

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Тернопільський національний технічний університет імені Івана Пулюя

(повне найменування вищого навчального закладу)

Факультет прикладних інформаційних технологій та електроінженерії

(назва факультету)

Кафедра автоматизації технологічних процесів та виробництв

(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА
на здобуття освітнього ступеня

бакалавр

(освітній рівень)

*Застосування нейронних мереж для контролю дефектності
металоконструкцій*

(комплексна тема)

Виконавли: студенти 4 курсу, групи КАс-41
Спеціальність 151

*“Автоматизація та комп’ютерно-інтегровані
технології”*

(шифр і назва спеціальності)

Поліник Назарій
Іванович

(підпис)

(прізвище та ініціали)
Рибець Сергій
Андрійович

(підпис)

(прізвище та ініціали)

Керівник

Марущак П.О.

(підпис)

(прізвище та ініціали)

Нормоконтроль

Козбур І.Р.

(підпис)

(прізвище та ініціали)

Рецензент

(підпис)

(прізвище та ініціали)

м. Тернопіль – 2023

А н о т а ц і я

Поліник Назарій Іванович, Рибець Сергій Андрійович «Застосування нейронних мереж для контролю дефектності металоконструкцій» (комплексна тема).

В дипломній роботі використані такі терміни: нейронні мережі, навчання, автоматизоване розпізнання, фон, дефект.

Об'єктом дослідження є процес автоматизації дефектомерії з застосування глибоких нейронних мереж.

Метою роботи є розробка та тестування методу інтелектуалізованої дефектометрії.

Виявлення металевих дефектів є складним завданням, оскільки зображення металевих поверхонь легко піддається впливу факторів навколишнього середовища, таких як освітлення та відбивання світла. У цьому дослідженні розроблено новий метод виявлення дефектів та застосовано на зображеннях металевих поверхонь. Запропонована система розпізнавання зображень працює ефективно, з точки зору виявлення дефектної області та класифікації типів дефектів.

Майбутні дослідження будуть зосереджені на системах з кількома підсвічуваннями та вхідними зображеннями, наприклад фотометричними стерео. За допомогою цього методу можна проводити дослідження виявлення дефектів для різних матеріалів поверхні (текстиль, скло тощо). Крім того, запропонована навчена модель, у перспективі буде доступна через веб-інтерфейс, що забезпечує онлайн виявлення та дефектометрію дефектів.

Annotation

Polynyk Nazariy Ivanovich, Rybets Serhiy Andriyovych "*Application of neural networks to control the defectiveness of metal structures*" (complex topic).

The following terms were used in the diploma work: neural networks, learning, automated recognition, background, defect.

The object of research is the process of automating defectometry using deep neural networks.

The purpose of the work is to develop and test the method of intellectualized defectometry.

Detecting metal defects is challenging because the imaging of metal surfaces is easily affected by environmental factors such as illumination and light reflection. In this study, a new defect detection method was developed and applied to images of metal surfaces. The proposed image recognition system works effectively in terms of defect area detection and defect type classification.

Future research will focus on systems with multiple illuminations and input images, such as photometric stereo. With the help of this method, it is possible to conduct research on the detection of defects for various surface materials (textiles, glass, etc.).

In addition, the proposed trained model will in the future be available through a web interface that provides online detection and defectometry of defects.

ЗМІСТ

ВСТУП	5
1. АНАЛІТИЧНА ЧАСТИНА	
1.1. Взаємозв'язок між штучним інтелектом, машинним навчанням і глибоким навчанням	7
1.2. Основні типи нейронних мереж та їх діагностичні можливості	13
2. ПРОЕКТНА ЧАСТИНА	
2.1. Оптико-цифрові методи контролю стану поверхні смуг металопрокату	22
2.2. Технічне оснащення фотозйомки та його параметри	25
2.3. Отримання зображень для дослідження	30
2.4. Розпізнавання зображень за допомогою CNN	33
2.5. Дослідження впливу освітлення на результат розпізнавання	38
2.6. Дефектоскопія зображень дефектів	40
2.7. Дефектометрія зображень дефектів	47
2.8. Вплив освітленості на параметри дефектів	50
2.9. Вплив частоти вібрації на параметри дефектів	52
3. СПЕЦІАЛЬНА ЧАСТИНА	
3.1. Алгоритм локалізації та класифікації поверхневих дефектів на сталевому листі	65
3.2. Дефекти та їх локалізація	67
4. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ	
4.1. Актуальність охорони праці	75
4.2. Правила техніки безпеки при роботі в лабораторіях	77
4.3. Санітарно-гігієнічні вимоги до лабораторного приміщення	78
ВИСНОВКИ	80
ПЕРЕЛІК ПОСИЛАНЬ	81

ВСТУП

Поява нейромереж вивела людські технології на якісно новий рівень, їх розвиток є безперервним. Ця технологія назавжди змінить суспільство, і, можливо, багато людей втратять свої робочі місця. Природно вважати, що нейромережі - результат епохи інтернету, проте насправді вони виникли раніше. Термін "нейромережа" сформульовано у середині минулого століття.

«Найперші спроби моделювати діяльність людського мозку ще в 1943 році описали нейрофізіолог Уоррен Мак-Каллок та математик Уолтер Піттс. Однак першу нейронну мережу, яка є попередником сучасного ШІ, запропонував нейрофізіолог Френк Розенблатт. Це була математична модель, що спрощено відтворювала принципи роботи біологічних нейронних мереж. Модель назвали перцептрон, іноді її також називають перцептрон Розенблатта.

Отже, підвалини сучасних нейромереж закладено у 40-50-х роках минулого століття. Вченими того часу реалізовано амбітну мету створення механізму (алгоритму), що відтворює функції людського мозку, зокрема здатність аналізувати інформацію. «Перцептрон, загалом це проста модель для машинного навчання, яку створено для допомоги обчислювальним машинам у навчанні на різному вигляді даних. Одним із перших застосувань нейромереж було розпізнавання зображень - саме для цього їх використав Френк Розенблатт. У ту епоху комп'ютери були громіздкими, повільними й дорогими, а доступ до обчислювальних ресурсів - обмеженим. Все це ускладнювало розвиток нейромереж, і вони «застигли» у розвитку на кілька десятиліть.

У 2006 році професор Джеффрі Хінтон з університету Торонто та його колеги представили статтю про глибокі нейронні мережі на конференції з комп'ютерного зору та обробки зображень. У цій статті було описано алгоритм навчання нейронних мереж із кількома шарами, який дозволяв значно покращити точність розпізнавання зображень. Цей алгоритм був названий «глибоким навчанням» (deep learning) і з того часу став широко застосовуватися в різних галузях, пов'язаних із аналізом даних та машинним навчанням».

Основним завданням нейромереж вчені бачать обробку великих масивів даних, що сильно скорочує витрати часу на ручні обчислення та інтерпретацію. І лише наприкінці ХХ – на початку ХХІ століття нейромережі почали розвивати для машинного навчання та створення штучного інтелекту.

Безконтактне діагностування поверхневих дефектів стає все більш важливою частиною виробничих систем, внаслідок високих вимог до якості поверхні. Машинне навчання досягло вражаючих показників розпізнавання в завданнях класифікації зображень. Для того, щоб використовувати ці можливості, у цьому проекті представлено виявлення та класифікацію поверхневих дефектів на гарячекатаній сталевій смузі за допомогою цифрових зображень за різних умов освітлення аналізованої поверхні.

1. АНАЛІТИЧНА ЧАСТИНА

1.1. Взаємозв'язок між штучним інтелектом, машинним навчанням і глибоким навчанням

У 1959 році Хабель і Візель провели експеримент, щоб зрозуміти, як зорова кора головного мозку обробляє візуальну інформацію (Хубель і Візель, 1959). Вони зафіксували активність нейронів зорової кори кота під час переміщення яскравої лінії перед котом. Вони помітили, що деякі комірочки спрацьовують, коли яскраву лінію показують під певним кутом і в певному місці (Вони назвали ці комірочки простими). Інші нейрони спрацьовували, коли яскрава лінія була показана незалежно від кута/розташування, і, здавалося, виявляли рух (Вони назвали ці складні клітини). Здавалося, що складні комірочки отримують вхідні дані від кількох простих комірок і мають ієрархічну структуру. Хьюбелі і Візель отримали Нобелівську премію за свої відкриття в 1981 році.

У 1980 році, натхненний ієрархічною структурою складних і простих комірок, Фукусіма запропонував Neocognitron (Fukushima 1988), ієрархічну нейронну мережу, яка використовується для розпізнавання рукописних японських символів. Neocognitron був першим CNN і мав власний алгоритм навчання. У 1989 році LeCun et. al. (LeCun та ін. 1989) запропонував CNN, який можна навчити за допомогою алгоритму зворотного поширення. CNN здобули величезну популярність, коли перевершили інші моделі на ILSVRC (ImageNet Large Scale Visual Recognition Challenge). ILSVRC — це змагання з класифікації та виявлення об'єктів на сотнях категорій об'єктів і мільйонах зображень. Конкурс проводиться щорічно з 2010 року по теперішній час, залучаючи участь понад п'ятдесят установ. Відомими архітектурами CNN, які виграли ILSVRC, є AlexNet у 2012 році, ZFNet у 2013 році, GoogLeNet та VGG у 2014 році і ResNet у 2015 році.

Глибоке навчання стало результатом досліджень штучного інтелекту та машинного навчання, рис. 1. Його розвиток зумовлений пошуком відповіді на

питання про те, коли програмовані комп'ютери зможуть стати такими ж інтелектуальними, як людський мозок, обробляти інформацію та вирішувати складні проблеми. Оскільки людський мозок здатний працювати з величезними обсягами даних і вирішувати проблеми на основі довготривалих знань, отже комп'ютер повинен генерувати достовірні знання з величезної кількості узагальненої інформації. Щоб імітувати процеси прийняття рішень і міркувань людиною, ШІ пробують «закласти» у машини, сприяючи їм навчатися з досвіду та адаптуватися до змін середовища.

На початку розроблення ШІ окремі проблеми у відносно простих середовищах швидко вирішувалися за допомогою списку формальних математичних правил, оскільки комп'ютер не вимагав багато знань про світ. Наприклад, шахова система Deep Blue була першим прикладом комп'ютерної системи, яка перемогла людину, перемігши чинного чемпіона світу в шаховому матчі за стандартного турнірного контролю часу. Однак шахи є відносно простим середовищем, що містить лише 64 локації та 32 фігури, які можуть рухатися певним чином.



Рис. 1.1. Взаємозв'язок між штучним інтелектом, машинним навчанням і глибоким навчанням

Відомо, що рух ігрових фігур може регулюватися набором формальних правил після розробки відповідної стратегії. Однак багато проблем у реальному світі є набагато складнішими. Вони не описуються простими правилами, тобто системи штучного інтелекту повинні навчитися на шаблонах необроблених даних. Ця можливість відома як машинне навчання (МН), що є а підвидом

штучного інтелекту, де алгоритми машинного навчання призначені для оптимізації виконання певного завдання на прикладах та/або досвіді. Навчання та висновок є двома кроками МН. Алгоритм обробляє набір даних під час навчання та вибирає функцію, яка найкраще відповідає шаблонам даних, із фіксованою моделлю після завершення навчання.

Потім модель використовують для виведення нових значень із нових прикладів. Поняття функції як детермінованої відображає входи і виходи, і мету МН, яка полягає в тому, щоб знайти функцію, яка відповідає відображенням вхідних характеристик і вихідних характеристик, які є шуканими в прикладах у наборі даних. Кілька аспектів ускладнюють вивчення завдань МН навіть за допомогою комп'ютера. МН є некоректним за умови, якщо набір можливих функцій перевищує кількість прикладів у множині даних: за таких умов, інформації, поданої в задачі, недостатньо для пошуку єдиного найкращого рішення; замість цього, є кілька можливих рішень.

Глибоке навчання (ГН) було введено як підвид МН, яке зосереджено на розробці та оцінюванні сучасних алгоритмів навчання нейронних мереж і архітектури моделей. Відомо, еволюцію глибокого навчання, де в перші роки воно не було «глибоким», оскільки застосовувало просту функцію активації. Воно мало лише один шар (шар - це структура або мережева топологія в архітектурі моделі), яка бере інформацію з попередніх шарів, а потім передає її на наступний рівень. Після того, як проблема «виключного або» (XOR) була порушена Мінським і Пейпером у 1969 році, стало зрозумілим, що багато завдань неможливо вирішити за допомогою лише лінійної функції. Як наслідок, з'єднуючи різні функціональні шари, об'єднання використовували для вирішення лінійних задач. Чим більше функціональних шарів необхідно для підключення та вирішення проблеми за аналізу великих даних і збільшення складності проблем тим складнішою є модель.

Алгоритми МН є машиноподібними, тобто їм все ще потрібна людина, щоб бути здатними виконати заплановану мету. Найбільша перевага алгоритмів глибокого навчання полягає в тому, що вони можуть вивчати категорії

поступово за допомогою своїх архітектур прихованого рівня, що переходить від категорій низького рівня до вищого рівня. Вони також можуть автоматично отримувати функції та класифікувати набори даних для подальшого оброблення на практиці, усуваючи потребу в експертних знаннях предметної області. Ще одна перевага глибокого навчання є те що воно обробляє величезні масиви даних забезпечуючи статистичну значимість результатів.

Проте, на відміну від традиційного МН, глибоке навчання вимагає високої продуктивності процесора і достатньо часу для навчання моделей завдяки великій кількості залучених параметрів.

«Можна сказати, що нейронна мережа – це «дерево» можливих варіантів, де кожна точка розгалуження є вузлом із математичною операцією. Однак, якщо уявляти нейронну мережу як дерево, то гілки у нього є різної товщини. Найчастіші результати будуть мати велику «вагу» при виборі варіантів. Деякі з них будуть вrostати в інші гілки через вагові коефіцієнти, які налаштовуються в процесі навчання і визначають ступінь впливу одного вузла на інший.

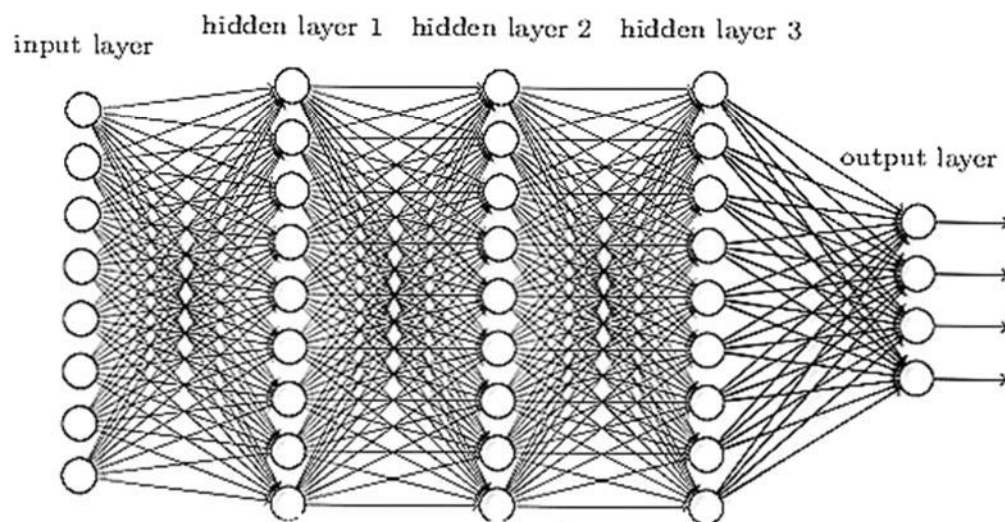


Рис. 1.2. Структура нейронної мережі з трьох шарів: вхідний шар; прихований шар, і вихідний шар.

Структура мережі традиційної CNN значною мірою залежить від експериментальних знань і не має встановлення параметра, який може бути основою, тому при визначенні оптимальної структури CNN потрібне

порівняння продуктивності через налаштування різних параметрів, іншу структуру CNN, а потім візьміть структуру оптимальної продуктивності CNN як остаточну мережеву структуру, це призводить до величезних втрат часу, перешкоджає подальшому розвитку CNN у великих обсягах даних. Для вищезгаданої ситуації пропонується багато методів покращення структури CNN, де основна частина полягає в створенні мережі заснований на традиційній структурі CNN, крім того, також є вчений, щоб запропонувати прийняти метод формули зростання від простого до складного для побудови мережі, але цей метод, нарешті, все ще потребує визначення остаточної структури мережі за допомогою порівняння продуктивності структури гетерогенних мереж, труднощі навчання розпізнавання великі, час навчання неконтрольований, неможливо реалізувати інтелектуальне розпізнавання роботизації.

