів, для цього скористаємося статистичною оцінкою Kruskal-Wallis.

Для дев'яти ступенів свободи ($df = 9$) значення цієї статистики склало $H = 697$; статистична значущість (Sig.) отримана близькою до 0.

Отриманий рівень значущості менший за промисловий стандарт p -значення на рівні 0,05, тому нульова гіпотеза H_0 про відсутність значних відмінностей між вибірками відхилена. Отже, відмінності між вибірками значущі. Тобто, залежність між методом комплексування та значенням ентропії Шенона статистично доведена.

Далі доцільно розглянути медіани отриманих результатів на які менше

впливають потенційні викиди (рис. 3.10).

Як видно з діаграми найгірший результат показав черезрядковий метод з ентропією на рівні 5.63 бітів. Метод підсилення відмінностей та степеневий показали схожі результати на рівні 6.1 бітів. Методи максимумів, усереднення, пороговий та піраміди Лапласа показали значення в околицях 6.35 бітів, що є досить не поганим результатом. Метод на основі вейвлет перетворення отримав медіану оцінки 6.43, що є 3 найкращим результатом. Метод вагової функції зайняв 2 місце з результатом 6.47 бітів ентропії Шенона. Найбільшу ентропію було отримано при використанні запропонованого методу комплексуювання зображень з адаптивним підбором вагових коефіцієнтів, а саме, медіану на рівні 6.56. Цей показник є на 1.4% більший я другий по величині, що є значною різницею для ентропії.

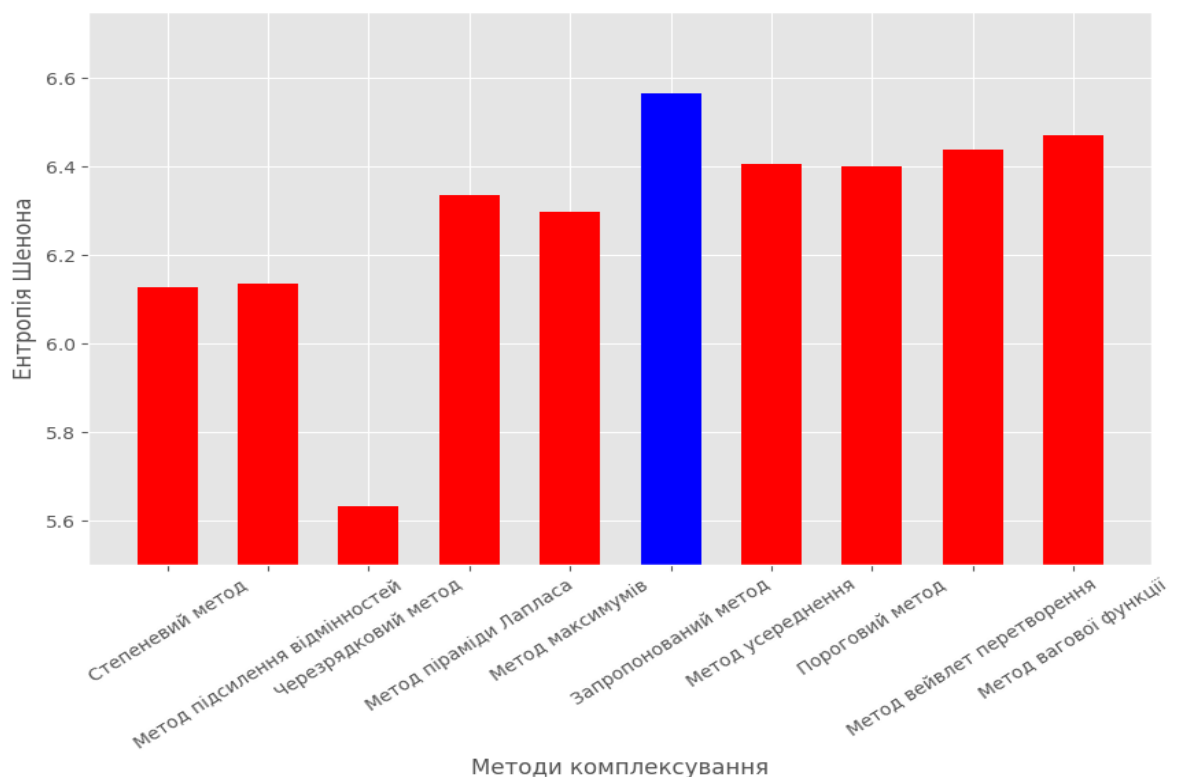


Рис. 3.10. Медіани отриманих значень ентропії Шенона для кожного проаналізованого методу комплексуювання. Синім позначено запропонований метод комплексуювання

Таким чином запропонований метод комплексуювання показав збільшення інформативності оціненої показником ентропії Шенона в порівнянні з аналогами, що говорить про ефективність розробленого методу комплексуювання.

ВИСНОВКИ ДО РОЗДІЛУ 3

1. Запропоновано та розроблено метод аугментації зображень на базі комплексування зображень з вейвлет перетворенням. Метод має допомогти вирішити проблеми пов'язані з браком об'єму та різноманітності даних в навчальній вибірці. Запропонований метод також має позитивний побічний ефект, а саме, збільшення інформативності результуючого зображення, що позитивно впливає на процес навчання нейронних мереж.

2. Проведено порівняння запропонованого методу аугментації з представленими раніше методами. Статистичний аналіз результатів підтвердив ефективність запропонованого методу при задачах класифікації. Новий метод аугментації підвищує точність класифікації зображень, а також, має меншу алгоритмічну складність в порівнянні з іншими методами змішування зображень на базі вейвлет перетворення.

3. Запропоновано метод комплексування зображень з адаптивним визначенням ваг, який повинен зменшити вплив не інформативних вхідних зображень на інформативність результуючого, тим самим підвищивши інформативність результуючого зображення.

4. Розроблено та реалізовано нейронну мережу адаптивного визначення ваг для запропонованого методу комплексування зображень. Яка дала можливість адаптивно визначати вагові коефіцієнти для подальшого використання їх в запропонованому методі комплексування. Обґрунтовано її структуру та проаналізовано ефективність на валідаційній вибірці.

5. Проведено порівняння запропонованого методу комплексування з представленими раніше методами. Статистичний аналіз результатів підтвердив ефективність запропонованого методу створювати результуюче зображення більшої інформативності, яка вимірювалась ентропією Шенона.

Список використаних джерел до розділу 3

- [1] Wronski, B., Garcia-Dorado, I., Ernst, M., Kelly, D., Krainin, M., Liang, C. K., ... & Milanfar, P. (2019). Handheld multi-frame super-resolution. *ACM Transactions on Graphics (ToG)*, 38(4), 1-18.
- [2] Storozhik, D. V., Protasov, A. G., Muraviov, O. V., Petrik, V. F., & Petrenko, D. V. *Technical Diagnostics and Non-Destructive Testing*, 2022, № 02.
- [3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [4] Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260), 583-621.
- [5] Storozhyk, D. V., Muraviov, O. V., Protasov, A. G., Bazhenov, V. G., & Bohdan, G. A. (2020). Комплексування мультиспектральних зображень як метод підвищення їх інформативності при бінарній сегментації. *KPI Science News*, (2), 82-87.

РОЗДІЛ 4. Розробка автоматизованої системи діагностики та ремонту дорожнього покриття

4.1. Формування навчальної вибірки зображень дефектів у видимому та інфрачервоному спектрах

Для реалізації нейронних мереж і оцінки ефективності елементів системи, було зібрано навчальний набір даних пар зображень дефектів дорожнього покриття у видимому та інфрачервоному спектрах.

Так для отримання видимих зображень використовувалось дві камери смартфонів Samsung Galaxy A71 і Samsung Galaxy S10+, дві камери використовувались, щоб показати адаптивність системи до різних камер. Зображення отримувалось в форматі RGB з роздільною здатністю 2312×1734 пікселів.

Для отримання теплових зображень використовувалась тепловізійна камера Wintact WT3320. Яка мала роздільну здатність детектора 320×240 пікселів та теплову чутливість 0.07°K , які дозволяли отримувати зображення при температурі повітря вищій за 15°C .

Зображення збирались для таких категорій дефектів (рис. 4.1): вибоїни, сітка тріщин, тріщини загальні поперечні, тріщини повздовжні, косі тріщини, без дефектна поверхня. Такий вибір обумовлено найчастішою появою саме таких дефектів дорожнього покриття.

Як видно з прикладів, при зборі даних акцент робився на дефекти в околицях яких присутні інформаційні шуми у вигляді тіней, забруднення, сторонніх предметів, в поганих умовах освітлення. Це все необхідно для ілюстрації надійності системи в умовах присутності таких завад.


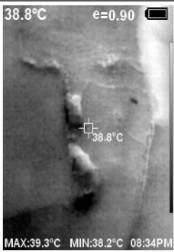

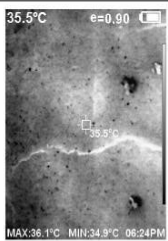

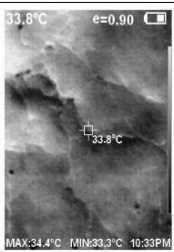

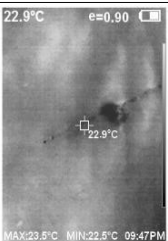

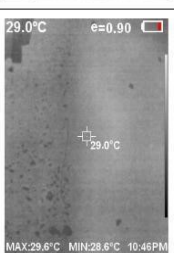

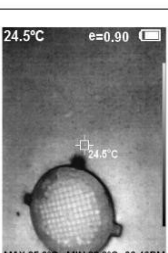
	Оптичне зображення	Теплове зображення		Оптичне зображення	Теплове зображення
Тріщини загальні поперечні		 38.8°C e=0.90 39.8°C MAX:39.3°C MIN:38.2°C 08:34PM	Тріщини повздовжні		 35.5°C e=0.90 35.5°C MAX:36.1°C MIN:34.9°C 06:24PM
Вибоїни		 33.8°C e=0.90 33.8°C MAX:34.4°C MIN:33.3°C 10:33PM	Косі тріщини		 22.9°C e=0.90 22.9°C MAX:23.5°C MIN:22.5°C 09:47PM
Сітка тріщин		 29.0°C e=0.90 29.0°C MAX:29.6°C MIN:28.6°C 10:46PM	Бездефектна поверхня		 24.5°C e=0.90 24.5°C MAX:25.0°C MIN:23.8°C 06:46PM

Рис. 4.1. Приклад зображень різних категорій дефектів зібраної навчальної вибірки.

Для кожного типу дефекту було зібрано 500 пар зображень видимого та інфрачервоного спектрів. Таким чином загальний об'єм даних становив 6000 зображень, а загальний розмір понад 12 гігабайтів.

4.2. Розробка структурної схеми автоматизованої системи діагностики та ремонту дорожнього покриття

На основі пророблених досліджень пропонується така структура автоматизованої системи діагностики та ремонту дорожнього покриття з двома мультимодальними датчиками, а саме, цифровою камерою та інфрачервоною камерою, і комплексуванням зображень (рис. 4.2).

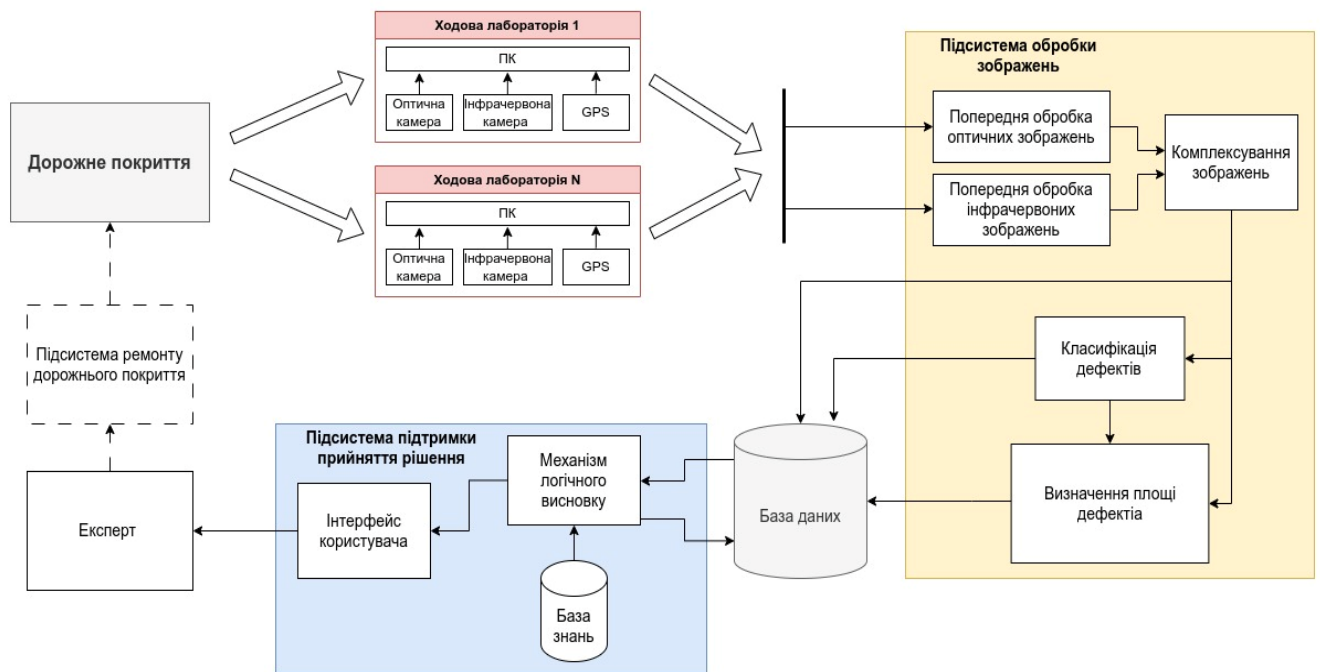


Рис. 4.2. Структурна схема автоматизованої системи діагностики та ремонту дорожнього покриття

Для збирання даних про наявний стан дорожнього покриття було обрано камери видимого та інфрачервоного випромінювання. Обрано саме такі типи, які мають властивості компенсувати вади одне одного та близькі до структури вихідного сигналу. Також, такі датчики навіть в парі в рази дешевші за лазерні сканери чи георадари.

Обрані типи датчиків обладнані на ходових лабораторіях, які є розподіленою підсистемою збору даних в системі. Ходові лабораторії представляють собою автомобіль оснащений камерами, GPS навігатором і бортовим комп'ютером для локального зберігання зібраних даних і асинхронної передачі їх на централізоване хмарне сховище не оброблених даних. Кількість ходових лабораторій обмежена лише бюджетом і потребою. Перевагою використання саме такої розподіленої системи є швидкість збору даних, відсутність необхідності перешкоджанню вільного руху транспорту, укомплектування лише необхідним обладнанням дає можливість зменшити вартість одиниці такої лабораторії.

Використання централізованого хмарного сховища дає можливість мати низько вартісне, швидко масштабоване рішення для проміжного зберігання

отриманих даних та забезпечує єдину вхідну точку для обробки зібраних даних.

Підсистема обробки зображень опрацьовує отримані пари зображень видимого та інфрачервоного спектрів пакетами для збільшення ефективності обробки. Головною ціллю цієї підсистеми є перехід від зображень до описових характеристик, таких як тип дефекту та цього площа. Ця мета досягається коштом двох окремих процесів попередньої обробки для кожного спектру, блоком комплексування зображень, нейромережею для класифікації (ідентифікація наявності дефекту і типу дефекту) та нейромережею для сегментації зображень (ідентифікація прощі дефекту). Окремі процеси попередньої обробки необхідні через різну природу сигналів, до прикладу, інфрачервоне зображення зазвичай є більш зашумленим. Комплексування зображень прибирає необхідність подальшої обробки зображень окремо, тим самим зменшуючи складність подальшого аналізу, а також, процес комплексування збільшує загальну інформативність наявних даних, що позитивно впливає на подальшу точність нейромереж.

Отримані оброблені зображення та описові характеристики зберігаються в базі даних для подальшого процесу оцінки стану дорожнього покриття та поповнення наявних знань для подальшої оптимізації моделей прийняття рішення.

Підсистема підтримки прийняття рішень на основі зібраних даних, історичних даних щодо діагностики і втручань (ремонтів), та нормативних документів (паспортів доріг, тощо) пропонує категорію стану дорожнього покриття на ділянці, термін в який необхідно зробити втручання, тип втручання та необхідні інструменти для вирішення несправностей. Усе це досягається за допомогою бази знань що містить в собі продукційну модель прийняття рішення, фрагмент моделі для визначення категорії наявного стану ділянки дорожнього покриття наведено на табл., та механізму логічного висновку. Отримані рекомендації та раніше зібрані дані в зручному для користувача виводяться на користувацький інтерфейс.

Таблиця. 4.1. Фрагмент продукційної моделі прийняття рішень

<p>ЯКЩО тип дефекту = "тріщина загальна поперечна", ТО характер дефекту = лінійний і коефіцієнт вагомості $Kv=0.08$ і коефіцієнт приведення до площі $Ks=0.15$</p> <p>ЯКЩО тип дефекту = "тріщина повздовжня", ТО характер дефекту = лінійний і коефіцієнт вагомості $Kv=0.1$ і коефіцієнт приведення до площі $Ks=0.1$</p> <p>ЯКЩО тип дефекту = "коса тріщина", ТО характер дефекту = лінійний і коефіцієнт вагомості $Kv=0.10$ і коефіцієнт приведення до площі $Ks=0.2$</p> <p>ЯКЩО тип дефекту = "сітка тріщин", ТО характер дефекту = площадний і коефіцієнт вагомості $Kv=0.10$</p> <p>ЯКЩО тип дефекту = "вибоїна", ТО характер дефекту = площадний і коефіцієнт вагомості $Kv=0.08$</p> <p>ЯКЩО дефектність дорожнього покриття $> 50\%$, ТО рівень дефектності = критичний</p> <p>ЯКЩО категорія дороги = I-II і дефектність дорожнього покриття = 5-10%, ТО рівень дефектності = Рівень 1 (ДП 1)</p> <p>ЯКЩО категорія дороги = I-II і дефектність дорожнього покриття = 10-20%, ТО рівень дефектності = Рівень 2 (ДП 2)</p>	<p>ЯКЩО категорія дороги = I-II і дефектність дорожнього покриття $> 20\%$, ТО рівень дефектності = Рівень 3 (ДП 3)</p> <p>ЯКЩО категорія дороги = III і дефектність дорожнього покриття = 10-15%, ТО рівень дефектності = Рівень 1 (ДП 1)</p> <p>ЯКЩО категорія дороги = III і дефектність дорожнього покриття = 15-25%, ТО рівень дефектності = Рівень 2 (ДП 2)</p> <p>ЯКЩО категорія дороги = III і дефектність дорожнього покриття $> 25\%$, ТО рівень дефектності = Рівень 3 (ДП 3)</p> <p>ЯКЩО категорія дороги = IV-V і дефектність дорожнього покриття = 15-20%, ТО рівень дефектності = Рівень 1 (ДП 1)</p> <p>ЯКЩО категорія дороги = IV-V і дефектність дорожнього покриття = 20-30%, ТО рівень дефектності = Рівень 2 (ДП 2)</p> <p>ЯКЩО категорія дороги = IV-V і дефектність дорожнього покриття $> 30\%$, ТО рівень дефектності = Рівень 3 (ДП 3)</p>
---	--

Експерт на основі отриманих даних і рекомендацій приймає остаточне рішення щодо втручання на ділянці. Така архітектура системи дозволяє збільшити продуктивність такого експерта.

Підсистема ремонту дорожнього покриття містить всі засоби та інструменти для проведення ремонтних робіт. Під всім цим розуміється все необхідне обладнання, спеціалісти, роботизовані системи та необхідні витратні матеріали.

4.3. Розробка підсистеми опрацювання зображень

4.3.1. Розробка процесу попередньої обробки зображень з комплексуванням зображень

Як вже було сказано раніше доцільно попередньо зробити обробку кожного каналу окремо через їх різну природу.

Так, для зображення видимого спектру через велику роздільну здатність, якість і показники оцінки сигнал-шум достатньо зменшити роздільну здатність

для зменшення необхідних обчислювальних потужностей для подальшої обробки.

Для інфрачервоного зображення ситуація складніше. Зображення низької роздільної здатності, часто сильно зашумлене і мало контрастне, через що важко розібрати детальні елементи зображення. Також часто як в нашому випадку, зображення містить маркування та інші інформативні елементи які не можливо прибрати налаштуваннями вбудованого ПЗ тепловізора, однак ці інформативні маркування несуть лише інформаційне забруднення для наших цілей.

Запропонований процес обробки зображено на рис. 4.3.