Нейронні мережі можна умовно розділити як за їх функціями, так і за іншими ознаками - наприклад, типами навчання. Розглянемо кілька популярних класифікацій та познайомимось з найвідомішими представниками штучного інтелекту.

Класифікації нейронних мереж – залежно від їх архітектури та методів навчання.

За структурою нейронної мережі:

- одношарові нейронні мережі (Perceptron);
- багатошарові нейронні мережі (Multi-layer Perceptron);
- рекурентні нейронні мережі (Recurrent Neural Networks);
- згорткові нейронні мережі (Convolutional Neural Networks);
- автоенкодера (Autoencoder);
- глибокі нейронні мережі (Deep Neural Networks);
- резервні нейронні мережі (Residual Neural Networks).

За завданнями, що вирішуються нейронною мережею:

- Класифікація (Classification);
- Регресія (Regression);
- Сегментація зображень (Image Segmentation);

- детектування об'єктів (Object Detection);
- Машинний переклад (Machine Translation);
- Аналіз тональності (Sentiment Analysis);
- генерація тексту (Text Generation);
- Управління роботами (Robot Control).

За типом навчання:

- Навчання з учителем (Supervised Learning);
- Навчання без учителя (Unsupervised Learning);
- Навчання з підкріпленням (Reinforcement Learning).

«Кожен тип нейронної мережі є ефективним для певних завдань та потребує своїх методів навчання. Наприклад, згорткові нейронні мережі добре підходять для обробки зображень, тоді як рекурентні нейронні мережі використовують для обробки послідовностей даних, таких як тексти. Класифікації нейронних мереж допомагають вибрати найкращу архітектуру для вирішення конкретного завдань.

Ряд спеціалістів, у свою чергу, також відзначають трансформерні (GPT-like, BERT, T5), генеративно-змагальні (Alpha-Go, Alpha-Zero) та дифузійні (Midjourney, Stable Diffusion) моделі нейромереж. Концептуально нейромережі для генерації картинок схожі на ChatGPT - і ті, й інші навчилися розуміти і виконувати інструкції, причому це вміння у них з узагальненням - тобто виконують вони навіть ті інструкції, яких у навчальних даних не було.

Крім того, різні нейромережі, як і будь-яке ПЗ, потребують різних технічних потужностей. Якщо говорити про архітектуру мовних мереж, то технологія GPT-3 для сервісу ChatGPT є унікальною, її архітектуру складно відтворити на широко доступній апаратній базі (тільки параметри мережі вимагають близько 800 ГБ оперативної пам'яті для розгортання). Генеративні мережі для зображень простіші: Stable Diffusion, наприклад, може бути розгорнута на домашньому комп'ютері.

Розглянемо ти основні види нейромереж, їх пристрої та функціональних особливості.

1. *Прямі нейронні мережі* є найпростішими та широко використовуються. Вони містять шари нейронів, з'єднаних таким чином, що дані проходять через мережу лише в одному напрямку;

2. *Рекурентні нейронні мережі* забезпечують оброблення послідовних даних, таких як звук, мовлення чи текст. Вони мають циклічний зв'язок між нейронами, що дозволяє їм запам'ятовувати інформацію про попередні стани та використовувати її для прийняття рішень у майбутньому.

3. *Згорткові нейронні мережі* часто використовують для обробки зображень. Вони мають спеціальні шари, які можуть виявляти певні ознаки у зображень, такі як краї чи кольорові плями. Автокодувальники можуть використовуватися для вивчення прихованих закономірностей даних. Вони працюють шляхом перетворення вхідних даних у якесь приховане уявлення, а потім відновлюють вихідні дані з цього уявлення.

1.2. Основні типи нейронних мереж та їх діагностичні можливості

Різні типи нейронних мереж можуть бути використані для вирішення різних завдань обробки зображень, починаючи від простої бінарної класифікації (перевірка на відповідність зображення певним критеріям) і закінчуючи сегментацією екземплярів. Вибір правильного типу та архітектури нейронної мережі відіграє важливу роль у створенні ефективного рішення обробки зображень. Розглянемо кілька популярних видів нейронних мереж і задачі, для яких вони використовуються.

Згорткова нейронна мережа (Convolutional Neural Network, CNN) – це клас мереж глибинного навчання, які були створені спеціально для обробки зображень. ЗНМ складається з шарів входу та виходу, а також із декількох прихованих шарів. Приховані шари ЗНМ складаються зі згорткових шарів, агрегувальних шарів, повноз'єднаних шарів і шарів нормалізації. CNN широко

використовуються для впровадження ШІ в обробку зображень та вирішення таких задач, як обробка сигналів, класифікація зображень і розпізнавання зображень [2].

Mask R-CNN – це нейронна мережа на основі CNN, яка може бути використана для розділення об'єктів в обробленому зображенні або відео [3]. Ця нейронна мережа працює у два етапи: Сегментація – нейронна мережа обробляє зображення, виявляє ділянки, які можуть містити об'єкти, та формує пропозиції. Генерація обмежувальних коробок (bounding box) та масок – мережа обчислює двійкову маску для кожного класу та генерує кінцеві результати на основі цих розрахунків.

U-Net – це згорткова нейронна мережа, яка дозволяє швидко і точно сегментувати зображення. На відміну від інших нейронних мереж у нашому списку, *U-Net* був розроблений спеціально для біомедичної сегментації зображень. Тому вважають, що *U-Net* є кращим за *Mask R-CNN*, особливо в таких складних завданнях, як обробка медичних зображень. Генеративна змагальна мережа – це клас алгоритмів штучного інтелекту, що використовуються в навчанні без учителя, реалізовані системою двох штучних нейронних мереж, які змагаються одна з одною в рамках гри з нульовою сумою [4]. Ця методика дозволяє створювати фотографії, які для побіжного огляду людиною виглядають як справжні та мають багато реалістичних елементів, проте є згенерованими або обробленими комп'ютером.

CNN — це тип алгоритму глибокого навчання, який використовується для обробки даних із сітковою топологією. CNN – це тип алгоритму глибокого навчання, який використовується для обробки даних, що мають просторові або часові зв'язки. CNN подібні до інших нейронних мереж, але вони мають додатковий рівень складності через те, що вони використовують серію **згорткових шарів**. Згорткові шари є важливим компонентом згорткових нейронних мереж (CNN). На зображенні нижче зображено типову архітектуру CNN

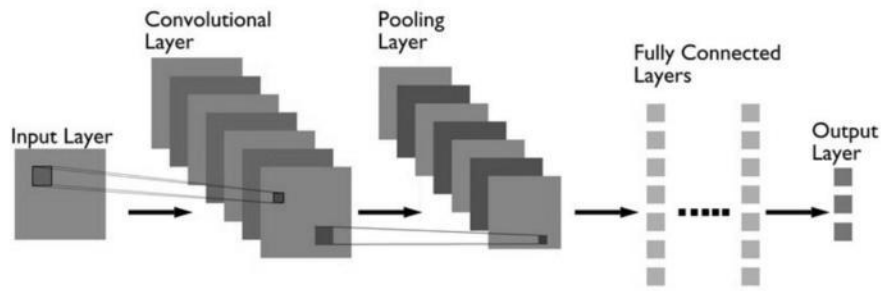


Рис. 1.3. Типова структура CNN

Нижче наведено визначення різних рівнів, показаних у наведеній вище архітектурі:

- **Згортковий шар** : згорткові шари складаються з набору фільтрів (також званих ядрами), які застосовуються до вхідного зображення. Результатом згорткового шару є карта функцій, яка є представленням вхідного зображення з застосованими фільтрами. Згорткові шари можна складати для створення більш складних моделей, які можуть вивчати більш складні функції із зображень.
- **Рівень об'єднання** : шари об'єднання – це тип згорткового шару, який використовується в глибокому навчанні. Об'єднання шарів зменшує просторовий розмір вхідних даних, полегшуючи обробку та вимагаючи менше пам'яті. Об'єднання також допомагає зменшити кількість параметрів і прискорює навчання. Існує два основних типи об'єднання: максимальне об'єднання та середнє об'єднання. Максимальне об'єднання бере максимальне значення з кожної карти функцій, а середнє об'єднання – середнє значення. Об'єднуючі шари зазвичай використовуються після згорткових шарів, щоб зменшити розмір вхідних даних перед тим, як вони подадуться на повністю пов'язаний рівень.
- **Повністю зв'язаний рівень** : повністю зв'язані шари є одним із найпростіших типів шарів у згортковій нейронній мережі (CNN). Як випливає з назви, кожен нейрон у повністю підключеному шарі є повністю підключеним до кожного іншого нейрона попереднього рівня. Повністю зв'язані рівні зазвичай використовуються наприкінці CNN, коли мета

полягає в тому, щоб взяти функції, отримані на попередніх рівнях, і використати їх для прогнозування. Наприклад, якби ми використовували CNN для класифікації зображень тварин, останній повністю пов'язаний шар міг би взяти ознаки, отримані попередніми шарами, і використовувати їх для класифікації зображення як собаки, kota, птаха тощо.

CNN часто використовують для завдань розпізнавання зображень і класифікації. Наприклад, CNN можна використати для ідентифікації об'єктів на зображенні або для класифікації зображення kota чи собаки. CNN також можна використовувати для більш складних завдань, зокрема генерування описів зображення або визначення точок інтересу на зображенні. CNN також можна використовувати для даних часових рядів, таких як аудіодані або текстові дані. CNN є потужним інструментом для глибокого навчання, і вони були використані для досягнення найсучасніших результатів у багатьох різних програмах.

Нижче наведено список різних типів архітектур CNN:

LeNet: LeNet є першою архітектурою CNN. Він був розроблений у 1998 році Янном Лекуном, Корінною Кортес і Крістофером Берджесом для проблем розпізнавання рукописних цифр. LeNet був одним із перших успішних CNN і часто вважається «Hello World» глибокого навчання. Це одна з найперших і найбільш широко використовуваних архітектур CNN, яка успішно застосовувалася для таких завдань, як розпізнавання рукописних цифр. Архітектура LeNet складається з кількох згорткових рівнів і рівнів об'єднання, за якими слідує повністю зв'язаний рівень. Модель має п'ять шарів згортки, за якими слідує два повністю з'єднані шари. LeNet був початком CNN у глибокому навчанні проблем комп'ютерного зору. Однак LeNet не міг добре тренуватися через проблему зникаючих градієнтів. Щоб вирішити цю проблему, рівень скороченого з'єднання, відомий як max-pooling, використовується між згортковими шарами, щоб зменшити просторовий розмір зображень, що допомагає запобігти переобладнанню та дозволяє CNN навчатися ефективніше. На діаграмі нижче представлена архітектура LeNet-5.

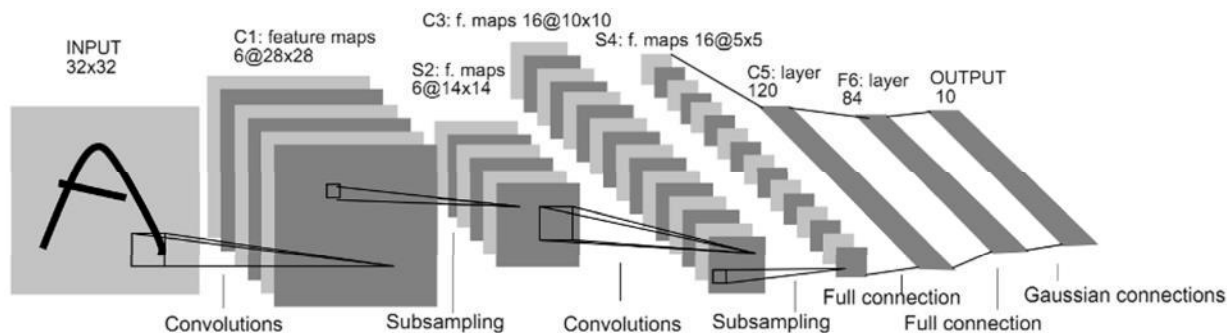


Рис. 1.4. LeNet CNN

LeNet CNN - це проста, але потужна модель, яка використовувалася для різних завдань, таких як розпізнавання рукописних цифр, розпізнавання дорожніх знаків і розпізнавання облич. Незважаючи на те, що LeNet було розроблено понад 20 років тому, його архітектура актуальна й донині та продовжує використовуватися.

AlexNet: AlexNet — це архітектура глибокого навчання, яка популяризувала CNN. Його розробили Алекс Крижевський, Ілля Суцкевер і Джефф Хінтон. Мережа AlexNet мала дуже схожу архітектуру на LeNet, але була глибшою, більшою та містила згорткові шари, складені один на одного. AlexNet був першим великомасштабним CNN і використовувався для перемоги у ImageNet Large Scale Visual Recognition Challenge (ILSVRC) у 2012 році. Архітектура AlexNet була розроблена для використання з великомасштабними наборами даних зображень і досягла найсучаснішого рівня. результати на момент публікації. AlexNet складається з 5 згорткових шарів із поєднанням шарів максимального об'єднання, 3 повністю підключених шарів і 2 шарів вилучення. Функція активації, яка використовується на всіх рівнях, Relu. Функція активації, яка використовується на вихідному рівні, — Softmax.

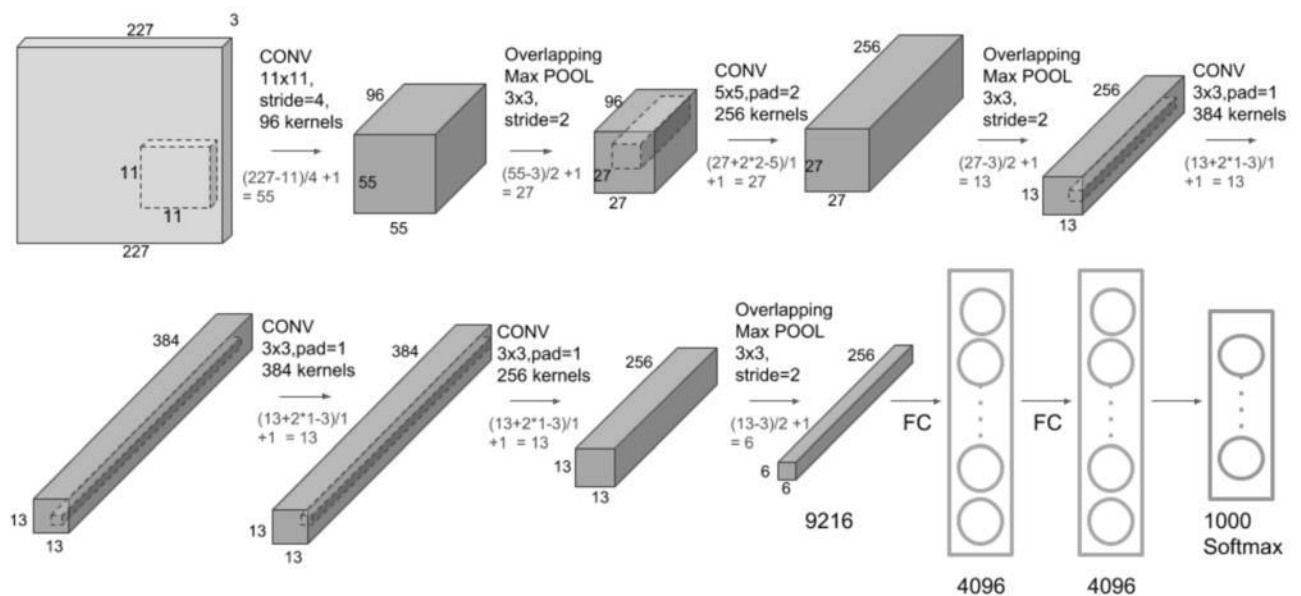


Рис. 1.5. Архітектура AlexNet

ZF Net: ZFnet — це архітектура CNN, яка використовує комбінацію повністю підключених рівнів і CNN. ZF Net був розроблений Метью Зейлером і Робом Фергусом. Це був переможець ILSVRC 2013. Мережа має відносно менше параметрів, ніж AlexNet, але все одно перевершує її в класифікаційному завданні ILSVRC 2012, досягаючи найвищої точності лише з 1000 зображеннями на клас. Це було вдосконалення AlexNet шляхом налаштування гіперпараметрів архітектури, зокрема шляхом розширення розміру середніх згорткових шарів і зменшення кроку та розміру фільтра на першому шарі. Він заснований на моделі Зейлера і Фергуса, яка була навчена на наборі даних ImageNet. Архітектура ZF Net CNN складається загалом із семи рівнів: згортковий рівень, рівень максимального об'єднання (зменшення масштабу), рівень конкатенації, згортковий рівень із функцією лінійної активації, і перший крок, видалення з метою регуляризації, застосоване перед повністю підключеним виходом. Ця модель CNN обчислювально ефективніша, ніж AlexNet, завдяки введенню етапу наближеного висновку через деконволюційні шари в середині CNN. Ось стаття на ZFNet .