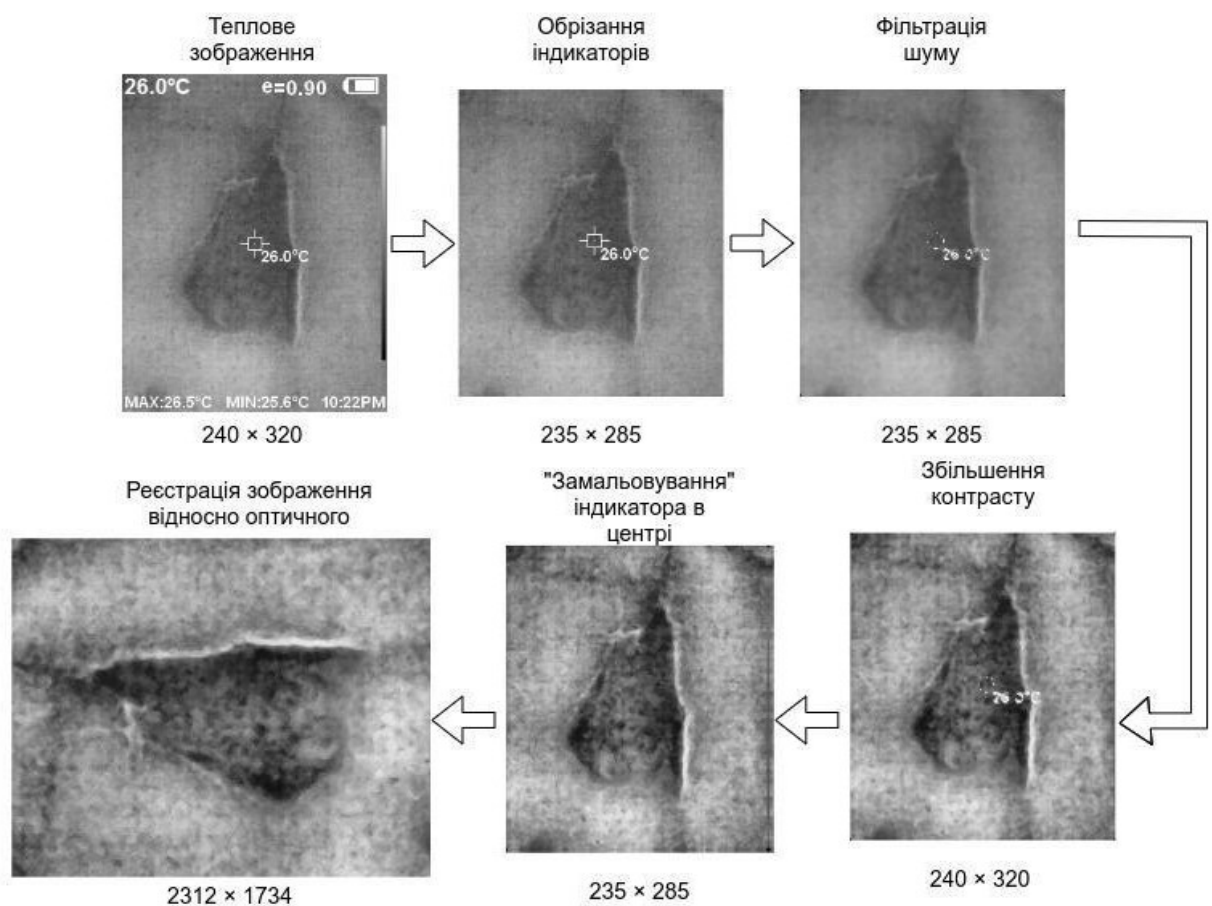


Рис. 4.3 Процес попередньої обробки теплових зображень

Спершу необхідно прибрати не інформативні індикатори для цього найлегшим методом є обрання зображення до розмірів 235 x 285 пікселів.

Зменшення шуму в інфрачервоних зображеннях є одним із критичних етапів попередньої обробки, оскільки воно покращує якість візуалізації та точність подальшого аналізу, тому його обрано другим етапом попередньої

обробки. Інфрачервоні зображення часто зашумлені через вплив зовнішнього середовища, теплові коливання та параметри сенсорів. Цей шум може спотворювати інформацію, ускладнюючи виявлення об'єктів, сегментацію та інші обчислювальні процеси. Тому ефективні методи фільтрації є важливими для збереження корисної інформації та підвищення точності аналізу.

Адаптивний медіанний фільтр є одним із найефективніших методів зменшення шуму в інфрачервоних зображеннях, оскільки він може пристосовуватися до локальних особливостей зображення, видаляючи шум, включаючи імпульсні завади, зберігаючи при цьому чіткість країв і деталі об'єктів. Це особливо важливо для інфрачервоних зображень, де збереження дрібних структур є критично важливим для подальшого аналізу та інтерпретації даних. Тому використання адаптивного медіанного фільтра підвищує якість зображень і, як наслідок, точність автоматизованих систем обробки та аналізу інфрачервоних даних [1].

Математична формалізація алгоритму адаптивного медіанного фільтра.

Нехай $W_s(x, y)$ — це вікно розміром $S \times S$ з центром у пікселі $I(x, y)$. Алгоритм працює за наступними етапами:

Визначення шуму на основі медіанного значення:

$$A_1 = I_{med} - I_{min}; A_2 = I_{med} - I_{max}, \quad (4.1)$$

де I_{med} — медіана вікна, I_{min} — мінімальне значення вікна, I_{max} — максимальне значення вікна.

Якщо $A_1 > 0$ і $A_2 < 0$, перейти до фази В. Інакше збільшити розмір вікна S і повторити обчислення.

Перевірка пікселя на шумовий вплив:

$$B_1 = I(x, y) - I_{min}; B_2 = I(x, y) - I_{max} \quad (4.2)$$

Якщо $B_1 > 0$ і $B_2 < 0$, залишити піксель $I(x, y)$ незмінним. Якщо ні, замінити $I(x, y)$ на I_{med} .

Інфрачервоні зображення зазвичай мають обмежений контраст через особливості випромінювання ІЧ зображення, що ускладнює виявлення та ідентифікацію об'єктів на зображенні. Одне з можливих рішень для цього —

підвищення контрасту, що допомагає підкреслити важливі деталі, покращуючи видимість критичних об'єктів та структур для завдань діагностики критичної інфраструктури. Щоб досягти цього, контраст включено в процес попередньої обробки.

Один з ефективних варіантів – адаптивне гістограмне вирівнювання з обмеженням контрасту (англ. Contrast-Limited Adaptive Histogram Equalization, CLAHE). За допомогою методу CLAHE контраст зображення зростає пропорційно приймачу, зберігаючи локальні деталі та запобігаючи пересиченню світлих ділянок. Таким чином, підвищення контрасту відбувається більш рівномірно в різних областях зображення, що важливо через те, що кадри ІЧ зображення мають нерівномірність освітленості та великий шум. Окрім того, використання методу CLAHE забезпечує більш чіткі та інформативні кадри ІЧ, що підвищує ефективність аналізу та обробки даних [2].

Наступним етапом для остаточного видалення не інформативних індикаторів, а саме, індикатора посередині, пропонується використовувати генеративну нейронну мережу для прибирання елементів на знімках LAMA. Ця генеративна штучна нейронна мережа з великою точністю на основі околиць елементу, який необхідно видалити, прогнозує інтенсивність RGB пікселя на заміну. Таким чином без сильних спотворень інформативних елементів вдалося позбавитися не інформативного, що впливав би на подальшу точність діагностики.

Останнім етапом обробки теплового зображення є реєстрація (вирівнювання) зображення відносно видимого.

Просторово-трансформаційні мережі (Spatial Transformer Networks, STN) є потужним інструментом для попередньої обробки зображень. Особливо це стосується випадків, коли потрібно внести поправки на геометричні спотворення або вирівняти зображення для подальшої легалізації. Використання STN надає можливість моделям самостійно навчитися виконувати набір просторових перетворень, таких як масштабування, обертання, зсув та перспективне перетворення, що покращує точність та стабільність наступних етапів обробки.

Для мультимодального злиття двох зображень STN особливо важливий, оскільки дозволяє забезпечити точне вирівнювання зображень з різних модальностей. Автоматична корекція спотворень необхідна, оскільки навіть відносно невелика невідповідність може призвести до втрати інформації та зниження якості кінцевого результату, що є критичним для досягнення високоякісного злиття.

Використання STN має велике перевагу, оскільки не потребує ручної розмітки контрольних точок чи попередніх знань про геометрію зображень. Крім того, STN автоматично адаптується до умов вхідних даних, що забезпечує мобільність та ефективний процес обробки зображень у різних умовах.

Таким чином, STN для попередньої обробки зображень у мультимодальному злитті допомагає вдосконалити точність та якість кінцевого результату, що робить його важливим інструментом у сучасній комп'ютерній візії та обробці зображень.

Загальний процес попередньої обробки зображень двох модальностей з комплексуванням зображень зображено на рис. 4.4.

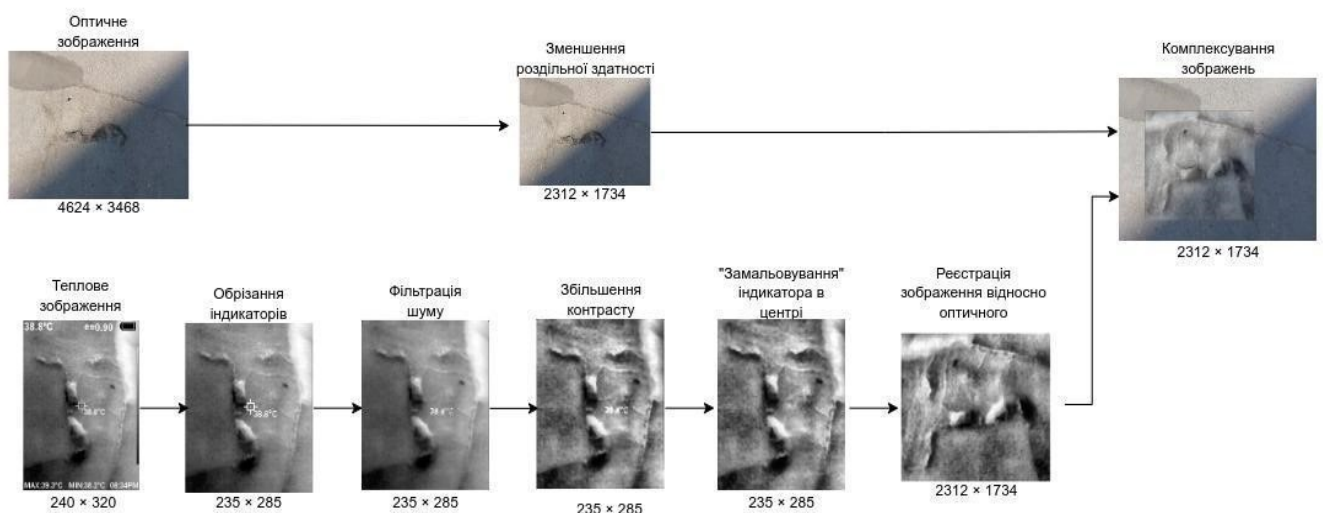


Рис. 4.4. Процес попередньої обробки зображень двох модальностей з комплексуванням зображень

Завершальним етапом попередньої обробки зображень є комплексування зображень. Воно необхідне для збільшення інформативності результуючого зображення через поєднання інформативних елементів зображень двох спектрів,

для забезпечення зменшення негативного впливу інформаційних шумів (не інформативних елементів на зображенні) та зменшення необхідних обчислювальних потужностей на наступних етапах діагностики.

Як метод комплексування обрано розроблений метод комплексування зображень з адаптивним визначенням ваг, який показав свою ефективність в задачах підвищення інформативності в умовах поганої інформативності початкових зображень.

Таким чином, маємо готовий процес попередньої обробки зображень двох спектрів що має підвищити точність виявлення інформаційних ознак зображень на наступних етапах діагностики.

4.3.2. Підбір та навчання нейронної мережі для класифікації типів дефектів

При виборі нейронної мережі для конкретної задачі необхідно керуватися правилом найменшого достатнього. Тому для підбору найкращої архітектури для розв'язання задачі класифікації дефектів дорожнього покриття, пропонується навчати поступово найрезультативніші моделі для класифікації сортуючи їх за кількістю параметрів, при однакових умовах.

Так, для проведення експерименту було обрано 4 мережі: ResNet (60.2 мільйонів параметрів), EfficientNet (118.5 мільйонів параметрів), VGG (138.4 мільйонів параметрів).

Мережі навчалися за однакових умов, а саме, швидкість навчання (англ. learning rate) склала 0.001, розмір партії (англ. batch size) - 10, кількість епох - 15, навчальна та валідаційна вибірки комплексований зображень видимого та теплового спектра — однакові для кожної мережі.

Результати навчання ResNet показані на рис. 4.5

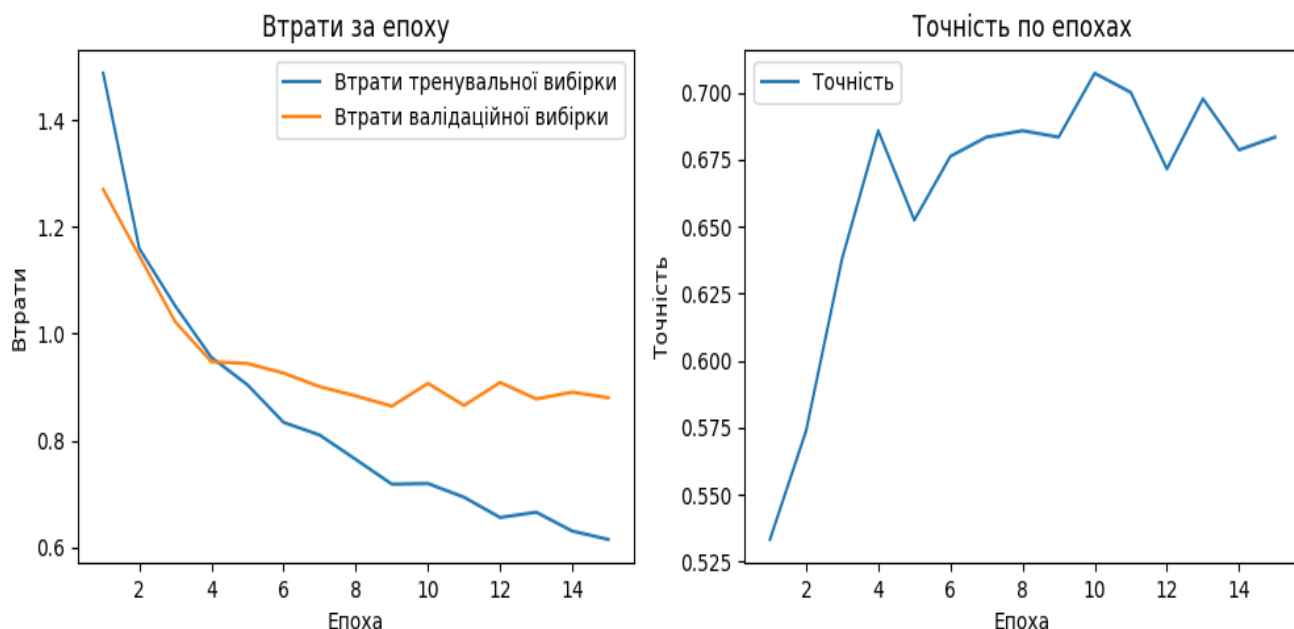


Рис. 4.5. Результати навчання ResNet: графік втрат — зліва, графік точності — справа.

Як видно з результатів після 3 епохи нейронна мережа починає перенавчатися (втрати тренувальної вибірки зменшуються в рази швидше за валідаційну вибірку) це говорить про менший потенціал приросту точності при продовженні навчання. По графіку змін точностей визначення класу варіаційної вибірки видно явну нестабільність результатів від епохи до епохи, що також вказує на перенавчання. Найкраща отримана точність це 79.15% на 13 епосі навчання. Виходячи з цього математична модель ResNet погано підходить для розв'язання задачі класифікації дефектів дорожнього покриття.

Результати навчання (рис. 4.6) нейромережі EfficientNet з такими ж умовами, конфігурацією гіперпараметрів та на тих самих даних вказують на більш позитивний прогноз щодо використання даної архітектури для класифікації дефектів дорожнього покриття в контурі системи автоматизованої діагностики та ремонту дорожнього покриття.

За 15 епох навчання втрати валідаційної вибірки залишалися меншими за втрати тренувальної вибірки, при цьому зберігаючи плавне зменшення показника.

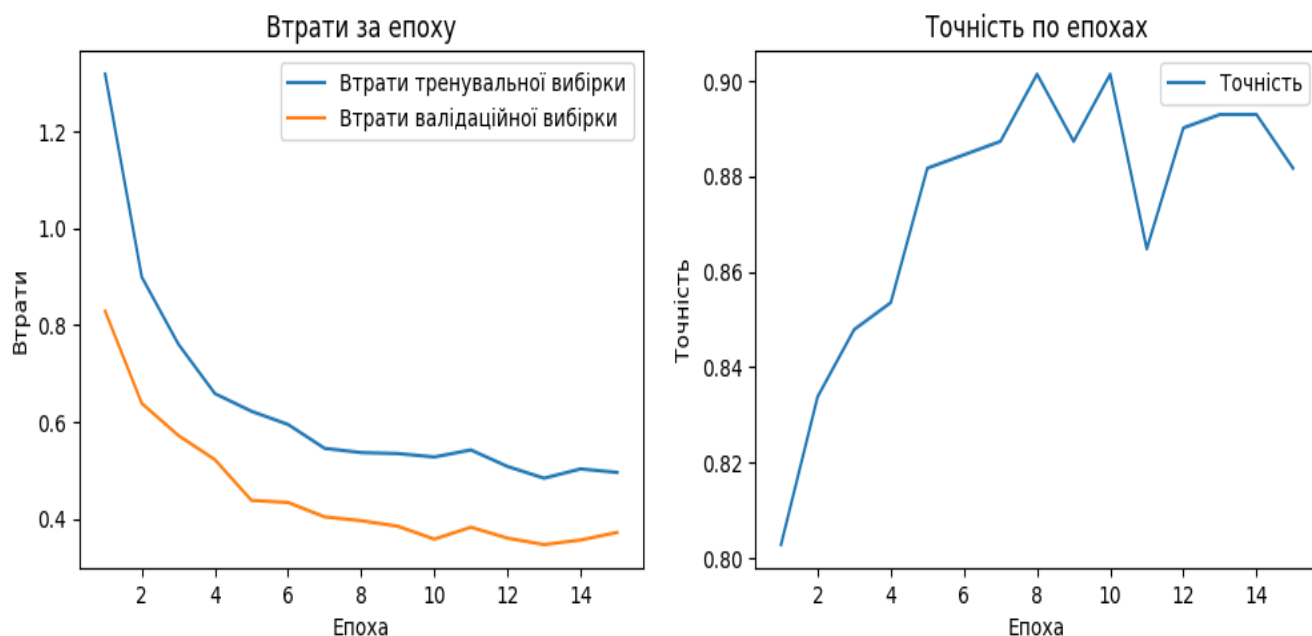


Рис. 4.6. Результати навчання EfficientNet: графік втрат — зліва, графік точності — справа.

Найкращий показник точності валідаційної вибірки отримано на рівні 90% на 8 епосі навчання.

Оскільки збільшення параметрів показало себе ефективно також було навчено більшу модель VGG, результати навчання зображено на рис. 4.7.

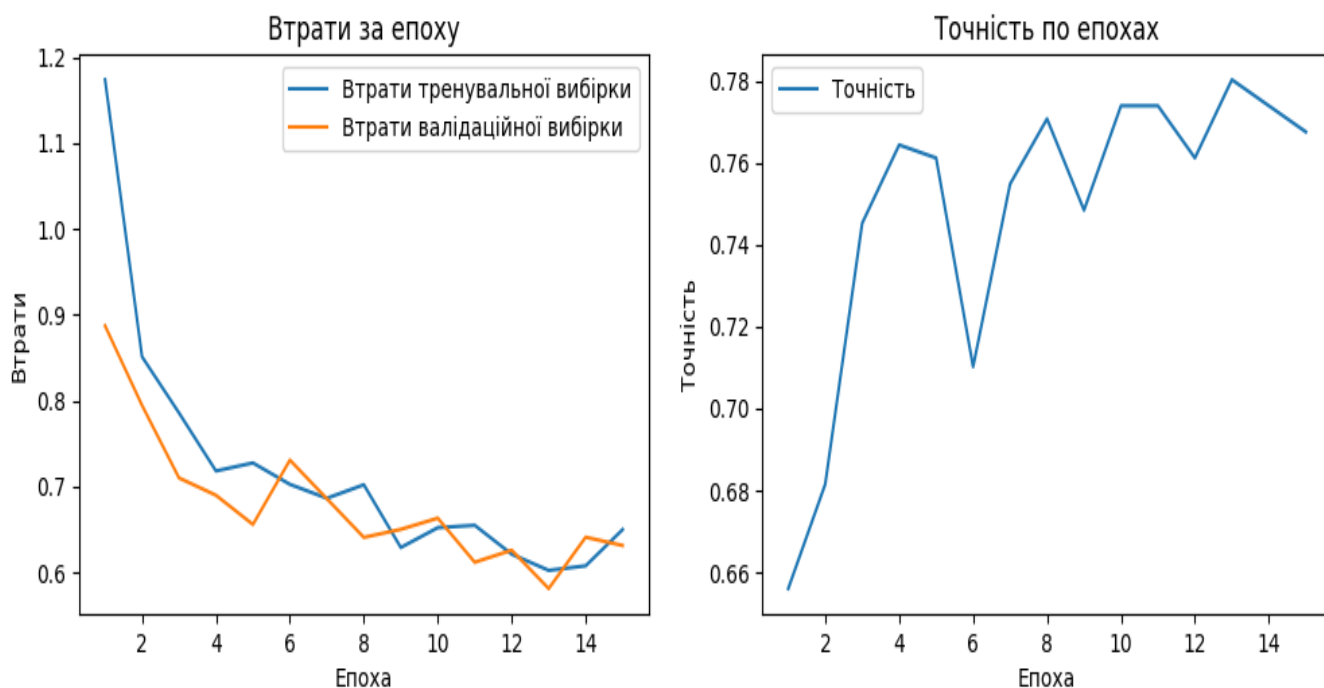


Рис. 4.7. Результати навчання VGG: графік втрат — зліва, графік точності — справа.