GoogLeNet: GoogLeNet — це архітектура CNN, яка використовується Google для перемоги в класифікаційному завданні ILSVRC 2014. Він був розроблений

Джеффом Діном, Крістіаном Сегеді, Александро Сегеді та іншими. Було показано, що він має значно менший рівень помилок у порівнянні з попередніми переможцями AlexNet (переможець IISVRC 2012) і ZF-Net (переможець IISVRC 2013). З точки зору частоти помилок, ця помилка значно менша, ніж у VGG (друге місце у 2014 році). Він досягає глибшої архітектури, використовуючи низку різних методів, включаючи згортку 1×1 і об'єднання глобального середнього. Архітектура GoogleNet CNN є дорогою з точки зору обчислень. Щоб зменшити кількість параметрів, які потрібно вивчати, він використовує важкі шари роз'єднання поверх CNN, щоб усунути просторову надлишковість під час навчання, а також пропонує швидкі з'єднання між першими двома згортковими шарами перед додаванням нових фільтрів у пізніші шари CNN. Реальні програми/прикладні архітектури GoogLeNet CNN включають задачу розпізнавання цифр номера будинку (SVHN), яка часто використовується як проксі для виявлення придорожніх об'єктів. Нижче наведено спрощену блок-схему, що представляє архітектуру GoogLeNet CNN:

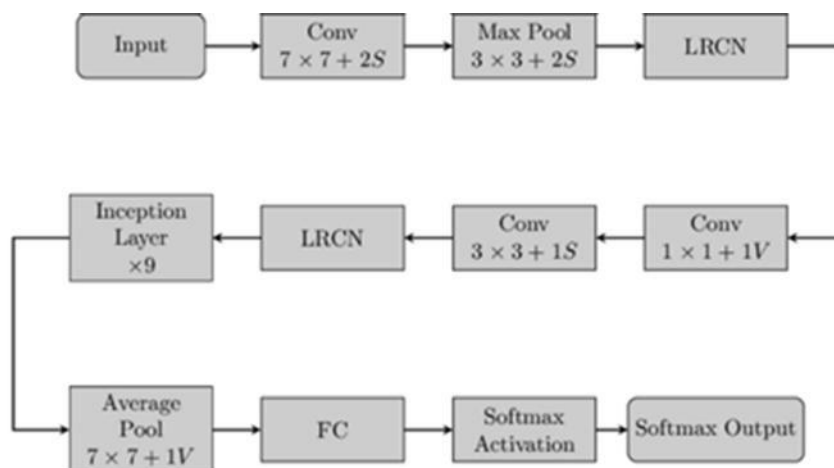


Рис. 1.6. Архітектура GoogLeNet CNN

VGGNet: VGGNet — це архітектура CNN, розроблена Кареном Симоньяном, Ендрю Зіссерманом та ін. в Оксфордському університеті. VGGNet — це 16-рівневий CNN із до 95 мільйонів параметрів, навчений на понад одному мільярді зображень (1000 класів). Він може приймати великі вхідні зображення розміром 224×224 пікселя, для яких він має 4096 згорткових функцій. CNN із такими великими фільтрами є дорогими для навчання та потребують багато

даних, що є головною причиною того, чому архітектури CNN, такі як GoogLeNet (архітектура AlexNet), працюють краще, ніж VGGNet, для більшості завдань класифікації зображень, де вхідні зображення мають розмір від 100 x 100 до пікселів і 350 x 350 пікселів. Реальні додатки/прикладі архітектури VGGNet CNN включають класифікаційне завдання ILSVRC 2014, яке також виграла архітектура GoogLeNet CNN. Модель VGG CNN є обчислювально ефективною та служить надійною основою для багатьох програм комп'ютерного зору завдяки її застосуванню для багатьох завдань, включаючи виявлення об'єктів. Його глибокі представлення функцій використовуються в багатьох архітектурах нейронних мереж, таких як YOLO, SSD тощо. Діаграма нижче представляє стандартну схему архітектури мережі VGG16:

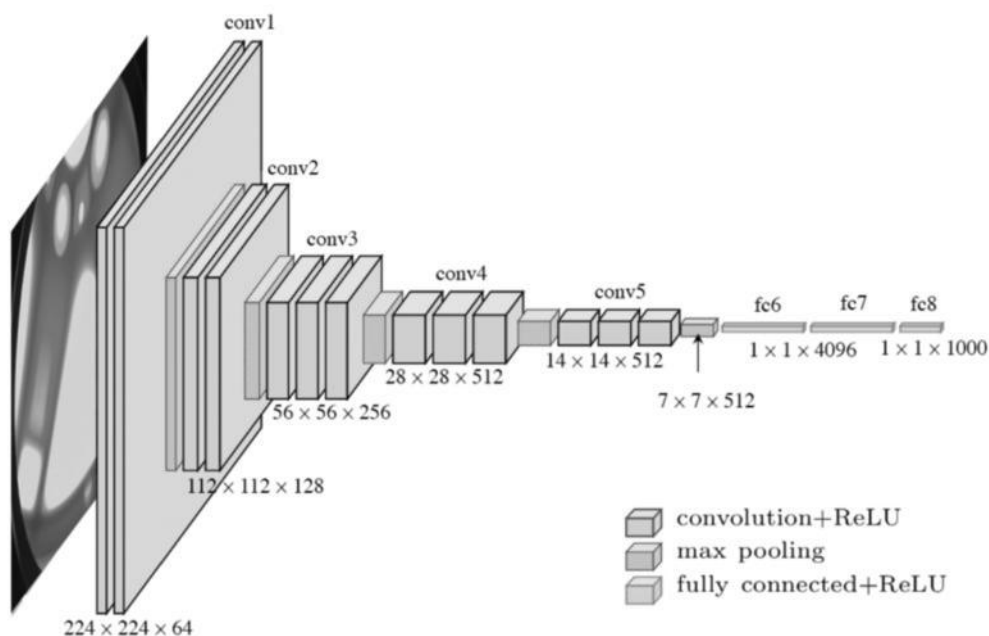


Рис. 1.7. Схема архітектури мережі VGG16

ResNet: ResNet — це архітектура CNN, розроблена Kaiming He та ін. перемогти в класифікаційному завданні ILSVRC 2015 із помилкою до п'ятірки лише 15,43%. Мережа має 152 рівні та понад один мільйон параметрів, що вважається глибоким навіть для CNN, оскільки для навчання мережі на наборі даних ILSVRC 2015 знадобилося б більше 40 днів на 32 графічних процесорах. CNN здебільшого використовуються для завдань класифікації зображень із 1000 класами, але ResNet доводить, що CNN також можна

успішно використовувати для вирішення проблем обробки природної мови, таких як завершення речень або машинне розуміння, де вони використовувалися командою Microsoft Research Asia у 2016 та 2017 роках відповідно. Реальні програми/прикладі архітектури ResNet CNN включають систему машинного розуміння Microsoft, яка використовувала CNN для отримання відповідей на понад 100 тисяч запитань у більш ніж 20 категоріях. Архітектура CNN ResNet ефективна з точки зору обчислень і може бути збільшена або зменшена відповідно до обчислювальної потужності графічних процесорів.

MobileNets: MobileNets — це CNN, які можна встановити на мобільному пристрої для класифікації зображень або виявлення об'єктів із низькою затримкою. MobileNets були розроблені Andrew G Trillion та іншими. Зазвичай це дуже маленькі архітектури CNN, що полегшує їх роботу в режимі реального часу за допомогою вбудованих пристроїв, таких як смартфони та дрони. Архітектура також є гнучкою, тому її перевірили на CNN із 100-300 шарами, і вона все ще працює краще, ніж інші архітектури, такі як VGGNet. Реальні приклади архітектури MobileNets CNN включають CNN, вбудовані в телефони Android для запуску Google Mobile Vision API, який може автоматично ідентифікувати мітки популярних об'єктів на зображеннях.

GoogLeNet_DeepDream: GoogLeNet_DeepDream — це архітектура глибокої мрії CNN, розроблена Олександром Мордвінцевим, Крістофером Олахом та іншими. Вона використовує мережу Inception для створення зображень на основі функцій CNN. Архітектура часто використовується з набором даних ImageNet для генерації психоделічних зображень або створення абстрактних творів мистецтва за допомогою людської уяви на семінарі ICLR 2017 Девіда Ха та ін.

2. ПРОЕКТНА ЧАСТИНА

2.1. Оптико-цифрові методи контролю стану поверхні смуг металопрокату

Оптико-цифрові методи контролю стану поверхні смуг металопрокату забезпечують швидке виявлення та локалізацію дефектів. Камери, які зараз застосовують забезпечують одержання кісних зображень смуг металопрокату, які є вхідної інформацією для оцінювання дефектності. Крім того встановлено та систематизовано технологічні параметри прокатування, узагальнено їх вплив на якість готової продукції. Встановлено причини виникнення поверхневих дефектів, описано їх класи, визначені їх ознаки (геометричні та оптико-цифрові), метрики вимірювання. Запропоновано низку параметрів їх кількісного аналізу. Разом з тим, швидкості прокатування смуг металу постійно зростають, що спричиняє зміни у технології, а відповідно й ставить нові задачі щодо врахування впливу технологічних факторів на коректність оцінювання параметрів дефектів з використанням глибоких нейронних мереж. Дослідження впливу вібродинамічних процесів у вузлах металургійного обладнання, за високих швидкостей прокатування є окремою науковим і практичним завданням. Підвищення вібрації до небезпечного рівня обмежує продуктивність металопрокатного обладнання, знижує якість готових смуг по поздовжньої різнотовщинності та формує періодичні дефекти поверхні типу «ребристість», «поперечні тіньові смуги», а також підвищує зношення валків. Наукові дослідження проблеми вібрації проводять у наступних напрямках:

- створення систем контролю вібрації та діагностики технічного стану станів;
- розробка способів та пристроїв демпфування коливань.

Врахування впливу динамічних процесів, що виникають при прокатуванні на точність розпізнавання дефектів смуги, є дуже складним завданням, вирішувати яке необхідно на міждисциплінарному рівні з інтеграцією знань у галузі теорії прокатування та підходів штучного інтелекту, тому в технічній літературі висвітлюються лише окремі аспекти цієї проблеми. На нашу думку,

важливо оцінити вплив частоти вібрацій на якість розпізнавання дефектів смуг металопрокату за різних умов освітлення. Це б дозволило сформулювати умови оптимального виявлення дефектів різної геометрії.

Традиційні методи розпізнавання зображень вимагають умов високої освітленості та мають погану адаптивність до недостатньої освітленості. Важко задовольнити вимоги розпізнавання зі зібраними зображеннями поверхневих дефектів за різних умов. Крім того, традиційні методи виявлення зображення в основному отримують функції зображення завдяки інтенсивній попередній обробці зображення. Іноді дефекти поверхні можуть бути дуже малі, що призводить до нечітко помітних елементів. Традиційними методами важко визначити точність визначити дефекти поверхні металопрокату

У даній роботі розроблено методику виявлення мікродефектів металевих поверхонь на глибоких згорткових нейронних мережах і побудована оптична платформа для отримання гвинтових зображень. Збираються зображення дефектних і бездефектних поверхонь, які використовуються для навчання проєктованих глибока згорткові нейронна мережа (CNN). Щоб забезпечити ефективне виявлення, спочатку знаходять дефектні поверхні на знімках, зроблених камерами, щоб можна було витягти зображення поверхонь, які потім вводяться в детектор дефектів на основі CNN. Пропонований спосіб не потребує придбання особливості зображень поверхні заздалегідь і стійкий до змін освітлення. Порівняння з традиційними техніками машинного зору, наприклад, технікою на основі зіставлення шаблонів, продемонструвати перевага запропонованого глибокого на основі CNN.

Відомі системи контролю, що складається зі звичайного мікроелектричного пластинчастого датчика з оптико-цифрової камери була розроблена Raytrix GmbH і модернізована aixACCT Systems GmbH (рис. 2.1). Ці камери фіксують інформацію про походження світлового променя в 3D-просторі, за допомогою сукупності мікролінз, встановлених близько перед звичайним фоточутливим чіпом.

Після завершення етапу калібрування камери об'єкт зображення обчислюється реконструйовано, внаслідок чого формується повністю сфокусоване 3D-зображення карти глибини. Ці камери компактні та мають збільшену глибину фокусування порівняно з камерами мікроскопів. Подібні оптичні пристрої, як і мікролінзи, мають різні фокусні відстані.

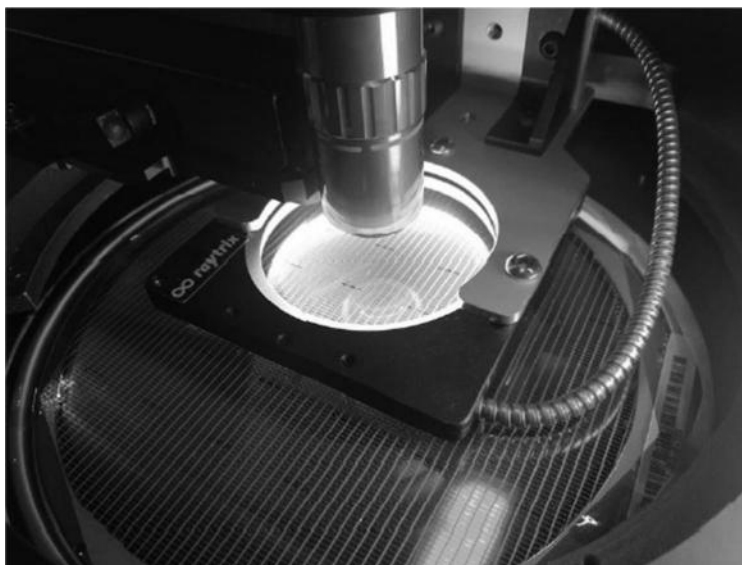


Рис. 2.1. Збільшений вигляд розробленої плівко-оптичної камери [1]

Зображення були зібрані за допомогою звичайного мікроелектричного пластинчастого зонда, модернізованого aixACCT Systems GmbH з пленоптичною камерою, розробленою Raytrix GmbH. Пленоптична камера зафіксувала інформацію про походження світлових променів у 3D-просторі за допомогою системи мікролінз, розташованих близько перед звичайним фоточутливим чіпом.

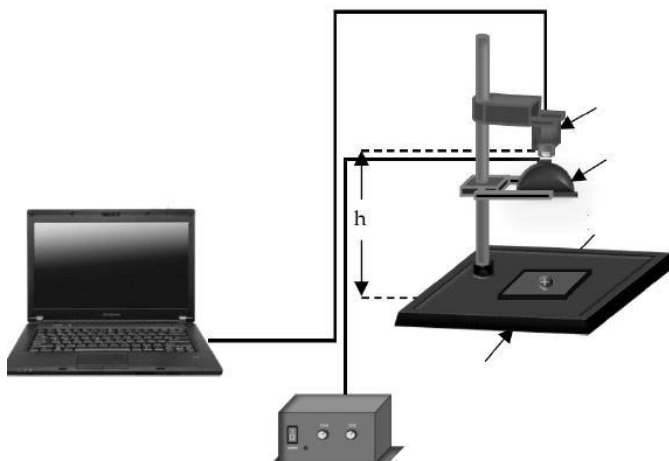
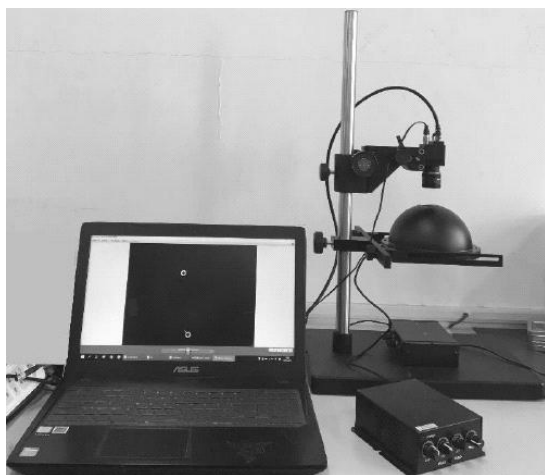


Рис. 2.2. Експериментальна система для отримання зображень. (а) Система виявлення, (б) Схема структури системи

На рисунку 2.1 показана система оптичної експериментальної платформи для отримання зображень. Було використано промислову камеру RS-A2300-GC50 (виробництва China - Microview, Пекін, Китай) з CMOS роздільною здатністю 1600x1200 пікселів і об'єктивом 16 мм (M0814-MP2). Відстань між камерами і об'єктом h становить близько 200 мм. Контролер джерела світла контролює яскравість рефлектора джерела світла. Великий кут від джерела світла DOME допомагає знімати зображення нерівної поверхні, і багаторазове відображення через внутрішню стінку півсфери може повністю усунути тіні, що корисно для перевірки металевих або дзеркальних поверхонь. Система захоплює зображення є 24-бітним зображенням розміром 1600x1200 пікселів у форматі BMP. Довжина пікселя на зображенні становить приблизно 0,0765 мм (тобто $122,35 \text{ мм} / 1600 \text{ пікселів} = 0,0765 \text{ мм/піксель}$).