При цій архітектурі навчальної вибірки стали меншими за валідаційну на

6 епосі, найкращий показник точності отриманий на 13 епосі і склав 78%. Спираючись на графіки змін точностей та втрат можна говорити по меншу стабільність навчання класифікації дефектів в порівнянні з EfficientNet.

Оскільки збільшення розмірів моделі не дало більшого результату пропонується зупинитися на архітектурі EfficientNet, яка показує більшу стабільність, найкращу точність та найбільшу протидію перенавчанню.

Підсилити результати точності моделі EfficientNet можливо завдяки більш детальному підбору гіперпараметрів навчання і додаванню методів аугментацій для збільшення об'єму даних, що дасть змогу навчати мережу більшу кількість циклів до перенавчання.

Після багатьох спроб навчання з різними конфігураціями результату з 99% точністю класифікації типу дефекту вдалося досягти завдяки такій конфігурації: кількість епох - 31, швидкість навчання - 0.001, розмір партії - 4 та використанню набору методів аугментації. Методи аугментацій були обрані в більшості однакові для всіх класів, а саме, запропонований метод аугментацій на основі комплексування зображень (0.8), Гаусове розмиття (з вірогідністю 0.4), шум сіл та перець (0.4), випадкове обрізання країв (0.9), випадкова зміна контрастності (0.4), віддзеркалення горизонтальне та вертикальне (0.6), випадкове зменшення роздільної здатності (0.5), випадкове вирізання прямокутника з зображення (0.2), випадкове видалення пікселів (0.2), випадкове коригування HSV (Hue, Saturation, Value) (0.3), Міхур (0.1). Також для всіх окрім, тріщин повздовжних та поперечних, через залежність цих типів дефектів від кута повороту, використовувався випадковий поворот на 90 градусів (0.6) та випадкова афінна трансформація (0.4)

Зміну втрат та точності на навчальній та валідаційній вибірках показано на рис. 4.8. Як видно така конфігурація дала найкращу стабільність навчання та результат по точності.

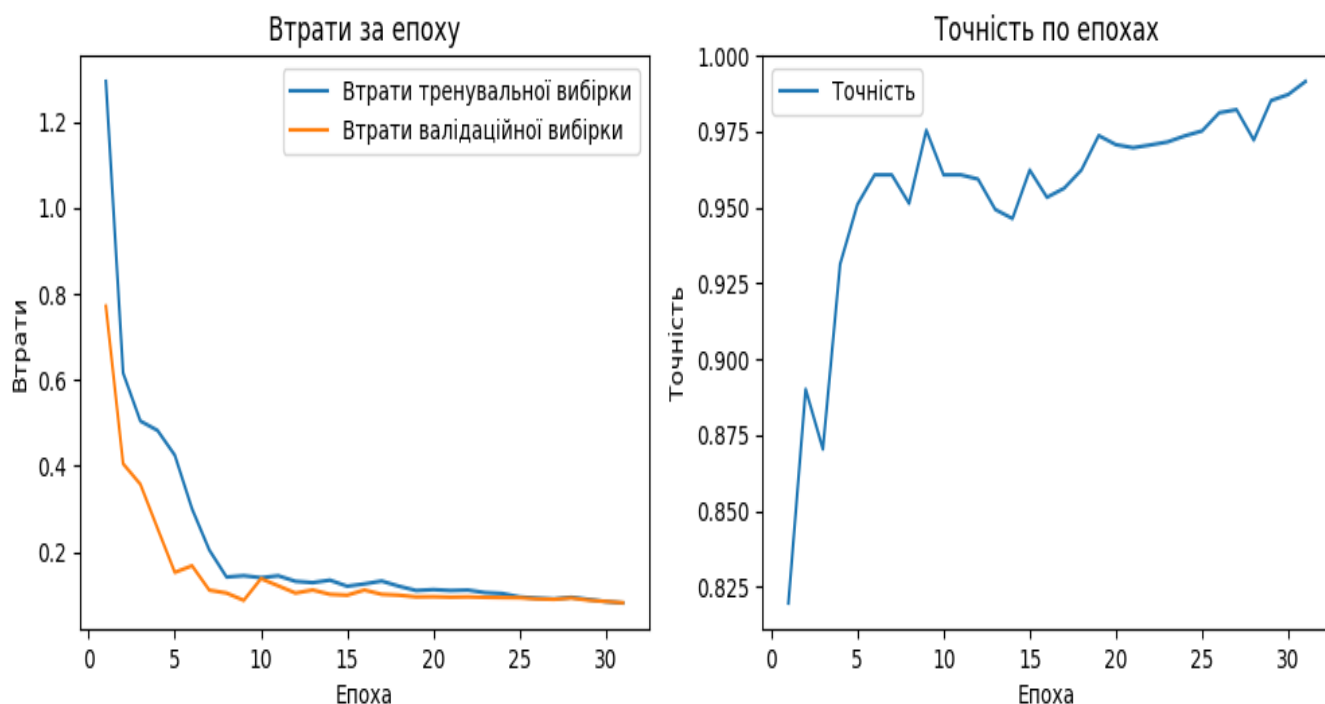


Рис. 4.8. Результати навчання моделі EfficientNet на фінальній конфігурації: графік втрат — зліва, графік точності — справа.

Для оцінки впливу набору методів аугментацій було проведено ще одне машинне навчання без її використання. Так, мережа навчалась 23 епохи без перенавчання і досягла показника точності на рівні 97.6%.

Для кращого розуміння розподілу точностей залежно від типу дефекту було побудовано матрицю заплутаності (рис. 4.9), яка показує який відсоток тестових даних було правильно чи помилково визначила математична модель. Як видно з діаграми лише 3 типи мали не 100% точність прогнозу типу дефекту, а саме, сітка тріщин, тріщини повздовжні та тріщини поперечні. Так, повздовжні та поперечні тріщини хибно вважалися сіткою тріщин, а сітка тріщин помилково визначалась як коса тріщина.

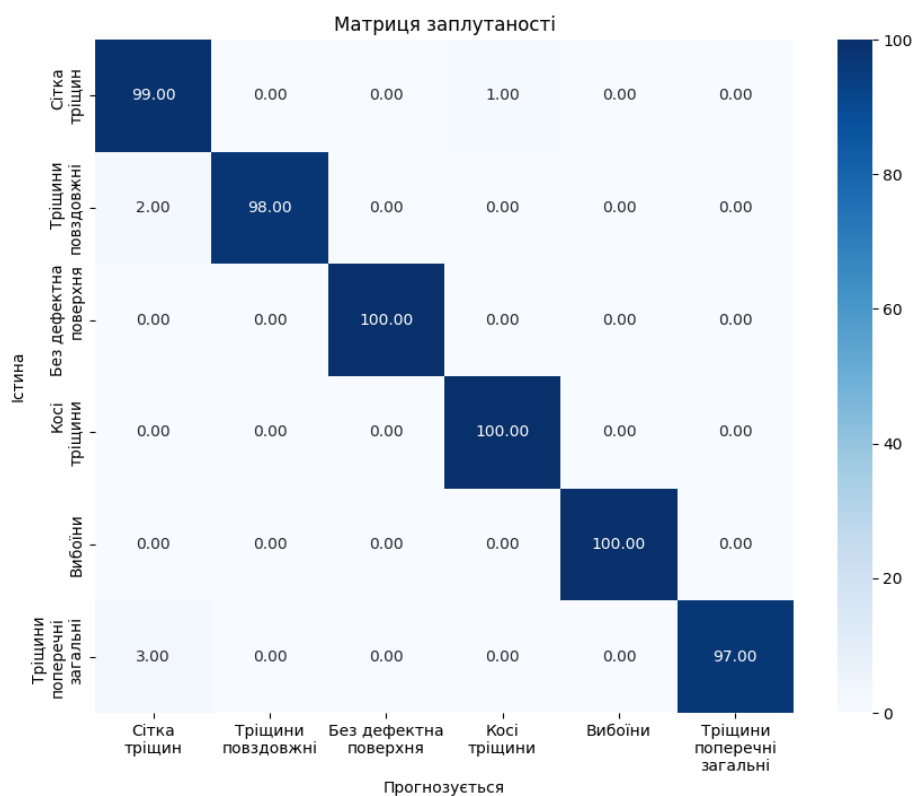


Рис. 4.9. Матриця заплутаності

Порівняння ефективності навченої нейронної мережі в порівнянні з аналогами можна побачити на таб. 4.2.

Таблиця. 4.2. Порівняльна таблиця точностей визначення типу дефектів дорожнього покриття

Назва/Опис	Точність
Запропонована конфігурація моделі з комплексуванням зображень з адаптивним визначенням ваг видимого та інфрачервоного спектрів	99%
Аналог з використанням комплексування зображень видимого та інфрачервоного спектрів [5]	98.34%
Використання лише видимого спектра [5]	96.57%
Використання лише теплового спектра [5]	93.83%
Використання лазерних сканерів [6]	99.6%
Використання 3D лазерних сканерів [7]	98%
Використання Kinect [8]	89.09%

Як видно з порівняльної таблиці точність запропонованої конфігурації моделі з комплексуванням зображень з адаптивним визначенням ваг видимого та

інфрачервоного спектрів перевищує аналоги, окрім, методу на основі лазерного сканування, однак, такі сканери в разі дорожчі за камери теплого та видимого спектрів.