Якщо необхідна роздільна здатність сенсора не відповідає стандартній роздільній здатності сенсора, виберіть камеру, роздільна здатність сенсора якої більша, ніж вам потрібна, або використовуйте кілька камер.

2.2. Технічне оснащення фотозйомки та його параметри

Визначаючи необхідну роздільну здатність датчика, звужується кількість параметрів камери, які відповідають вашим потребам. Іншим важливим фактором, який впливає на вибір камери, є фізичний розмір датчика, відомий як розмір датчика. На малюнку 3 показано розміри датчиків для стандартних датчиків 1/3 дюйма, 1/2 дюйма та 2/3 дюйма. Зауважте, що назви датчиків не відображають фактичних розмірів датчика.

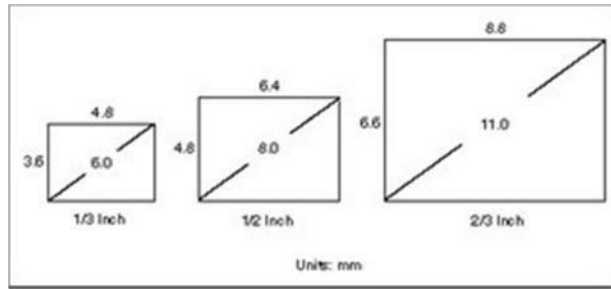


Рис. 2.3. Типові розміри датчиків та їх фактичні розміри

У більшості випадків розмір датчика є фіксованим для заданої роздільної здатності датчика. Якщо ви знайдете камери з однаковою роздільною здатністю, але з різними розмірами датчика, ви можете визначити потрібний вам розмір датчика на основі наступної вказівки.

Об'єктив в першу чергу визначається його фокусною відстанню. На рисунку 2.4 показано зв'язок між фокусною відстанню об'єктива, полем зору, розміром сенсора та робочою відстанню.

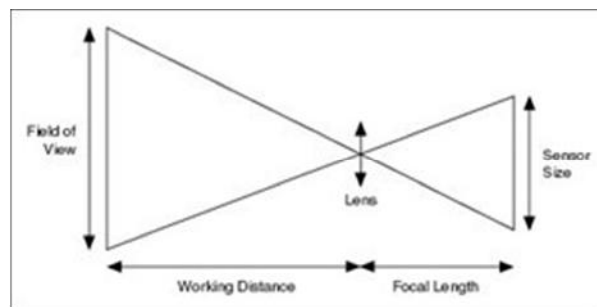


Рис. 2.4. Зв'язок між фокусною відстанню, полем зору, розміром датчика та робочою відстанню

Робоча відстань - це відстань від передньої частини об'єктива до об'єкта, що перевіряється. Якщо ви знаєте поле зору, розмір датчика та робочу відстань, ви можете обчислити фокусну відстань потрібного об'єктива за такою формулою:

$$\text{фокусна відстань} = \text{розмір датчика} \times \text{робоча відстань} / \text{поле зору}$$

Об'єктиви виготовляються з обмеженою кількістю стандартних фокусних відстаней. Загальні фокусні відстані об'єктивів включають 6 мм, 8 мм, 12,5 мм,

25 мм і 50 мм. Вибравши об'єктив, фокусна відстань якого найближче до фокусної відстані, оптимальної для системи обробки зображень, потрібно лише налаштувати робочу відстань, щоб об'єкт аналізу перебував у фокусі.

Лінзи з короткою фокусною відстанню (менше 12 мм) створюють зображення зі значним спотворенням. Якщо ваша програма чутлива до спотворення зображення, спробуйте збільшити робочу відстань і використовувати об'єктив із більшою фокусною відстанню. Якщо ви не можете змінити робочу відстань, ви дещо обмежені у виборі об'єктива.

Під час налаштування системи потрібно налаштували різні параметри, що входять до рівняння визначення фокусної відстані, поки не було досягнуто правильної комбінації значень, що відповідатимуть вашим умовам перевірки та вимогам щодо вартості.

Контраст

Роздільна здатність і контраст є тісно пов'язаними між собою факторами, що впливають на якість зображення. Контраст визначає різницю значень інтенсивності між досліджуваним об'єктом і фоном. Система обробки зображень повинна мати достатній контраст, щоб відрізнити об'єкти від фону. Правильні методи освітлення можуть підвищити контраст вашої системи.

Глибина різкості

Глибина різкості об'єктива - це його здатність утримувати у фокусі об'єкти різної висоти або розташовані на різній відстані від камери. Якщо потрібно аналізувати об'єкти різної висоти, вибирають об'єктив, що може підтримувати задану якість зображення, за умов наближення об'єкту до об'єктива та віддалення від нього. Можна збільшувати глибину різкості, закривши діафрагму об'єктива та забезпечивши більш потужне освітлення.

Телецентричні лінзи працюють із широкою глибиною різкості. За допомогою телецентричного об'єктива можна знімати об'єкти на різних відстанях від об'єктива, і об'єкти залишатимуться у фокусі.

Перспектива

Помилки перспективи виникають, коли вісь камери не перпендикулярна до аналізованого об'єкту. На рис. 5а показано ідеальне положення камери. На рис. 5 зображено камеру, яка знімає об'єкт під кутом.

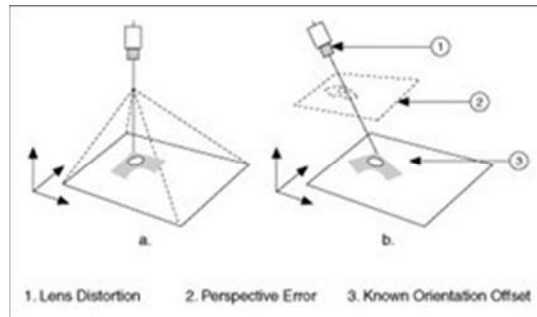


Рис. 2.5. Кут камери відносно аналізованого об'єкта

На рис. 2.6а показано сітку точок, рис. 2.6 ілюструє помилки перспективи, спричинені камерою, яка відображає сітку під кутом.

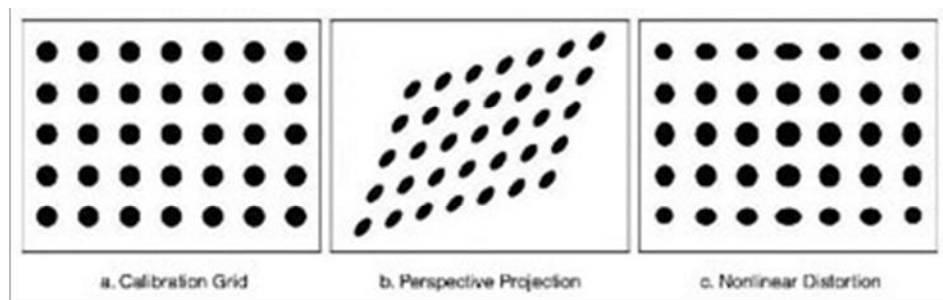


Рис. 2.6. Помилки перспективи та спотворення

Слід розташувати камеру перпендикулярно досліджуваному об'єкту, щоб зменшити помилки перспективи. Обмеження інтеграції можуть перешкодити встановити камеру перпендикулярно сцені. За цих обмежень можна проводити точні вимірювання, виправляючи помилки перспективи за допомогою методів просторового калібрування.

Спотворення

Нелінійне спотворення - це геометрична аберація, спричинена оптичними помилками в об'єктиві камери. Звичайний об'єктив камери створює радіальне спотворення. Це призводить до того, що точки, віддалені від оптичного центру лінзи, виглядають далі від центру, ніж вони є насправді. Рисунок 4с ілюструє

вплив викривлення на мережу точок. За умови виникнення спотворення, інформація на зображенні зміщується відносно центру поля зору, але інформація не обов'язково втрачається. Є можливість виправлення зображення за допомогою просторового калібрування.

Освітлення

Одним із найважливіших аспектів налаштування середовища для фотографування є належне освітлення. Зображення, отримані за належних умов освітлення, полегшують розробку програмного забезпечення для обробки зображень і зменшують загальний час обробки. Однією з цілей освітлення є відокремлення об'єкта або аналізованої частини, від навколишнього фону якомога більшою кількістю рівнів сірого. Ще однією метою є контроль світла в сцені. Освітлювальні пристрої налаштовували таким чином, щоб зміни навколишнього освітлення (наприклад, зміна сонячного світла залежно від погоди чи часу доби) не впливали на результати аналізу й обробку зображень. Поширені типи джерел світла є: галогенні, світлодіодні, люмінесцентні та лазерні.

Тип обраного освітлювального приладу впливає на якість дефектоскопії та дефектометрії. Неправильне освітлення може спричинити появу тіней і відблисків, які погіршують ефективність обробки зображень. Наприклад, деякі об'єкти відбивають велику кількість світла через свою кривизну або текстуру поверхні. Сильно спрямовані джерела світла збільшують чутливість дзеркальних відблисків (відблисків), як показано на штрих-коді на рис. 2.7а. На рис. 2.7б показано зображення того самого штрих-коду, отримане за розсіяного освітлення.



Рис. 2.7. Використання розсіяного освітлення для усунення відблисків

Підсвічування - ще один метод освітлення, який може допомогти покращити продуктивність вашої системи технічного зору. Якщо діагностична

проблема полягає у аналізі лише форми об'єкта, можна створити силует об'єкта, розмістивши джерело світла за об'єктом, який діагностують. Освітлюючи об'єкт ззаду, створюють різкий контраст, що робить пошук країв і вимірювання відстаней швидким і легким. На рис. 2.8 показано штаповану металеву деталь, отриману в установці з використанням підсвічування.

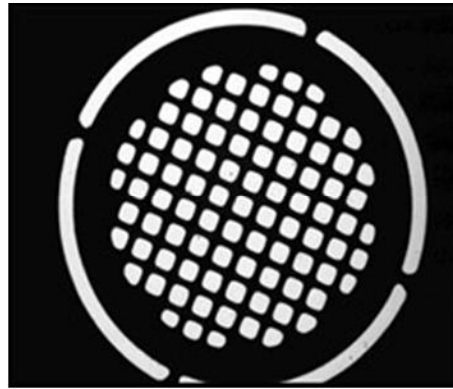


Рис. 2.8. Використання фонового освітлення для створення силуету об'єкта

Багато інших факторів, наприклад тип обраної камери, впливають на технічні рішення щодо відповідного освітлення. Наприклад, можна вибрати джерела освітлення та фільтри, довжина хвилі яких відповідає чутливості ПЗЗ-сенсора обраної камери та кольору об'єкта, що перевіряють. Крім того, можна використати спеціальні світлофільтри або лінзи.

2.3. Отримання зображень для дослідження

Для дослідження впливу вібрацій та освітлення на результат розпізнавання дефектів металеву пластину з дефектами типу "подряпина" фотографували у лабораторних умовах. При цьому змінювали освітленість, а пластину циклічно переміщували у вертикальному напрямку з частотою 10-30 Гц та амплітудою переміщень 0,1-03, мм. Циклічне переміщення пластини виконували на сервогидравлічній установці СТМ-100. Фіксували значення вертикального переміщення штока. Схема дослідної установки для отримання зображень металеві поверхні з пошкодженнями типу "подряпина" зображена на рис. 1. Освітлення поверхні аналізованої металеві смуги, розділеної на ділянки, виконано зверху. Зображення одержується внаслідок того, що різні ділянки

металевої смуги неоднаково відхиляють падаюче на них світло, а відбиті промені мають різну інтенсивність [18]. За пропонованої схеми освітлення камера одержує більшу частину відбитого світла, це так званий аналіз "у світлому полі". Напрявлене освітлення є майже перпендикулярним до досліджуваної металевої поверхні. За таких умов зйомки поверхня буде яскраво сірою, а дефекти – добре видимими.

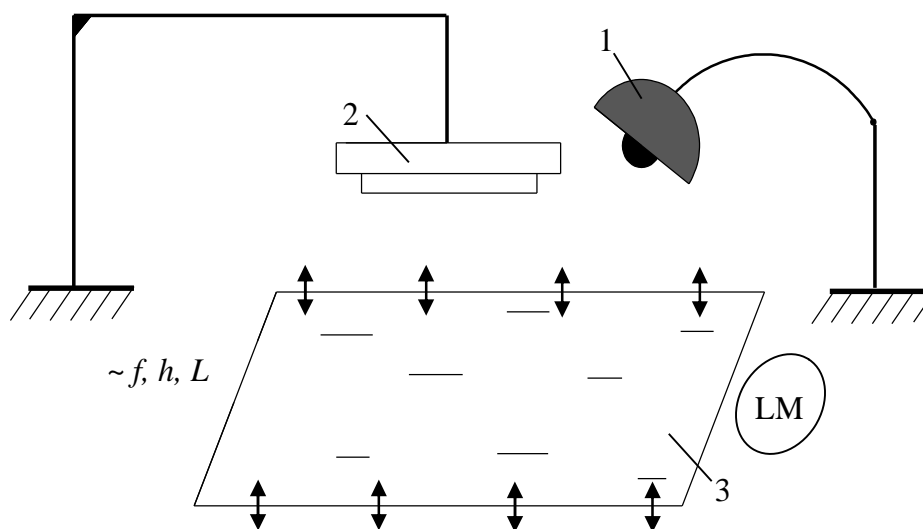


Рис. 2.9. Схема одержання зображень ділянок металевої смуги під час циклічного вертикального переміщення: 1 – лампа; 2 – фотокамера; 3 – металева смуга з пошкодженнями; LM – люксметр

Для одержання зображень цих зон використовували камеру Nikon D5600. Для освітлення металевої пластини використовували лампу розжарювання потужністю 60 Вт. Під час одержання зображень яскравість світла змінювали регулятором та контролювали люксметром, який залишався нерухомим. Фотозображення всіх ділянок були одержані у фіксованій точці зйомки, так як показано на рис. 2.9. Зображення отримували за параметрів поданих у таблиці 1. На рис. 2.10 приведено приклади отриманих зображень при різних рівнях освітленості.

Таблиця 2.1. Параметри випробувань та умови одержання зображень

Номер зразка	Освітленість, люкс	Частота випробувань, Гц	Амплітуда переміщень, мм
--------------	--------------------	-------------------------	--------------------------

0	300	0	0
1		10	0,1
2		10	0,2
3		10	0,3
4		20	0,1
5		20	0,2
6		20	0,3
7		30	0,1
8		30	0,2
9		30	0,3
0	400	0	0
10		10	0,1
11		10	0,2
12		10	0,3
13		20	0,1
14		20	0,2
15		20	0,3
16		30	0,1
17		30	0,2
18		30	0,3
0	500	0	0
19		10	0,1
20		10	0,2
21		10	0,3
22		20	0,1
23		20	0,2
24		20	0,3
25		30	0,1
26		30	0,2

27		30	0,3
----	--	----	-----

Розглядатиму все ж таки питання «кількості дефектів» (як кількість рядків таблиці з даними – підсумок роботи нейромережі). Під час отримання зображення яскравість світла регулювалася та контролювалася люксометром, який залишався нерухомим. Фотографії всіх областей були отримані у фіксованій точці зйомки, як показано на рис. 2.10 для зони II. Це зробило умови зйомки ідентичними та дозволило порівнювати зображення як в одній зоні, так і між різними областями. Зображення були отримані при рівнях освітленості 2, 50, 100, 200, 300, 400, 500, 600, 700 і 800 лк. На рисунку 2 показані приклади зображень, отриманих при різних рівнях освітлення.