4.3.3. Обґрунтування вибору моделі для сегментації зображень

Segment Anything Model (SAM) від Meta AI є новою передовою математичною моделлю семантичної сегментації зображень, яка не потребує попереднього навчання на специфічних наборах даних із зазначеними об'єктами та сценами. Однією з її найвидатніших особливостей є підтримка всіх типів підказок: точок, рамок чи текстових запитів для сегментації об'єктів. За даними Meta AI, модель демонструє IoU в межах від 0,5 до 0,8 на стандартних наборах даних, таких як COCO та LVIS [3].

Модель практично інтегрується з такими технологіями, як Lang-Segment Anything, що забезпечує можливість використання текстових запитів для автоматичної сегментації об'єктів. Це особливо корисно для завдань, де необхідно швидко ідентифікувати об'єкти на зображенні без будь-яких візуальних підказок або спеціального навчання.

SAM забезпечує можливість подальшого збільшення IoU до 0,9 у складних сценаріях за допомогою текстових або точкових підказок, що дозволяє досягати високої надійності сегментації для визначення пошкоджень [4]. Це робить її ефективним інструментом для систем моніторингу дорожнього стану. У таких випадках модель довела свою ефективність у точному виявленні пошкоджень завдяки здатності виявляти дефекти.

Таким чином, завдяки здатності працювати з мінімальною кількістю підказок і адаптуватися до різних типів зображень, SAM є ідеальним рішенням для автоматизації процесу сегментації в складних умовах, таких як інфраструктурний моніторинг стану дорожнього покриття.

4.4. Розробка алгоритмів системи

4.4.1. Розробка алгоритму руху ходової лабораторії

На рис. 4.10 представлено алгоритм руху ходової лабораторії, який містить стратегію збору та зберігання зібраних даних ходовою лабораторією. Так, локальне сховище здатне накопичувати масив даних, які згодом пакетно відправляються на централізоване хмарне сховище.

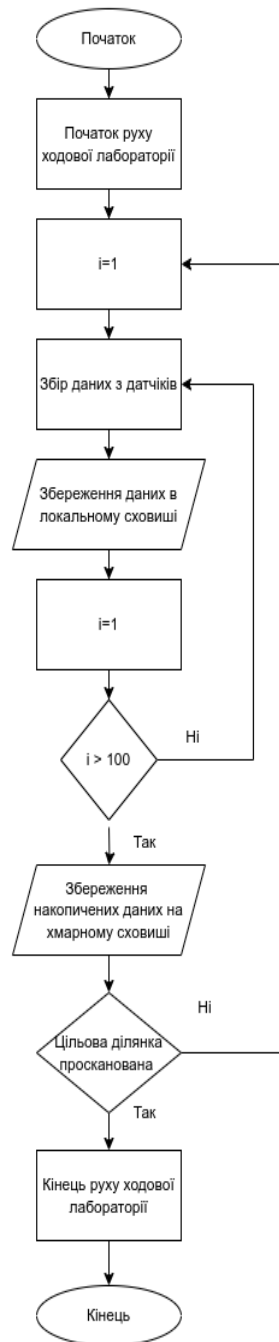


Рис. 4.10. Алгоритм руху ходової лабораторії

4.4.1. Розробка алгоритму підсистеми обробки зображень

Розроблений алгоритм підсистеми обробки зображень наведений на рис.

4.11.



Рис. 4.11. Алгоритм підсистеми обробки зображень

ВИСНОВКИ ДО РОЗДІЛУ 4

1. Зібрано та оброблено навчальний набір даних з пар видимого та інфрачервоного спектрів з дефектами дорожнього покриття. Що дало змогу подальшої розробки підсистеми опрацювання зображень отриманих з камер ходових лабораторій.

2. Розроблено структурну схему автоматизованої системи діагностики та ремонту дорожнього покриття з елементами штучного інтелекту.

3. Розроблено та обґрунтовано процеси попередньої обробки для обох спектрів (видимого та інфрачервоного), що дало змогу прибрати інформаційні шуми та деякі дефекти зображень для підвищення точності діагностики в системі.

4. Навчено модель для визначення типу дефектів дорожнього покриття. Запропонована конфігурація показала підвищення отриманої точності визначення типів дефектів в порівнянні з аналогами.

5. Обрано та обґрунтовано вибір нейронної мережі Segment Anything Model для сегментації, яка необхідна для визначення площі дефектів. Така модель показує високу точність 90% сегментації при використанні текстових підказок.

6. Розроблено алгоритми руху ходової лабораторії та підсистеми обробки зображень.

Список використаних джерел до розділу 4

- [1] Hwang, H., & Haddad, R. A. (1995). Adaptive median filters: new algorithms and results. *IEEE Transactions on image processing*, 4(4), 499-502.
- [2] Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In *Graphics gems IV* (pp. 474-485).
- [3] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4015-4026).
- [4] Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., ... & Zhang, L. (2023). Grounding dino: Marrying dino with grounded pre-training for open-set object detection. *arXiv preprint arXiv:2303.05499*.
- [5] Chen, C., Chandra, S., Han, Y., & Seo, H. (2021). Deep learning-based thermal image analysis for pavement defect detection and classification considering complex pavement conditions. *Remote Sensing*, 14(1), 106.
- [6] Zhou, S., & Song, W. (2020). Deep learning-based roadway crack classification using laser-scanned range images: A comparative study on hyperparameter selection. *Automation in Construction*, 114, 103171.
- [7] Zhong, M., Sui, L., Wang, Z., & Hu, D. (2020). Pavement crack detection from mobile laser scanning point clouds using a time grid. *Sensors*, 20(15), 4198.
- [8] Zhang, Y., Chen, C., Wu, Q., Lu, Q., Zhang, S., Zhang, G., & Yang, Y. (2018). A kinect-based approach for 3D pavement surface reconstruction and cracking recognition. *IEEE Transactions on Intelligent Transportation Systems*, 19(12), 3935-3946.

РОЗДІЛ 5. Експериментальна перевірка роботи системи

5.1. Формування тестової вибірки для імітаційного моделювання

Для проведення імітаційного моделювання зі сформованої в 4 розділі тестової вибірки (300 пар зображень) було сформовано 3 набори даних, які складаються з пар зображень видимого та інфрачервоного зображень, об'єм кожного набору 100 пар зображень що симулює проаналізований проміжок дороги площею 100 м^2 , детальний опис наборів даних надано в табл 5.1.

Щоб спростити моделювання вважаємо що тестовий проміжок дорожнього покриття належить до першої категорії дороги.

Таблиця. 5.1. Опис наборів даних для імітаційного моделювання системи діагностики та ремонту дорожнього покриття

Назва	Опис
Набір з низьким рівнем дефектів	90% цього набору це без дефектна ділянка, 10% різнотипні тріщини
Набір з середнім показником дефектності	40% набору складають різнотипні дефекти, решту бездефектна ділянка
Набір з високим рівнем дефектності	10% набору без дефектна ділянка, 90% складають різнотипні дефекти (більшість з яких вибоїни)

Таке розбиття дає перевірити працездатність системи на більшості випадків її використання.

5.2. Реалізація конвеєра обробки зображень та системи прийняття рішень

Дані елементи системи було реалізовано мовою програмування

Python на базі хмарних сервісів компанії Amazon (англ. Amazon Web Services (AWS)).

Для зберігання не структурованих файлів найкраще підходить сервіс Amazon Simple Storage Service (S3), надає дешевий спосіб зберігання масивних не структурованих файлів та має підтримувальний та задокументований програмний інтерфейс для взаємодії, що підходить для зберігання великої кількості зображень поверхні дорожнього покриття.

Конвеєр обробки зображень реалізовано на базі сервісу Amazon Lambda, який надає змогу користуватися без серверними хмарними технологіями (Serverless), така хмарна архітектура дає можливість отримувати доступ до обчислювальних потужностей лише на час запуску програми, що значно скорочує витрати на хмарні обчислення.

Механізм логічного висновку також доречно реалізувати на Amazon Lambda.

Усі бази даних системи це PostgreSQL, який є надійною реляційною базою даних з відкритим кодом. Ця база даних є стандартом в розробці програмного забезпечення, до того ж її використання безкоштовне. Для розгортання баз і доступності до даних з всіх пристроїв, доцільно скористатися Amazon Relational Database Service (Amazon RDS).

Таким чином було реалізовано централізовані підсистеми обробки зображень та підсистему прийняття рішень на базі AWS. Хмарна архітектура дала змогу гнучко налаштовувати обчислювальні потужності та швидко впроваджувати нові версії підсистем.

5.3. Аналіз результатів імітаційного моделювання автоматизованої системи діагностики та ремонту дорожнього покриття

5.3.1. Результати імітаційного моделювання на наборі з низьким рівнем дефектів

Після проходження даних через підсистему обробки зображень було

отримано результати, які показано в табл 5.2.

Таблиця. 5.2. Результати, що отримані після проходження підсистеми обробки зображень

Тип дефекту	Реальний % в наборі	Визначений % в наборі	Реальна площа дефектів, м ²	Визначена площа дефектів, м ²
Без дефектна поверхня	90	90	-	-
Косі тріщини	2	1	0.13	0.06
Тріщини і повздовжні	1	1	0.05	0.04
Тріщини і поперечні	3	2	0.14	0.09
Сітка тріщин	4	6	3.2	3.78
Сума	100	100	3.52	3.97

Отримані результати демонструють не відповідність визначених та реальних типів тріщин, що спостерігалось і при розробці нейронної мережі, так, 2 зображення з поперечними та косими тріщинами були визначені як сітка тріщин. Цей приклад демонструє значний вплив похибки визначення типу тріщин на кінцевий результат, через невеликий відсоток вмісту дефектів в наборі даних. Таким чином, було отримано 98% точність визначення класу дефекту, через те, що тип дефекту використовується як текстова підказка для сегментації. Похибка визначення типу дефекту також впливає на точність визначення площі, отримана похибка склала 0.41 м².

Отже, за результатами перевірки дорожнього покриття площею 100

м² категорії I, виявлено такі дефекти:

- Косі тріщини: 0.06 м²
- Тріщини поперечні: 0.09 м²
- Сітка тріщин: 3.78 м²

Загальна площа дефектів становить 3.93 м².

Рекомендації підсистеми прийняття рішень.

Оскільки поточна дефектність менша за **5%**, рівень дефектності не досягає порогових значень для планових ремонтних робіт, пропонується:

1. Критичність та необхідність ремонту:

- а. Поточний стан дорожнього покриття не є критичним.
- б. Провести профілактичні заходи для запобігання подальшому розвитку дефектів.

2. Тип ремонту:

- а. Локальний поверхневий ремонт з герметизацією наявних тріщин.
- б. Застосування захисних покриттів для підвищення зносостійкості.

3. Необхідна техніка та обладнання:

- а. Машини для очищення та підготовки поверхні.
- б. Обладнання для герметизації тріщин.
- с. Набір ручних інструментів для дрібних ремонтних робіт.

4. Терміни виконання:

- а. Рекомендується виконати роботи протягом найближчих **12 місяців**.

5. Додаткові рекомендації:

- а. Повторна діагностика стану покриття через **12 місяців** після ремонту для оцінки його ефективності.

5.3.2. Результати імітаційного моделювання на наборі з середнім рівнем наявності дефектів

Результати отримані на наборі з середнім рівнем наявності дефектів показано в табл 5.3.

Таблиця. 5.3. Результати, що отримані після проходження підсистеми обробки зображень

Тип дефекту	Реальний % в наборі	Визначений % в наборі	Реальна площа дефектів, м ²	Визначена площа дефектів, м ²
Без дефектна поверхня	60	60	-	-
Вибоїни	10	10	12.05	11.7
Косі тріщини	10	9	0.61	0.52
Тріщини повздовжні	10	10	0.39	0.36
Тріщини поперечні	5	5	0.22	0.25
Сітка тріщин	5	6	3.85	4.62
Сума	100	100	17.12	17.45

Як видно з отриманих результатів точність визначення типу дефектів склала 99%, так само як в попередньому прикладі, похибку вносить не правильне визначення типу тріщини. Похибка визначення загальної площі склала 0.38 м², як бачимо, похибка присутня при визначенні площі кожного типу дефекту, однак, в сумі вона не досягає критичних значень.

Отже, за результатами перевірки дорожнього покриття площею 100 м² категорії I, виявлено такі дефекти:

- Вибоїни: 11.7 м²
- Косі тріщини: 0.52 м²
- Тріщини повздовжні: 0.36 м²
- Тріщини поперечні: 0.25 м²
- Сітка тріщин: 4.62 м²

Загальна площа дефектів: 17.45 м².

Рекомендації підсистеми прийняття рішень.

Оскільки дефектність становить **17.45%**, це відповідає **Рівню 2 (ДП 2)**, пропонується:

1. Критичність та необхідність ремонту:

- a. Стан дорожнього покриття є незадовільним і потребує планового ремонту.
- b. Необхідно запобігти подальшому розвитку дефектів, що можуть призвести до погіршення безпеки руху.

2. Тип ремонту:

- a. Необхідний ремонт із частковою заміною верхнього шару покриття на дефектних ділянках.
- b. Латання вибоїн з використанням гарячих асфальтобетонних сумішей.
- c. Ремонт тріщин шляхом герметизації та заповнення спеціальними матеріалами.
- d. Укріплення основи в місцях значних пошкоджень.

3. Необхідна техніка та обладнання:

- a. Машини для очищення та підготовки поверхні.
- b. Асфальтоукладачі для нанесення нового покриття
- c. Катки (вібраційні та статичні) для ущільнення асфальтобетону.
- d. Самоскиди для транспортування матеріалів.

- е. Обладнання для герметизації тріщин.
 - ф. Набір ручних інструментів для дрібних ремонтних робіт.
4. Терміни виконання:
- а. Рекомендується виконати роботи протягом найближчих **3 місяців**.
5. Додаткові рекомендації:
- а. Повторна діагностика стану покриття через **6 місяців** після ремонту для оцінки його ефективності.
 - б. Після завершення ремонту провести маркування дорожнього покриття та встановити необхідні дорожні знаки.

5.3.3. Результати імітаційного моделювання на наборі з високим рівнем наявності дефектів

Результати отримані на наборі з високим рівнем наявності дефектів показано в табл 5.4.

Таблиця. 5.4. Результати, що отримані після проходження підсистеми обробки зображень

Тип дефекту	Реальний % в наборі	Визначений % в наборі	Реальна площа дефектів, м²	Визначена площа дефектів, м²
Без дефектна поверхня	10	10	-	-
Вибоїни	50	49	30.6	29.7
Косі тріщини	10	10	0.62	0.64
Тріщини повздовжні	10	10	0.37	0.33

Тріщини поперечні	10	9	0.34	0.3
Сітка тріщин	10	12	7.2	8.4
Сума	100	100	39.14	39.37

На цьому наборі даних бачимо хибне визначення дефекту типу вибоїна, як сітку тріщин, це обумовлено наявністю великої кількості малих вибоїн на одному зображенні. Загалом точність виявлення типів дефектів на наборі з високим рівнем дефектності склала 98%, а похибка визначення площі 0.23 м².

Отже, за результатами перевірки дорожнього покриття площею 100 м² категорії I, виявлено такі дефекти:

- Вибоїни: 29,7 м²
- Косі тріщини: 0,64 м²
- Тріщини повздовжні: 0,33 м²
- Тріщини поперечні: 0,3 м²
- Сітка тріщин: 8,4 м²

Загальна площа дефектів: 39,37 м².

Рекомендації підсистеми прийняття рішень.

Оскільки дефектність становить 39,37%, це відповідає Рівню 3 (ДП 3), пропонується:

1. Критичність та необхідність ремонту:

- a. Стан дорожнього покриття є незадовільним і вимагає негайного втручання.
- b. Висока концентрація вибоїн та тріщин може призвести до аварійних ситуацій та негативно впливає на безпеку дорожнього руху.

2. Тип ремонту:

- a. Капітальний ремонт з повною заміною верхнього шару покриття.
 - b. Регенерація основи в місцях значних пошкоджень.
 - c. Ремонт тріщин шляхом герметизації та заповнення спеціальними матеріалами.
 - d. Поліпшення дренажної системи для запобігання накопиченню води та подальшому руйнуванню покриття.
 - e. Зміцнення узбіч та ремонт обочин для забезпечення цілісності дорожнього полотна.
3. Необхідна техніка та обладнання:
- a. Машини для очищення та підготовки поверхні.
 - b. Грейдери та бульдозери для підготовки основи.
 - c. Асфальтоукладачі для нанесення нового покриття
 - d. Катки (вібраційні та статичні) для ущільнення асфальтобетону.
 - e. Самоскиди для транспортування матеріалів.
 - f. Екскаватори для земляних робіт.
 - g. Обладнання для дренажних робіт.
4. Терміни виконання:
- a. Рекомендується розпочати роботи протягом 1 місяця.
5. Додаткові рекомендації:
- a. Після завершення ремонту провести маркування дорожнього покриття та встановити необхідні дорожні знаки.
 - b. Повторна діагностика стану покриття через 6 місяців після ремонту для оцінки його ефективності.
 - c. Провести аналіз причин швидкого зношення попереднього покриття для запобігання повторенню таких ситуацій у майбутньому.

По результатам імітаційного модулювання підсистеми обробки даних і підтримки прийняття рішень моделювання отримано точність виявлення

основних типів дефектів дорожнього покриття на рівні 98-99%. Похибка визначення площі дефектної поверхні не перевищила 1 м² на ділянці 100 м². Результати говорять про значний вплив на визначення дефектної площі поверхні дороги, що діагностується, точності визначення типу дефекту. Якість отриманих рекомендацій говорить про зменшення необхідного часу експертів для остаточного рішення, щодо ремонту ділянки дороги. Таким чином було підтверджено працездатність запропонованої системи.

Грунтуючись на отриманих результатах моделювання можна зробити загальні висновки щодо працездатності запропонованої системи та порівняти її системою запропонованою у роботі [1] та підсистемою класифікації дефектів дорожнього покриття описано в роботі [2].

Головною відмінністю від системи [1] є використання камер двох спектрів інфрачервоного та видимого, а також використання блоку комплексування для зменшення впливу інформаційних завад і зменшення складності подальшої обробки. Таким чином було замінено використання дорого вартісного лазерного сканеру на більш дешевий тепловізійний прилад.

Запропонована в роботі [2] підсистема підсистеми класифікації дефектів дорожнього покриття з попереднім мультиспектральним комплексуванням зображень видимого та інфрачервоного спектрів набула подальшого розвитку коштом заміни методу комплексування на запропонований в цій роботі метод злиття зображень з адаптивним визначенням вагових коефіцієнтів на основі вейвлет перетворення, що на 1.4% підвищує ентропію Шенона(інформативність) в порівнянні з аналогами, а також, завдяки запропонованому методу аргументації навчальних даних на основі злиття зображень для задач класифікації. Така конфігурація дала збільшити вірогідність правильного визначення типу дефектів на 0.66% (до 99%) та зменшити час необхідний на навчання нейромережевого класифікатора 4.5 рази коштом скорочення необхідного

об'єму навчальних даних з 13,5 тисяч до 3 тисяч зображень.

ВИСНОВКИ ДО РОЗДІЛУ 5

1. Реалізовано на базі хмарних сервісів Amazon конвеєри обробки зображень та підсистему прийняття рішень. Що дозволило скоротити час на впровадження нових версій і значно прискорило проведення моделювання.

2. Проведено імітаційне моделювання системи, за результатами якого підтверджено працездатність запропонованої системи. Отримано практичні результати точності виявлення типу дефектів на рівні 98-99% та визначено що похибка розрахунку площі не перевищує 1м².

Список використаних джерел до розділу 5

- [1] Roberts, R. A. (2021). A new approach to Road Pavement Management Systems by exploiting Data Analytics, Image Analysis, and Deep Learning (Doctoral dissertation, Doctoral dissertation, University of Palermo).
- [2] Chen, C., Chandra, S., Han, Y., & Seo, H. (2021). Deep learning-based thermal image analysis for pavement defect detection and classification considering complex pavement conditions. *Remote Sensing*, 14(1), 106.

ВИСНОВКИ

1. Проаналізовано сучасний стан розвитку методів автоматизованого діагностування стану дорожнього покриття, що є актуальним завданням для забезпечення безпеки дорожнього руху та оптимізації витрат на ремонтні роботи. Серед основних проблем виділено високу вартість точних методів діагностування та значний вплив інформаційних завад на точність результатів. Запропоновано мультимодальний підхід із комплексуванням зображень, який дозволяє ефективно протидіяти інформаційним завадам, знижуючи їхній вплив на точність діагностики, а також зменшує складність подальшої обробки даних у різних спектральних діапазонах.

2. Для проведення досліджень зібрано 6000 зображень дефектів дорожнього покриття у двох спектральних діапазонах. Розроблено та обґрунтовано процес обробки цих зображень з використанням нейронних мереж, який полягає в двох паралельних процесах попередньої обробки зображень обох спектральних діапазонів, комплексування зображень єдине більш інформативне, класифікації типу дефекту дорожнього покриття нейромережевим класифікатором та визначенні площі дефекту завдяки сегментації на базі нейронної мережі.