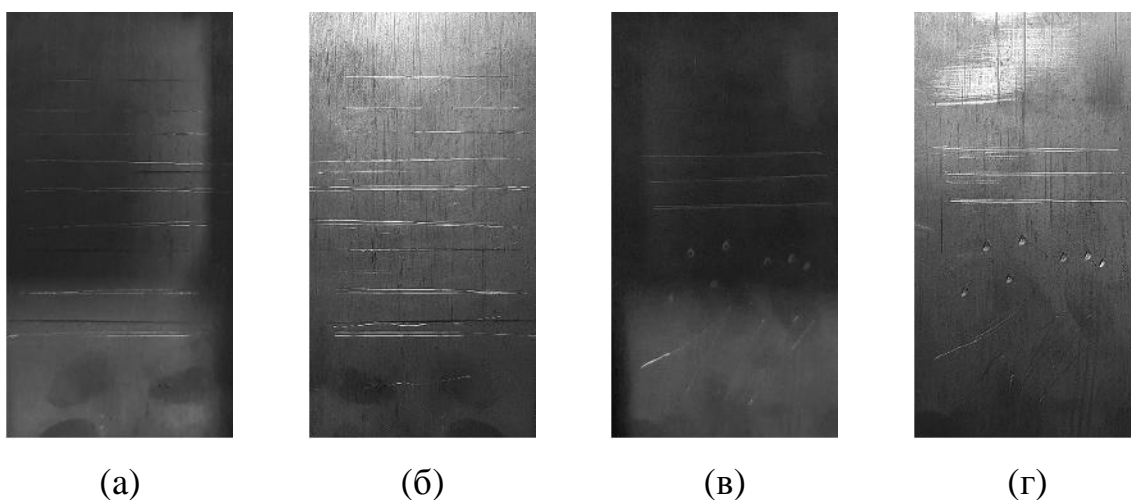


Рис. 2.10. Зображення пошкоджень у зонах I (а, б) та II (в, г) при двох рівнях освітлення: 2 лк (а, в) та 600 лк (б, г)

2.4. Розпізнавання зображень за допомогою CNN

Для розпізнавання пошкоджень використовувалася модель нейронної мережі на основі архітектури U-net з декодером на базі ResNet152 [11]. Модель була навчена на зображеннях поверхні прокату, наданих сталеливарною компанією у рамках Конкурсу з аналітики та моделювання, організованого на платформі Kaggle у 2019 році. З масиву зображень ми вибрали ті, які відповідають проаналізованим дефектам, схожим на подряпини.

Нейронна мережа U-net була розроблена в 2015 році для обробки медичних зображень. Її архітектура є результатом розробки та вдосконалення звичайних згорткових нейронних мереж. Що стосується завдань дефектометрії, то потрібно розділити зображення на дефектні та бездефектні. Локалізація пошкодження в цьому випадку є більш точною. Окрім покращення контролю якості продукції, це дозволить розраховувати кількісні характеристики пошкодження (площа, напрямок тощо). Це, у свою чергу, дасть нам змогу краще зрозуміти природу пошкодження та дозволить розробити кроки для його усунення. Нейронна мережа U-net забезпечує семантичну сегментацію зображення, в якій кожен піксель класифікується як належний до пошкодженої або непошкодженої області. У цьому випадку вхідне і вихідне зображення мають однаковий розмір.

Архітектура U-net складається з кодера, який виявляє карти ознак, притаманні пошкодженню, що представляє інтерес, і декодера, який проектує ті області, які найкраще відповідають виявленим функціям оригінальному зображенню (рис. 2.11). Кодер містить 5 ступенів. Кожен етап містить послідовно з'єднані згорткові шари, за допомогою яких виконується пакетна нормалізація. Декодер також містить п'ять каскадів, кожен з яких підключений до виходів відповідних каскадів кодера і виконує підвищення дискретизації та узагальнення своїх карт ознак. Вихідний шар нейронної мережі має форму 256×256 нейронів із сигмовидною функцією активації. Це дає змогу отримати значення на нейронах вихідного шару в діапазоні $\left[0,1\right]$, що відображає ступінь надійності, з якою модель розпізнає відповідний піксель як частину пошкодження.

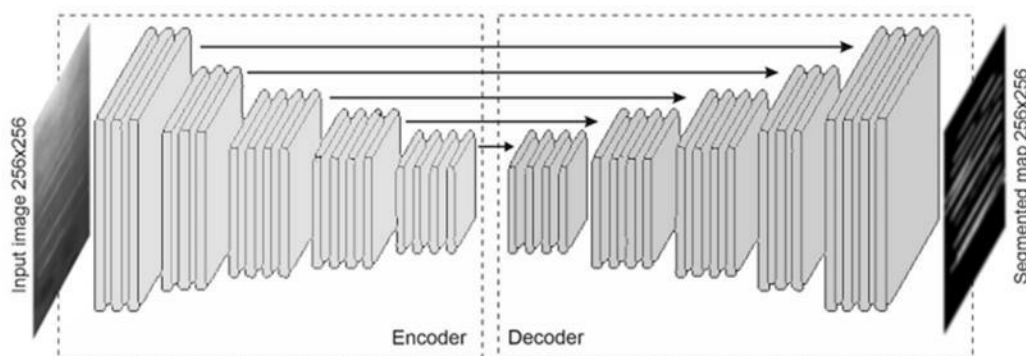


Рис. 2.11. Загальна архітектура моделі сегментації на основі U-net

Використано розроблену та навчену в ході попередніх досліджень модель семантичної сегментації [10, 11]. При розпізнаванні зображень з тестової вибірки модель показала коефіцієнт Дайса 0,93. Коефіцієнт Dice показує площу перекриття відносно фактичної площі розпізнаного об'єкта у відсотках. На вхід нейронної мережі надходило зображення розміром 256×256 пікселів. Нейронну мережу було навчено за допомогою оптимізатора SGD (Stochastic Gradient Descent) і функції втрат бінарної крос-ентропії.

Модель розпізнавання пошкоджень реалізована на Python 3.8 за допомогою бібліотек TensorFlow і Keras версії 2.6.0. Для роботи з зображеннями використовувалася бібліотека OpenCV, розроблена компанією Intel. Він містить зручні інструменти для обробки та аналізу вмісту зображень.

4 показана загальна послідовність операцій, що використовуються при розпізнаванні дефектів нейронною мережею. На вхід нейронної мережі можуть надходити лише зображення розміром 256×256 пікселів. Тим не менш, зображення металевої поверхні можуть бути будь-якого розміру. Тому, щоб розпізнати всю поверхню, фрагменти були вирізані з вхідного зображення так, щоб вони покривали всю його площу. Для цього використовувалося ковзне вікно розміром 256×256 пікселів, яке переміщувалося з певним фіксованим кроком (рівним $2/3$ ширини вікна) спочатку по горизонтальній осі, а після досягнення краю зображення – по вертикальній осі. Оскільки розпізнавання зображення часто призводить до краєвих ефектів, які спричиняють низьку точність розпізнавання біля країв, фрагменти були вирізані так, щоб вони накладалися. Таким чином, для вхідного зображення був побудований тензор з фрагментами, який потім подавався на вхід нейронної мережі для розпізнавання.

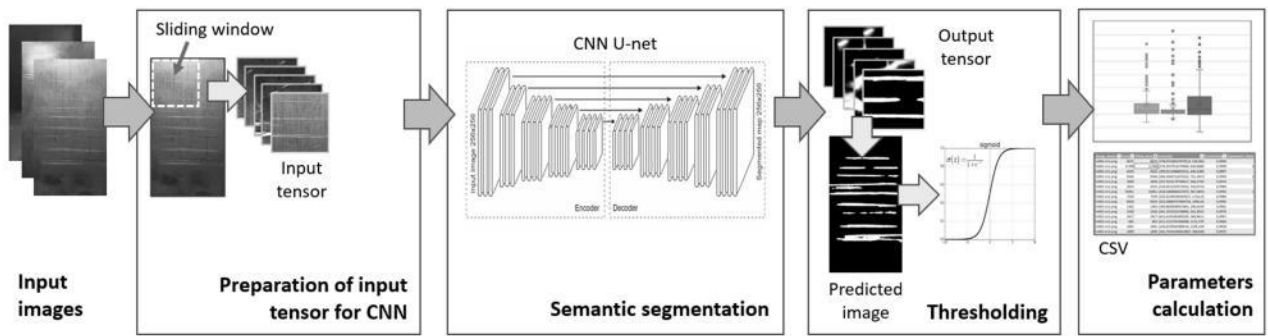


Рис. 2.12. Загальна послідовність операцій при розпізнаванні дефектів за допомогою нейронної мережі на зображенні довільної форми

Після розпізнавання набору фрагментів у вхідному зображенні за допомогою моделі нейронної мережі на його виході виходить набір фрагментів, у якому кожен розпізнаний фрагмент відповідає вхідному фрагменту. Щоб отримати картину пошкодження оригінального зображення, його необхідно синтезувати з фрагментів, розпізнаних шляхом згортання (і накладання) у правильному порядку. Пікселі фрагментів зображення, що накладаються, матимуть різні значення інтенсивності. Тому було використано та досліджено дві стратегії, які передбачають синтез загального образу з пошкодженням:

- усереднення. При накладенні двох фрагментів інтенсивність пікселя розраховується як середнє арифметичне. Це узгоджується з гіпотезою про те, що обидва фрагменти вважаються однаково точними в розпізнаванні. Тоді вихідне значення буде усереднено з двох отриманих;

- вибір макс. У цьому випадку для пікселя в області накладання фрагментів максимальне значення інтенсивності вибирається із значень, отриманих для двох відповідних фрагментів пікселів. Це узгоджується з гіпотезою про те, що якщо одна з моделей більш впевнена в пошкодженні, то вона «правильна».

Експериментально було встановлено, що максимальний метод працює краще для проаналізованих зображень. Він усуває краєві ефекти, які в основному проявляються таким чином, що деякі дефекти з країв фрагмента майже невідомі. Тому всі подальші дослідження проводились за допомогою max-методу, який застосовувався для синтезу розпізнаних зображень. Приклад вихідного зображення пошкодженої поверхні та остаточного синтезованого

зображення наведено на рис. 2.13, а,б. Як видно, різні розпізнані ділянки пошкодження мають неоднакову яскравість. Крім того, більш яскраві області відповідають пошкодженням, які класифікуються нейронною мережею як належні більше до набору ознак, які вона сформувала. Тому з деяким наближенням такі ділянки можна з більшою достовірністю вважати належними до пошкоджень.

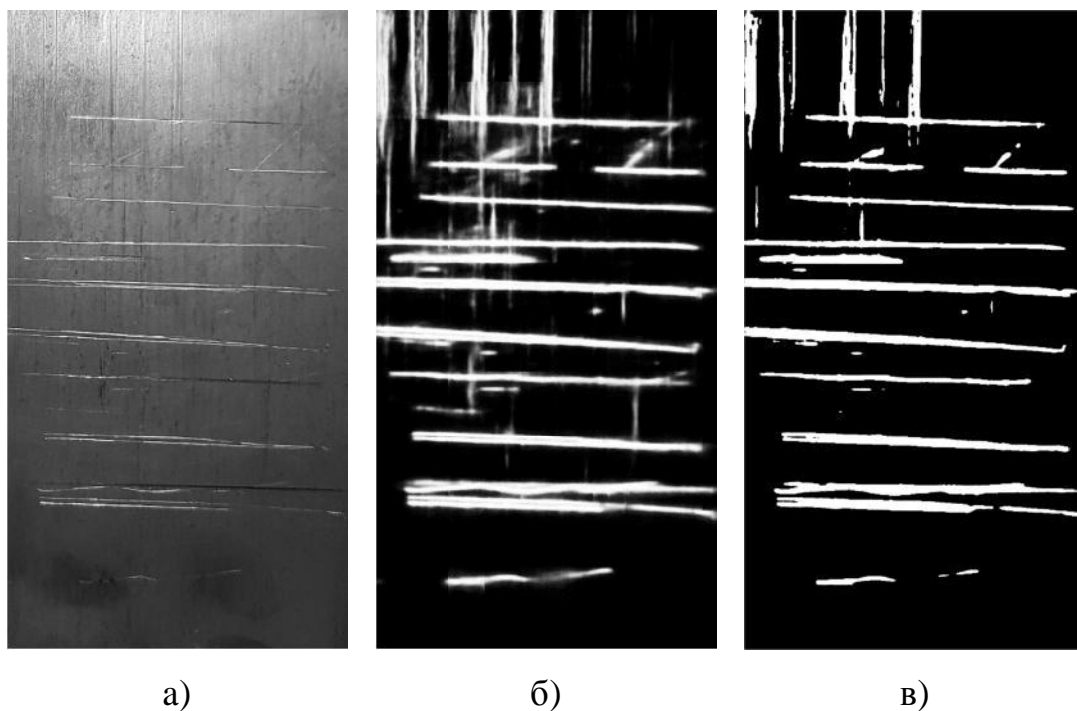


Рис. 2.13. Початкове (а), розпізнане та синтезоване (б) і бінарне (в) зображення поверхні при 50 лк

Для об'єктивної кількісної оцінки пошкодженої поверхні важливо отримати числові показники, що характеризують тип наявних пошкоджень. Окремі об'єкти (об'єднані групи пікселів), розпізнані моделлю, являють собою фрагменти поверхневих дефектів. З метою ідентифікації та диференціації типів дефектів визначено їх геометричні параметри, які описують особливості кожного виявленого фрагмента пошкодження. В першу чергу була визначена площа як основний параметр, що характеризує розмір пошкодження. Для розрахунку інших параметрів використовувався підхід «еквівалентного» еліпса (рис. 2.14), запропонований авторами [10, 22]. Еквівалентний еліпс має другий момент, рівний моменту розпізнаного дефекту. Цей підхід є спрощеним і не

завжди описує морфологію випадкових дефектів. Однак він ефективний для дефектів розглянутих класів.

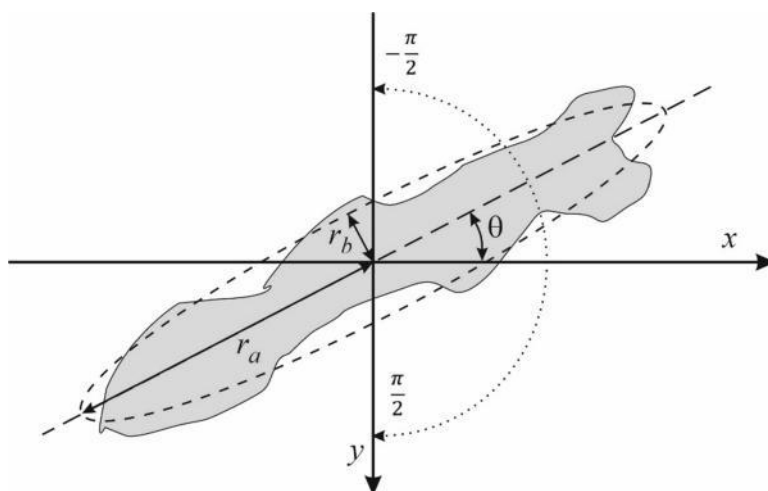


Рис. 2.14. Дефект, еквівалентний еліпс і його параметри

Ми припускаємо, що $(x_i, y_i) \in F$ - набір пікселів, які складають об'єкт. Площа пошкодження була розрахована як загальна кількість пікселів у наборі F :

$$a = \text{len}(F). \quad (2.1)$$

Координати центру мас для окремого пошкодження розраховували за формулами:

$$x_c = \frac{\sum_F x_i}{\text{len}(F)}, \quad y_c = \frac{\sum_F y_i}{\text{len}(F)} \quad (2.2)$$

Еквівалентний діаметр розраховували як діаметр кола, площа якого дорівнює площі пошкодження:

$$d_{eq} = 2 \sqrt{\frac{a}{\pi}}. \quad (2.3)$$

Головна $2r_a$ вісь еквівалентного еліпса (рис. б) проходить через центр мас об'єкта (x_c, y_c) і має нахил θ відносно осей. Орієнтація θ може змінюватися в межах діапазону $[-\frac{\pi}{2}; \frac{\pi}{2}]$ і має тенденцію до збільшення за годинниковою стрілкою, оскільки вісь у спрямована вниз.

2.5. Дослідження впливу освітлення на результат розпізнавання

При отриманні зображення поверхні важливим фактором, який впливає на результат розпізнавання дефектів, є освітлення поверхні. Тому важливо розуміти, як кількісні параметри, що використовуються для оцінки пошкодження поверхні, можуть змінюватися при зміні освітлення. На рис. 7 наведено вихідні зображення поверхні І зони (рис. 7, а) та ділянки пошкодження, розпізнаного нейронною мережею (рис. 7, б).