3. У результаті проведення досліджень було запропоновано та обґрунтовано метод злиття зображень з адаптивним визначенням вагових коефіцієнтів на основі вейвлет перетворення. Метод полягає в розкладанні початкових зображень на вейвлет коефіцієнти та визначені згортковою нейронною мережею вагових коефіцієнтів, що визначають частку початкового зображення в результуючому, і дозволяє збільшити інформативність результуючого зображення при проведенні діагностування дорожнього покриття. Було досліджено приріст інформативності при використанні запропонованого методу. За результатами порівняння запропонованого методу з наявними було

отримано приріст ентропії Шенона (інформативності) на 1.4%, що є основною оцінкою якості злиття.

4. У результаті проведених досліджень було розроблено метод аугментації навчальних даних на основі злиття зображень для задач класифікації. Метод полягає в поєднанні вейвлет коефіцієнтів отриманих з початкових зображень одного об'єкту та зворотному перетворенню для отримання аугментованого зображення. Статистичне порівняння з іншими методами змішування зображень виявило залежність ймовірності правильної класифікації від обраного методу аугментації під час навчання нейронної мережі. Найбільший приріст ймовірності (до 22%) було досягнуто за допомогою запропонованого методу та методу випадкової заміни вейвлет-коефіцієнтів, водночас статистично значущих відмінностей між вибірками отриманих ймовірностей не виявлено. Запропонований метод продемонстрував утричі вищу продуктивність порівняно з методом випадкової заміни вейвлет-коефіцієнтів, оскільки для його реалізації достатньо двох вихідних зображень, тоді як метод випадкової заміни вейвлет-коефіцієнтів потребує шести зображень.

5. Удосконалено систему автоматизованого діагностування стану дорожнього покриття на основі запропонованих підсистем опрацювання (обробки) зі злиттям зображень двох спектрів з адаптивним визначенням вагових коефіцієнтів та методів аугментації, що дало підвищення ймовірності коректного визначення типу дефекту дорожнього покриття на 0.66% та скоротити об'єм навчальних даних до 3 тисяч зображень (аналог 13,5 тисяч).

6. У результаті проведених досліджень було досягнуто підвищення точності оцінки стану дорожнього покриття при використанні теплових та видимих камер на 0.66%, завдяки високій ймовірності (99%) коректного визначення типу дефекту та похибці у визначенні площі дефекту — менш ніж 1 м² на ділянці площею 100 м². Ці показники є ключовими аспектами

якості системи, оскільки висока точність у класифікації дефектів і мінімальні відхилення у розрахунках площі підвищують довіру до результатів діагностики. Завдяки вдосконаленню автоматизованої системи діагностики було також знижено обсяг ручної необхідної для корекції похибки системи, що дозволило скоротити трудомісткість процесу корекції похибки при оцінці стану дорожнього покриття на 39%.

ДОДАТКИ

Додаток А

**Діагностичні
Прилади**
НАУКОВО-ВИРОБНИЧА ФІРМА

Адреса офісу:

м. Київ, 03061, вул. Патріотів, 103
Тел./факс: (+38044)-239-12-50
(+38044)-239-12-46
e-mail: support@ndt-ua.com
www.ndt-ua.com



Вих. № 1503/1 від «15» 03 2024

1503/1 big 15.03

АКТ ВПРОВАДЖЕННЯ результатів наукових досліджень «Використання методу аугментації для навчання нейронних мереж при застосуванні їх у тепловому неруйнівному контролі»

Ми, що підписалися нижче, склали цей акт про те, що результати наукової роботи з використання методу аугментації на базі комплексування зображень дозволили розробити та протестувати методику обробки та аналізу інформативних параметрів сигналів тепловізора для подальшого їх застосування в нейронних мережах. В процесі досліджень розроблено метод аугментації на базі комплексування зображень, який успішно пройшов експериментальні дослідження на серіях зразків дефектів, що підтвердило можливість використатися цього методу в задачах контролю. Завдяки цьому є можливість зменшити собівартість процесу автоматичної класифікації дефектів тепловим методом та підвищити точність контролю.

Науково-виробнича фірма «Діагностичні Прилади» підтверджує готовність до налагодження співпраці з розробниками цієї методики для подальшого впровадження результатів цього дослідження у виробничому процесі за умови регулярної підтримки працездатності вказаного методу.

Генеральний директор
«НВФ «ДІАГНОСТИЧНІ ПРИЛАДИ»



Олександр ПАВЛІЙ
«15» 03 2024 р.

Виконавці:

[Signature] Анатолій ПРОТАСОВ

[Signature] Денис СТОРОЖИК

«15» 03 2024 р.

Додаток Б



УКРАЇНА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

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

ПРИЛАДОБУДІВНИЙ ФАКУЛЬТЕТ

03056, м. Київ, Берестейський пр-т, 37; тел. (+38 044) 204-94-55 тел./факс (+38 044) 236-10-43
<http://pbf.kpi.ua> e-mail: pbf@kpi.ua

25.11.2014 № 30
на № _____ від _____

АКТ ВПРОВАДЖЕННЯ

результатів наукових досліджень аспіранта Сторожика Дениса
Володимировича в навчальний процес кафедри Автоматизації та систем
неруйнівного контролю Приладобудівного факультету

Ми, що підписалися нижче, склали цей акт про те, що частина результатів дисертаційної роботи Сторожика Д.В., а саме методика опрацювання та аналізу теплових зображень, були впроваджені в навчальний процес у вигляді лабораторної роботи для студентів кафедри Автоматизації та систем неруйнівного контролю. Лабораторна робота на тему «Дослідження алгоритмів опрацювання теплових зображень» виконується в рамках дисципліни «Основи теплобачення та систем технічного зору роботів» та присвячена дослідженню алгоритмів опрацювання теплових зображень, отриманих за допомогою тепловізійних систем.

В роботі розглядається використання функціоналу бібліотек, які відповідають за зчитування зображень, визначення їх параметрів, роботу з пікселями області інтересу, а також розглянуто фільтраційні можливості для покращення та деталізації елементів зображень. Опанування вказаної лабораторної роботи допомагає студентам набутти практичних навичок роботи з результатами теплового контролю, а також сприяє розвитку дослідницьких здібностей.

Декан

Приладобудівного факультету

Завідувач кафедри

«Автоматизації та систем
неруйнівного контролю»



Григорій ТИМЧИК

Юрій КИРИЧУК

Додаток В

Функція Python для аугментації запропонованим методом

```
import cv2
import numpy as np
import pywt
import random

def wavelet_transform_rgb(ir_img: np.ndarray, vs_img: np.ndarray, fusion=None) -> np.ndarray:
    if fusion is None:
        fusion = conditional_averaging_rgb
    if vs_img.shape[2] != 3 or ir_img.shape[2] != 4:
        raise ValueError("Перше зображення повинно мати 3 канали (RGB), а друге 4 канали (RGBA).")
    channels = []
    for i in range(3):
        channel_img1 = vs_img[:, :, i]
        channel_img2 = ir_img[:, :, i]
        alpha2 = ir_img[:, :, 3]

        channel_out = _wavelet_channel_Transform_rgb(channel_img1, channel_img2, alpha2, vs_img.shape[:2], fusion)
        channels.append(channel_out)

    outImage = cv2.merge(channels)
    return outImage.astype(np.uint8)

def _wavelet_channel_Transform_rgb(channel_img1: np.ndarray,
                                   channel_img2: np.ndarray,
                                   alpha2: np.ndarray,
                                   shape: tuple,
                                   fusion) -> np.ndarray:

    cooef1 = pywt.dwt2(channel_img1, 'db5', mode='periodization')
    cooef2 = pywt.dwt2(channel_img2, 'db5', mode='periodization')

    cA1, (cH1, cV1, cD1) = cooef1
    cA2, (cH2, cV2, cD2) = cooef2

    cA = fusion(cA1, cA2, alpha2)
    cH = fusion(cH1, cH2, alpha2)
    cV = fusion(cV1, cV2, alpha2)
    cD = fusion(cD1, cD2, alpha2)

    fincoC = cA, (cH, cV, cD)
    outImageC = pywt.idwt2(fincoC, 'db5', mode='periodization')

    outImageC = np.multiply(np.divide(outImageC - np.min(outImageC), (np.max(outImageC) - np.min(outImageC))), 255)
    outImageC = cv2.resize(outImageC, (shape[1], shape[0]))

    # Маскування результату з альфа-каналом
    outImageC = np.where(alpha2 > 0, outImageC, channel_img1)
    return outImageC

def conditional_averaging_rgb(matrix1: np.ndarray, matrix2: np.ndarray, alpha2: np.ndarray) -> np.ndarray:
    alpha2_resized = cv2.resize(alpha2, (matrix1.shape[1], matrix1.shape[0]))
    result = np.where(alpha2_resized > 0, (matrix1 + matrix2) / 2, matrix1)
    return result
```

Додаток Г

Функція Python для комплексування запропонованим методом

```
import cv2
import numpy as np
import pywt
import random

def wavelet_transform(image, wavelet='db5'):
    cA, (cH, cV, cD) = pywt.dwt2(image, wavelet, mode = 'periodization')

    concatenate_coeffs = np.stack((cA, cH, cV, cD), axis=0)

    return concatenate_coeffs

def adaptive_image_fusion(ir_img, vs_img, gray_ir_img, gray_vs_img, alpha):
    if not alpha:
        ir_w = torch.tensor(wavelet_transform(gray_ir_img)).unsqueeze(0)
        vs_w = torch.tensor(wavelet_transform(gray_vs_img)).unsqueeze(0)

        with torch.no_grad():
            alpha = MODEL(ir_w.float(), vs_w.float())
    print(alpha)
    f_img = imgfusion.adaptive_wavelet_transform(ir_img, vs_img, alpha if alpha else alpha.item())
    return f_img

def adaptive_wavelet_transform(ir_img: np.ndarray, vs_img: np.ndarray, alpha: float) -> np.ndarray:
    fision = lambda vs_matrix, ir_matrix, image_alpha: _adaptive_conditional_averaging(vs_matrix, ir_matrix, image_alpha, alpha)
    return wavelet_transform_rgb(ir_img, vs_img, fision)

def _adaptive_conditional_averaging(vs_matrix: np.ndarray,
                                     ir_matrix: np.ndarray,
                                     image_alpha: np.ndarray,
                                     alpha: float) -> np.ndarray:
    image_alpha_resized = cv2.resize(image_alpha, (vs_matrix.shape[1], vs_matrix.shape[0]))
    result = np.where(image_alpha_resized > 0, (alpha * ir_matrix + (1-alpha) * vs_matrix), vs_matrix)
    return result
```