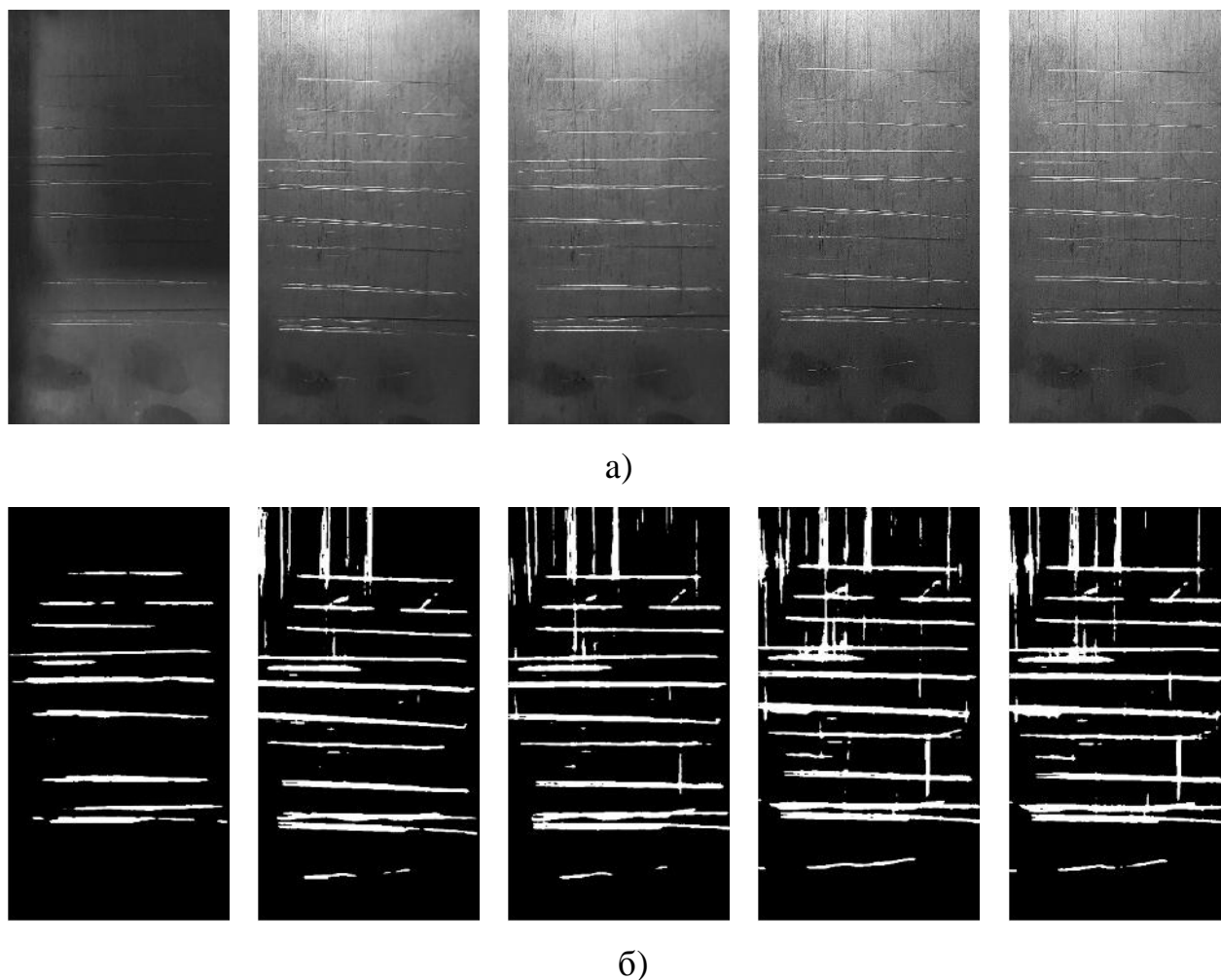


Рис. 2.15. Дефекти, їх зображення та їх бінаризовані зображення

В цілому зі збільшенням освітленості поверхні збільшується площа розрізів, визнаних пошкодженням. Також зростає кількість знайдених фрагментів пошкоджень (рис. 2.16). В першу чергу це пов'язано з дрібними подряпинами та елементами текстури поверхні, які стають краще помітними. У нашому випадку особливо помітний перехід від дуже низького рівня освітлення

(2 лк) до більш високого (50 лк). У першому випадку вертикальні подряпини на зображенні взагалі не видно, тоді як при 50 лк більшість із них уже розпізнаються. Подальше збільшення освітленості до 800 лк практично не сприяє виявленню нових подряпин (крім дрібних), але при візуальному спостереженні площа вже розпізнаних об'єктів збільшується.

Однак чутливість нейронної мережі, яка використовується для розпізнавання пошкоджень, особливо помітна навіть при дуже слабкому освітленні: усі найбільш виражені подряпини добре розпізнаються. Тому скрізь, де необхідно виявити основне, найбільш виражене пошкодження, може бути доречним слабке освітлення. У той же час рівень шуму на зображенні, спричинений підвищеним рівнем деталізації поверхневих утворень, які можуть не являти собою пошкодження, значно нижчий [21, 22].

На рис. 2.15 наведено кількісні характеристики впливу рівня освітленості на рівень розпізнавання. Розподіл за площею окремих розпізнаних фрагментів показано коробчатими діаграмами (рис. 2.15, а). Основні прямокутники коробчастої діаграми містять 50% площі фрагментів, розпізнаних у межах інтерквартильного діапазону $IQR = a_{75} - a_{25}$ (де a_{75}, a_{25} це 75-та та 25-та перцентилі, відповідно). Аналіз наведених коробкових діаграм показав, що більшість фрагментів мають площу до 10 тис. пікселів, тоді як області розміром понад 30 тис. пікселів мають лише поодинокі фрагменти. Лінії позначають значення в інтерквартильному діапазоні $\pm 1,5$, причому довжина верхнього вуса обмежена максимальним значенням із зразка, що потрапляє в цей діапазон; довжина нижнього вуса обмежена мінімальним значенням із зразка, що потрапив у його довжину.

На рис. 2.15,б наведено залежність сумарної площі розпізнаних дефектів від рівня освітленості. Є два стрибки загальної площі, які відбуваються при збільшенні рівня освітленості до 50 лк і 300 лк відповідно. При рівні освітленості 300 лк і вище помітна відносна стабілізація загальної площі розпізнаного пошкодження.

У наступних таблицях наведено деякі статистичні підсумки для подальшого аналізу.

2.6. Фіксована освітленість

Таблиця 2.2. Результати аналізу зображень за різної освітленості

300Лк

Total Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	841106			
	10		828421	826824	827303
	20		448877	723956	614456
	30		711583	775263	414642

300Лк

Average Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	5570			
	10		5082	5300	5590
	20		4676	5284	6335
	30		6134	6741	4459

300Лк

Number		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	151			
	10		163	156	148
	20		96	137	97
	30		116	115	93

всього записів - 1272

400Лк

Total Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	897102			
	10		949523	843811	944755
	20		757323	937241	393975
	30		661456	786561	371531

400Лк

Average Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	7294			
	10		7979	6297	7381
	20		5610	8599	7163
	30		6068	7086	6518

400Лк

Number		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	123			
	10		119	134	128
	20		135	109	55
	30		109	111	57

всього записів - 1080

500Лк

Total Area			Амплітуда, мм			
			0	0,1	0,2	0,3
ст	от	0	970622			

	10		893819	888781	883473
	20		905624	814957	549184
	30		690434	487346	442743

500Лк

Average Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	6788			
	10		6524	8229	6577
	20		6333	7087	5492
	30		8740	8122	7634

500Лк

Number		Амплітуда, мм			
		0	0,1	0,2	0,3
Частота	0	143			
	10		137	108	135
	20		143	115	100
	30		79	60	58

Всього записів – 1078; Освітленість – 300 Лк.

**Залежність загальної площі розпізнаних дефектів від амплітуди
коливань за різної частоти**

Тут і в наступних двох пунктах наведено одну лінійну та дві об'ємні діаграми, що відображають одні і ті ж дані.

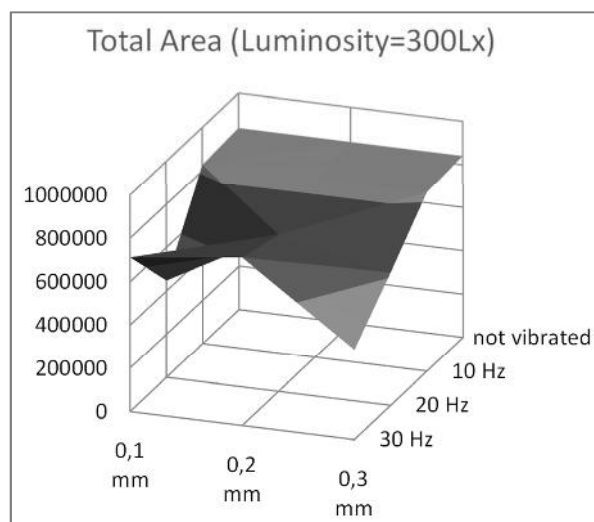
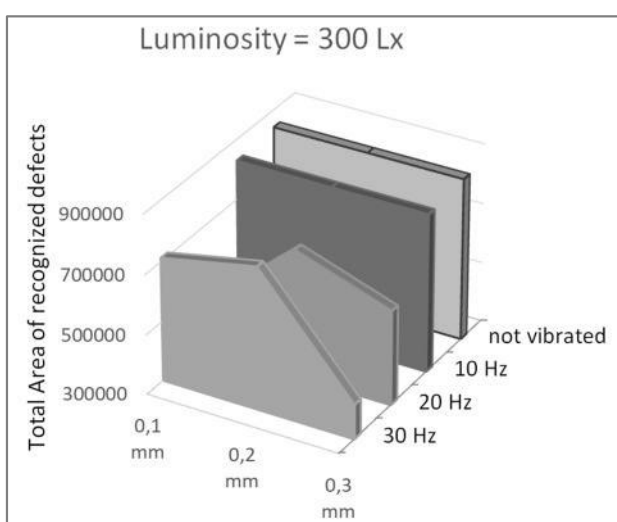
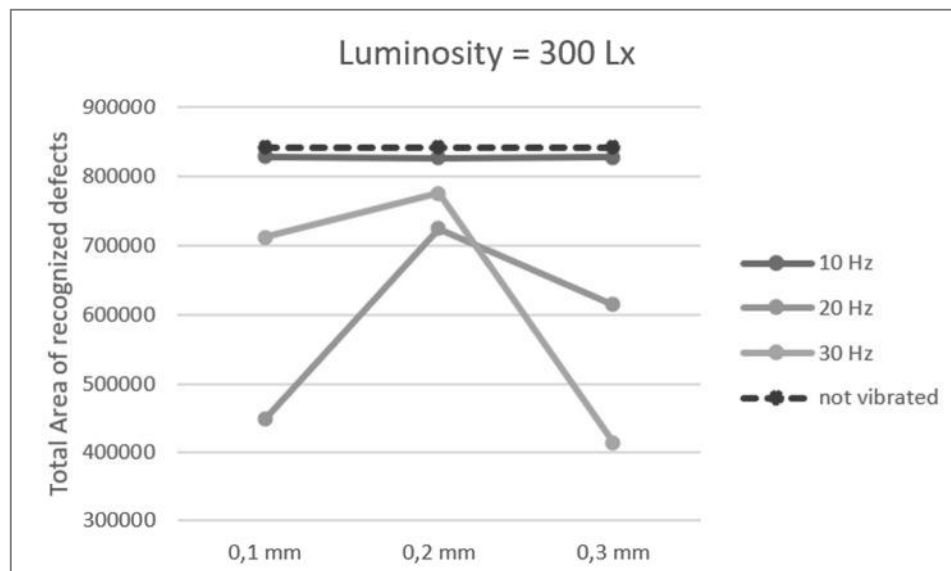


Рис. 2.16. Залежність загальної площі розпізнаних дефектів від амплітуди коливань поверхні аналізованої смуги за різної частоти

За вібрації поверхні із частотою 10 Гц значення загальної площі розпізнаних дефектів практично не змінюється порівняно із результатом, отриманим без вібрації поверхні.

Із збільшенням частоти коливань до 20 Гц та 30 Гц найбільша площа розпізнається за амплітуди коливань 0,2 мм. (Усі «зменшення» та «збільшення» можна порахувати у відсотках, тут і далі).

Найменша загальна площа дефектів розпізнається за частоти вібрації 30 Гц та амплітуди коливань 0,3 мм, – це значення загальної площі удвічі менше від еталонного, отриманого без вібрації.

Залежність кількості розпізнаних дефектів від амплітуди коливань за різної частоти

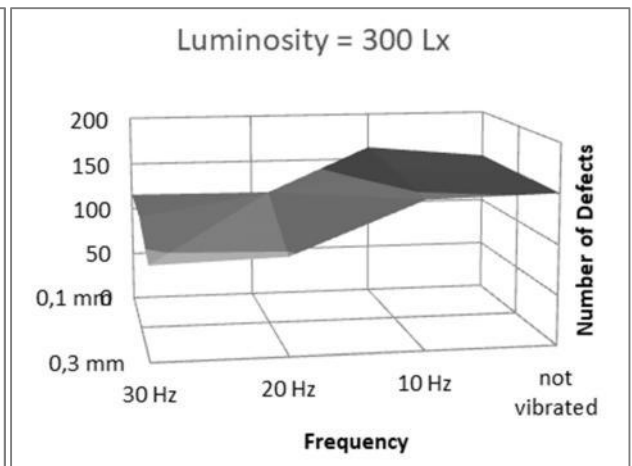
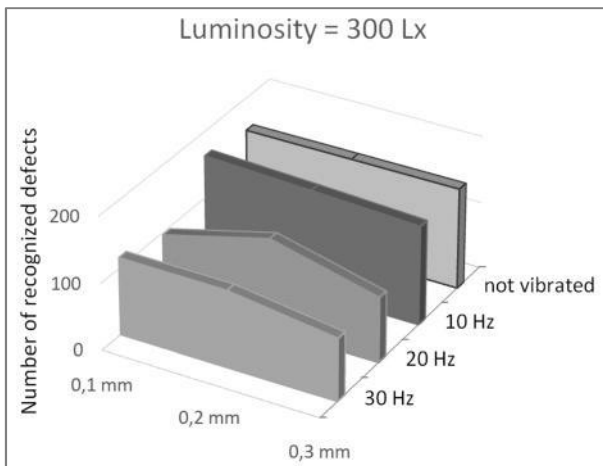
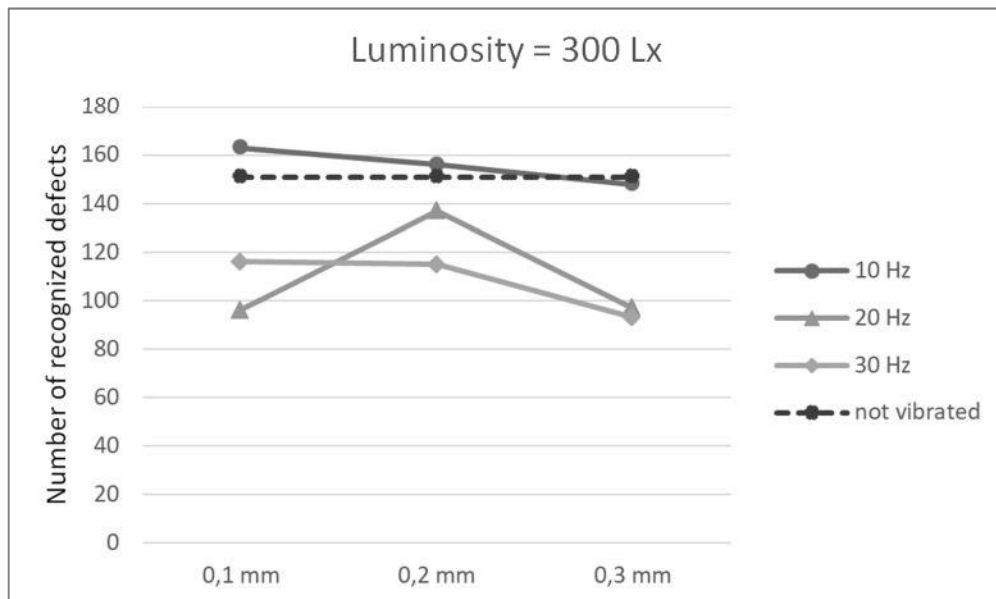


Рис. 2.17. Залежність кількості розпізнаних дефектів від амплітуди коливань за різної частоти

За вібрації поверхні з частотою 10 Гц число розпізнаних дефектів збільшується. Це можна пояснити деяким розмиттям тонких ліній («перешийків», «сполучних ланок») та сприйняттям нейронною мережею одного «дефекту» (дефектної ділянки без розривів ліній) як кількох окремих.

Із збільшенням частоти до 20 Гц та 30 Гц число розпізнаних об'єктів зменшується. Це можна пояснити тим, що нейронна мережа:

- або «втрачає» деякі дрібні дефекти (через розмиття їхніх країв)

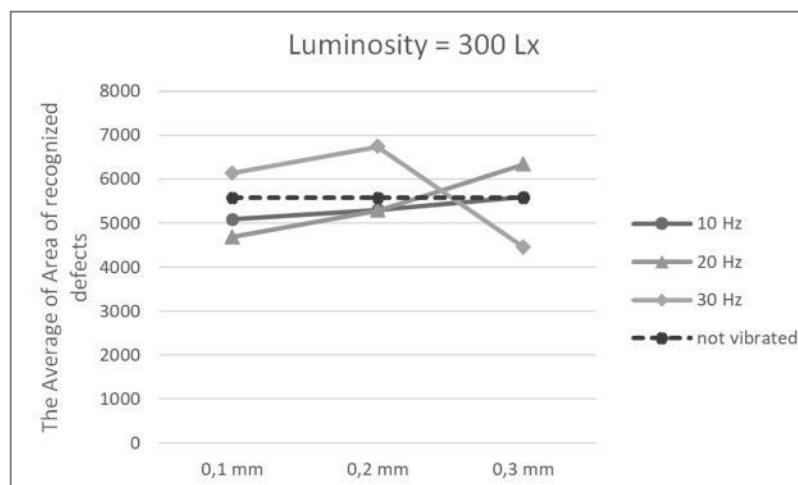
– або починає сприймати як цілісний дефект групу дефектів, розташованих досить близько один до одного.

В обох випадках очікується збільшення значення середньої площі розпізнаних об'єктів. Слід також розглянути гіпотезу про вплив параметрів коливань на ступінь «відблискування» як дефектних ділянок поверхні, так і ділянок без дефектів. Також варто врахувати можливий вплив робочих налаштувань фотокамери.

Найбільш відповідне число розпізнаних дефектів до числа, отриманого без вібрації поверхні, досягається за частоти 10 Гц та амплітуди коливань 0,2 мм та 0,3 мм.

За частоти 20 Гц кількість розпізнаних дефектів, найближча до значення, отриманого без вібрації, спостерігається за амплітуди 0,2 мм.

Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної частоти



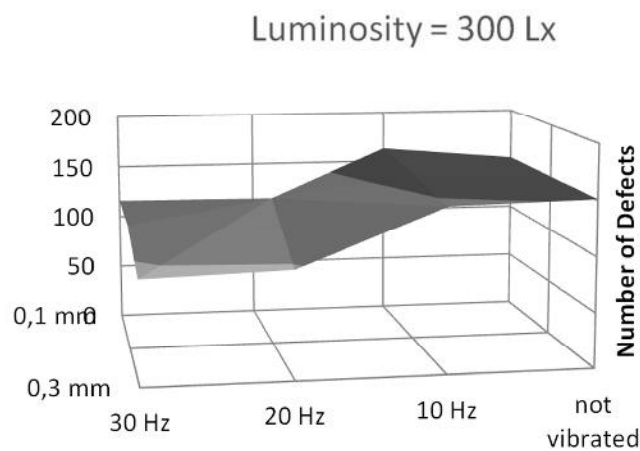
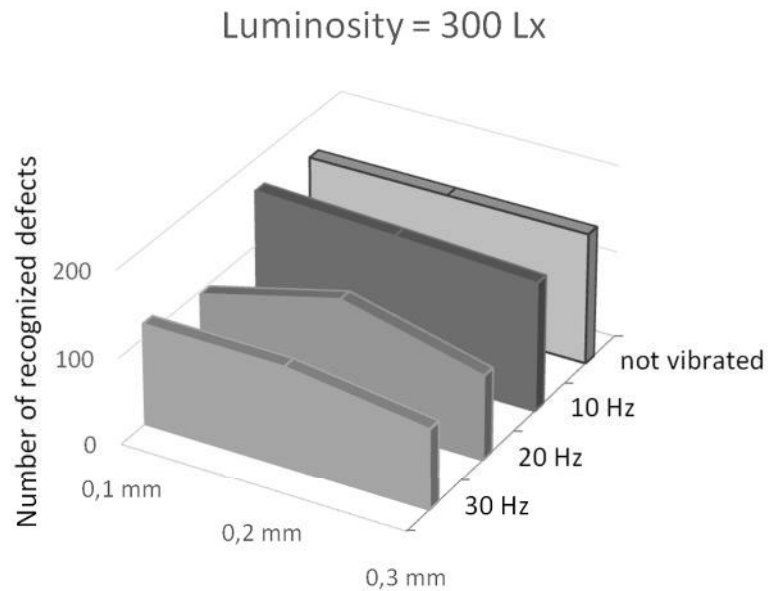


Рис. 2.18. Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної частоти

За частоти коливань 10 Гц значення середньої площі розпізнаних об'єктів зменшується. Це узгоджується із припущенням про розмиття тонких ліній та розпізнавання одного цілісного дефекту як кількох окремих, менших за площею. Якщо амплітуда коливань становить 0,3 мм, а частота – 10 Гц, то кількісні результати розпізнавання є найближчими до значень, отриманих без вібрації.

2.7. Освітленість – 400 Лк

Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної частоти

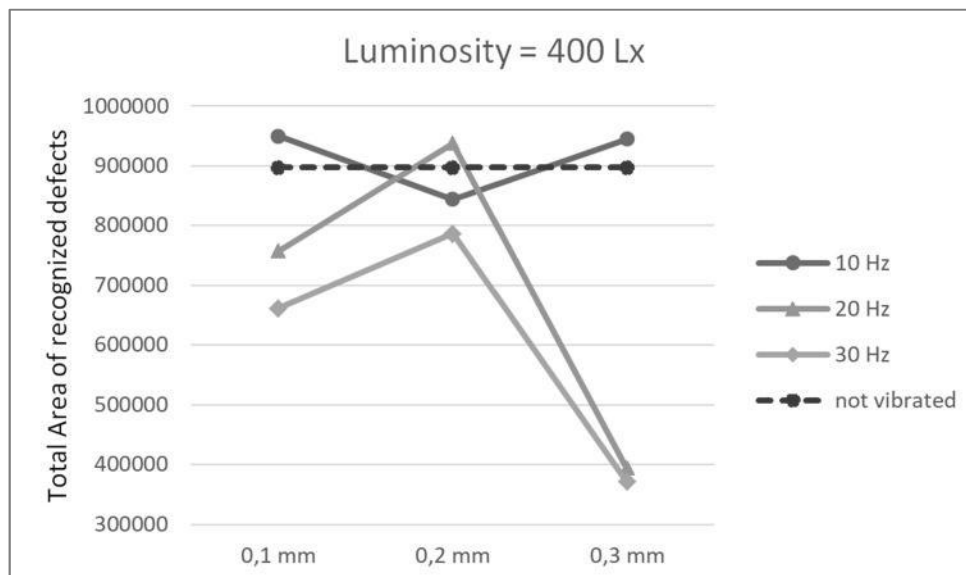


Рис. 2.19. Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної частоти

За трьох пар значень частоти вібрації та її амплітуди – (10 Гц; 0,1 мм), (20 Гц; 0,2 мм) та (10 Гц; 0,3 мм) спостерігаються дещо більші значення загальної розпізнаної площі дефектів у порівнянні із еталонним значенням, отриманим без вібрації. Це можна пояснити, зокрема, відблискуванням деяких дефектів або їх фрагментів, не розпізнаних в статичному положенні, у певному положенні відносно джерела світла при вібрації.

Таким чином, за деяких параметрів вібрування поверхні можна покращити показники розпізнавання дефектів нейромережею або ж отримати помилкові позитивні відповіді.

Найбільше значення загальної площі досягається за частоти 10 Гц та амплітуди 0,1 мм. Найменший розкид значень загальної розпізнаної площі отримано для значення амплітуди 0,2 мм, незалежно від частоти вібрації.

У випадку, коли частота вібрації встановлена на рівні 20 Гц чи 30 Гц, значення розпізнаної площі дефектів за амплітуди 0,3 мм зменшується у 2,5 рази порівняно із значенням, отриманим без вібрації.

Залежність кількості розпізнаних дефектів від амплітуди коливань за різної частоти

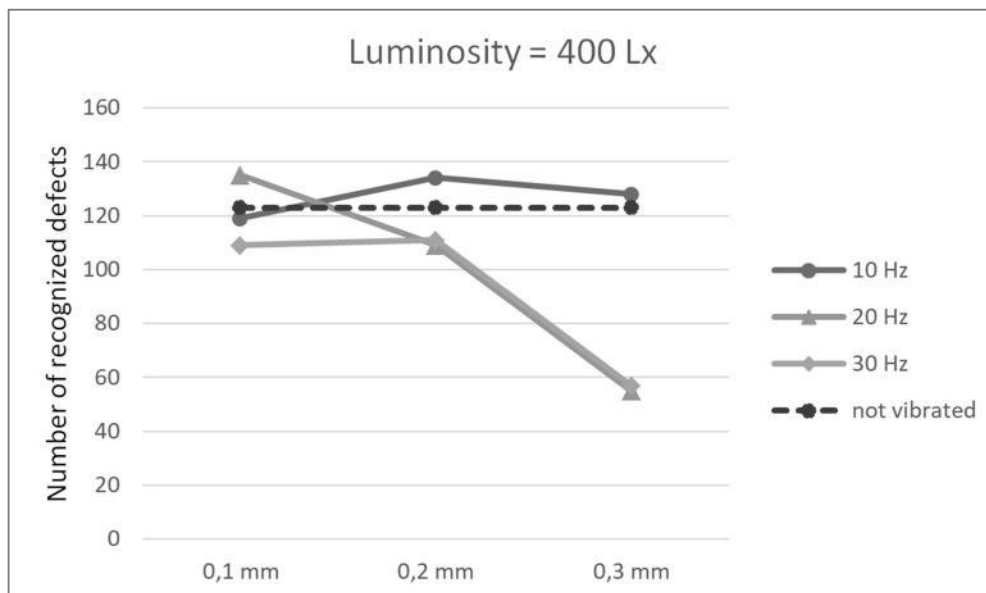


Рис. 2.20. Залежність кількості розпізнаних дефектів від амплітуди коливань за різної частоти

За певних параметрів вібрування – (10 Гц; 0,2 мм), (10 Гц; 0,3 мм) та (20 Гц; 0,1 мм) мережею розпізнається більше число дефектів, ніж за відсутності вібрації. Це можна пояснити розмиттям на зображенні дуже тонких з'єднань в межах цілісного дефекту, за рахунок чого один об'єкт розпізнається як два чи більше. Найближчі до еталонних значень отримано за частоти коливань 10 Гц, незалежно від амплітуди коливань. Із збільшенням частоти до 20 Гц чи 30 Гц та амплітуди до 0,3 мм кількість окремих розрізнених дефектів зменшується удвічі. Це можна пояснити спільним впливом на отримане зображення робочих налаштувань фотокамери та параметрів вібрації поверхні.

Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної частоти.

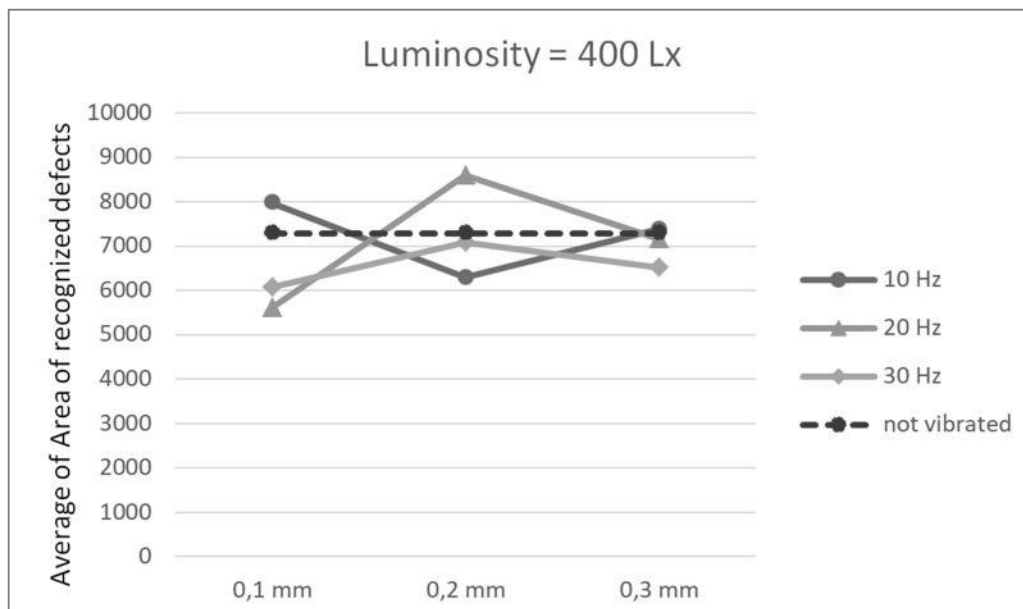


Рис. 2.21. Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної частоти

Діаграма є наслідком з попередніх двох. Якщо за певних умов вібрації фотокамера «губить» окремі дрібні об'єкти, то середня площа виявлених об'єктів збільшується. Якщо камера не фіксує тонкі з'єднувальні ланки у цілісних дефектах, то число дефектів штучно збільшується і їх середня площа стає меншою.

2.8. Освітленість – 500 Лк

Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної частоти

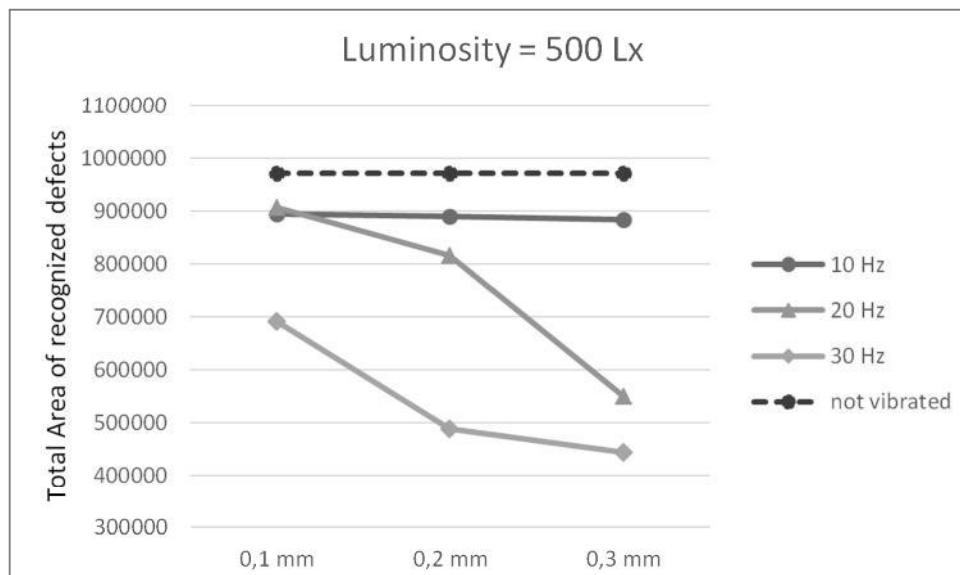


Рис. 2.22. Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної частоти

Усі дев'ять можливих налаштувань вібрації, що розглядаються, погіршують кількісні показники розпізнавання дефектів.

За вібрації поверхні із частотою 10 Гц значення загальної площі розпізнаних дефектів зменшується порівняно із результатом, отриманим без вібрації поверхні, та є практично стабільним відносно амплітуди коливань.

Із збільшенням частоти коливань до 20 Гц та 30 Гц та збільшенням амплітуди загальна площа дефектів суттєво зменшується. За комбінації значень (30 Гц; 0,3 мм) загальна площа зменшилась більше, ніж удвічі, порівняно із значенням, отриманим без вібрації поверхні.

Оптимальними параметрами вібрації за рівня освітленості 500 Лк є (10 Гц; 0,1 мм) або (20 Гц; 0,1 мм).

Цікавим є факт, що найменше дрібне дрижчання поверхні за даного рівня освітленості (вібрація 30 Гц та амплітуда 0,1 мм) значно погіршує кількісні показники розпізнавання поверхневих дефектів порівняно із еталонними.

Залежність кількості розпізнаних дефектів від амплітуди коливань за різної частоти.

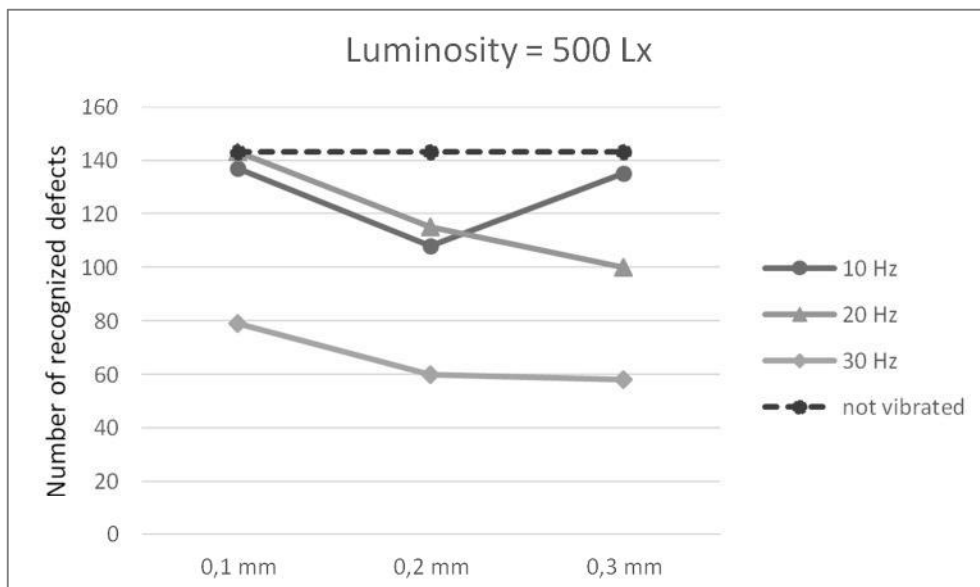


Рис. 2.23. Залежність кількості розпізнаних дефектів від амплітуди коливань за різної частоти

За комбінацій (10 Гц; 0,1 мм), (20 Гц; 0,1 мм) та (10 Гц; 0,3 мм) розрізнено практично ту ж кількість окремих дефектів, що й в експерименті без вібрації.

Із збільшенням амплітуди коливань від 0,1 мм до 0,2 мм число розрізнених дефектів стрімко зменшується. Ця тенденція зберігається для наступного збільшення амплітуди коливань до 0,3 мм, якщо частота коливань становить 20 Гц або 30 Гц.

Цікавим є факт протилежної динаміки у розрізненні кількості дефектів за частот 10 Гц та 20 Гц при зміні амплітуди коливань від 0,2 до 0,3 мм.

Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної частоти

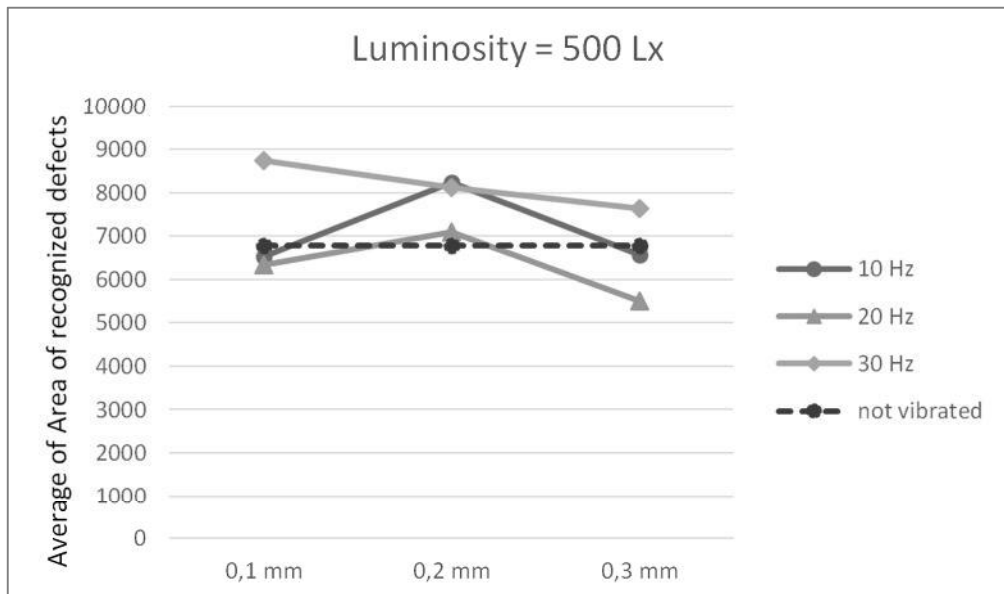


Рис. 2.24. Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної частоти

Діаграма отримана як наслідок з двох попередніх результатів. За режимів вібрації (10 Гц; 0,1 мм) та (20 Гц; 0,1 мм) кількісні показники розпізнавання площі та кількості дефектів практично співпадають із отриманими без вібрації.

2.9. Фіксована частота

Таблиця 2.3. Результати аналізу зображень за сталої освітленості (300Лк)

Total Area		10 Гц			
		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	841106	828421	826824	827303
	400 Лк	897102	949523	843811	944755
	500 Лк	970622	893819	888781	883473

Number of Defects		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	151	163	156	148
	400 Лк	123	119	134	128
	500 Лк	143	137	108	135

10 Гц

Average Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	5570	5082	5300	5590
	400 Лк	7294	7979	6297	7381
	500 Лк	6788	6524	8230	6544

20 Гц

Total Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	841106	448877	723956	614456
	400 Лк	897102	757323	937241	393975
	500 Лк	970622	905624	814957	549184

20 Гц

Number of Defects		Амплітуда, мм			
		0	0,1	0,2	0,3

Luminosity	300 Лк	151	96	137	97
	400 Лк	123	135	109	55
	500 Лк	143	143	115	100

20 Гц

Average Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	5570	4676	5284	6335
	400 Лк	7294	5610	8599	7163
	500 Лк	6788	6333	7087	5492

30 Гц

Total Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	841106	711583	775263	414642
	400 Лк	897102	661456	786561	371531
	500 Лк	970622	690434	487346	442743

30 Гц

Number of Defects		Амплітуда, мм			
		0	0,1	0,2	0,3

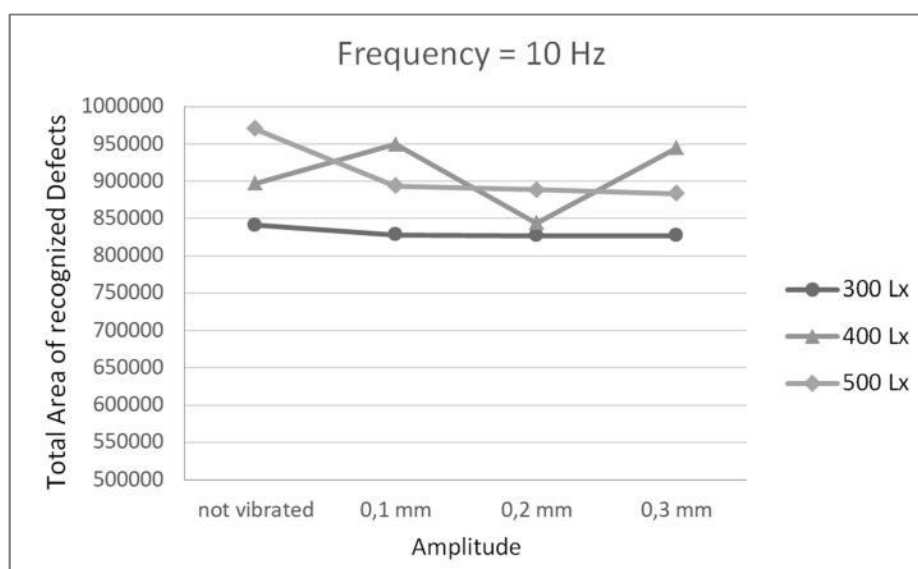
Luminosity	300 Лк	151	116	115	93
	400 Лк	123	109	111	57
	500 Лк	143	79	60	58

30 Гц

Average Area		Амплітуда, мм			
		0	0,1	0,2	0,3
Luminosity	300 Лк	5570	6134	6741	4459
	400 Лк	7294	6068	7086	6518
	500 Лк	6788	8740	8122	7634

Б.І. Частота – 10 Гц

Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної освітленості



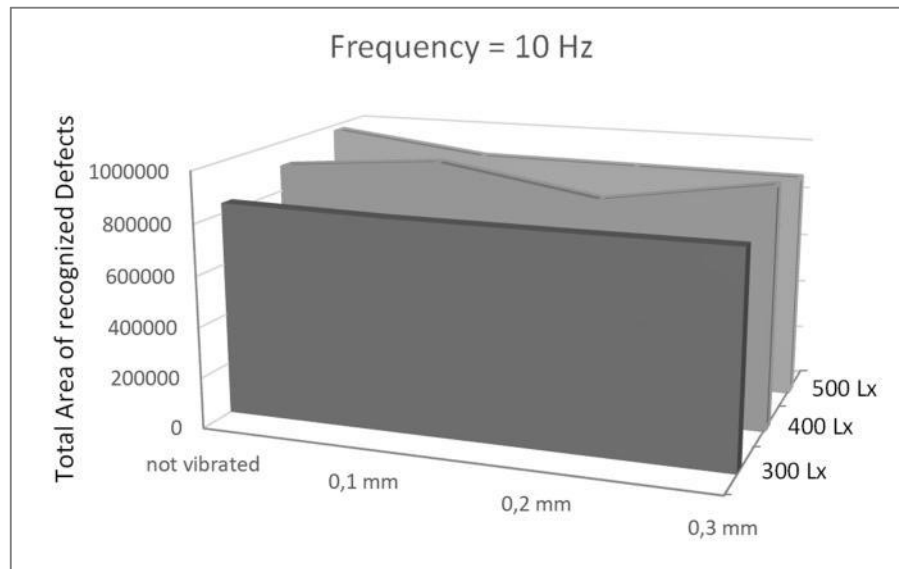


Рис. 2.25. Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної освітленості

Якщо для рівнів освітленості 300 Лк і 500 Лк амплітуда коливань збільшується, то значення загальної площі розпізнаних дефектів зменшується порівняно із результатом, отриманим без вібрації поверхні. Причому за освітленості на рівні 500 Лк поява вібрацій поверхні відразу суттєво зменшує показник загальної розпізнаної площі. За освітленості на рівні 400 Лк та амплітуди коливань на рівні 0,1 мм та 0,3 мм загальна площа дефектів є більшою у порівнянні з невіброваною поверхнею.

Залежність кількості розпізнаних дефектів від амплітуди коливань за різної освітленості

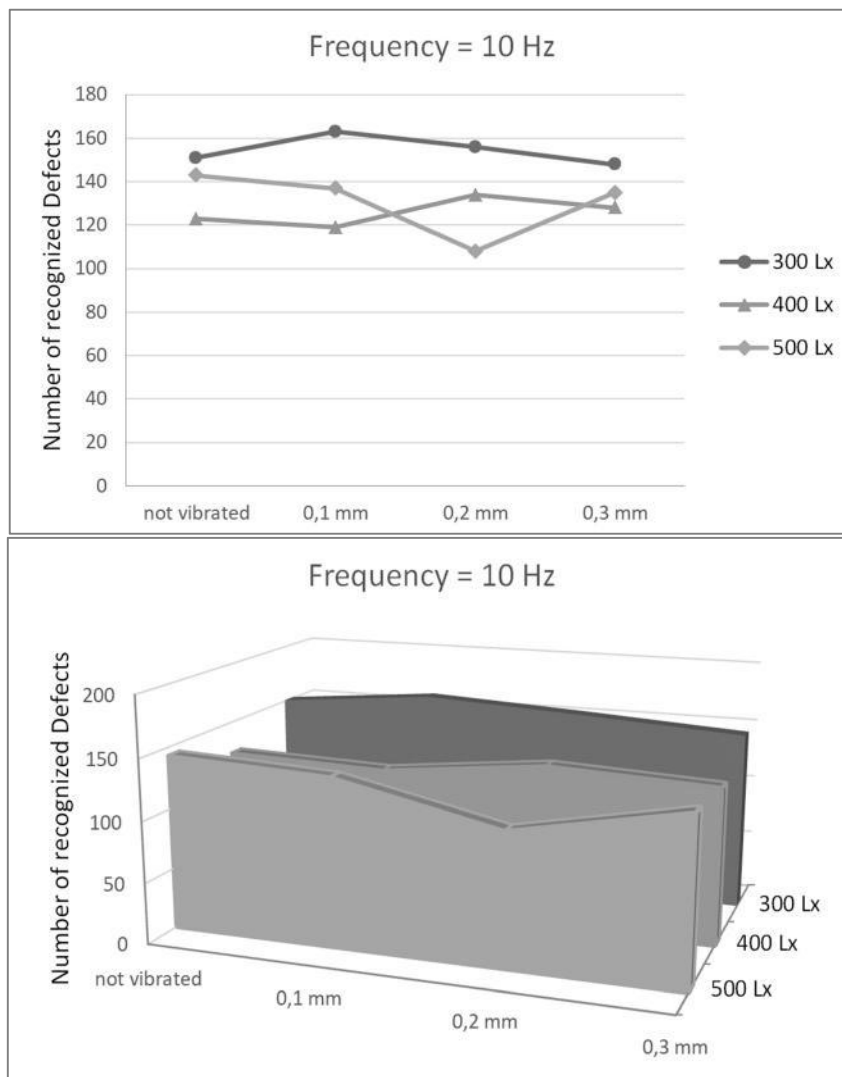


Рис. 2.26. Залежність кількості розпізнаних дефектів від амплітуди коливань за різної освітленості

Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної освітленості

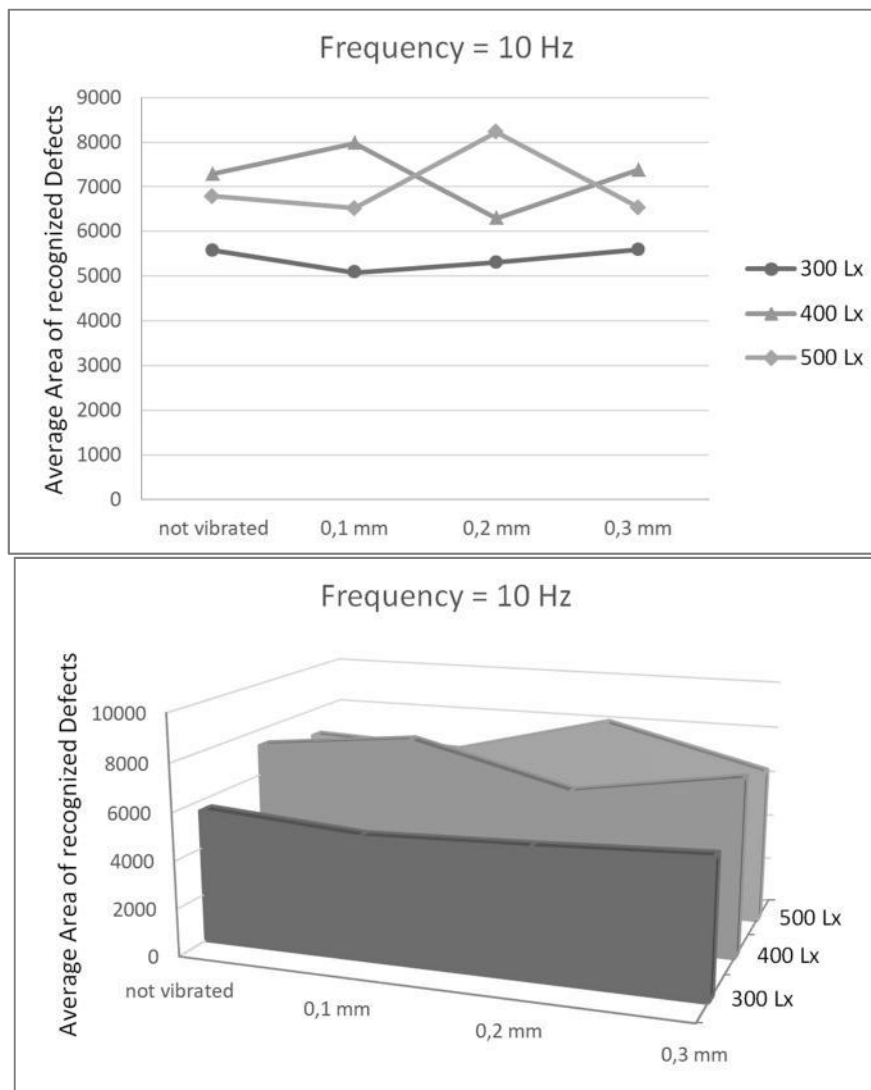


Рис. 2.27. Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної освітленості

За рівня освітленості 300 Лк отримано найстабільніші показники щодо загальної площі, кількості та середньої площі розпізнаних дефектів, незалежно від амплітуди коливань.

Проте за освітленості 400 Лк і 500 Лк існують такі режими вібрації, за яких можна значно збільшити значення розпізнаної площі дефектів.

Б.ІІ. Частота – 20 Гц

Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної освітленості

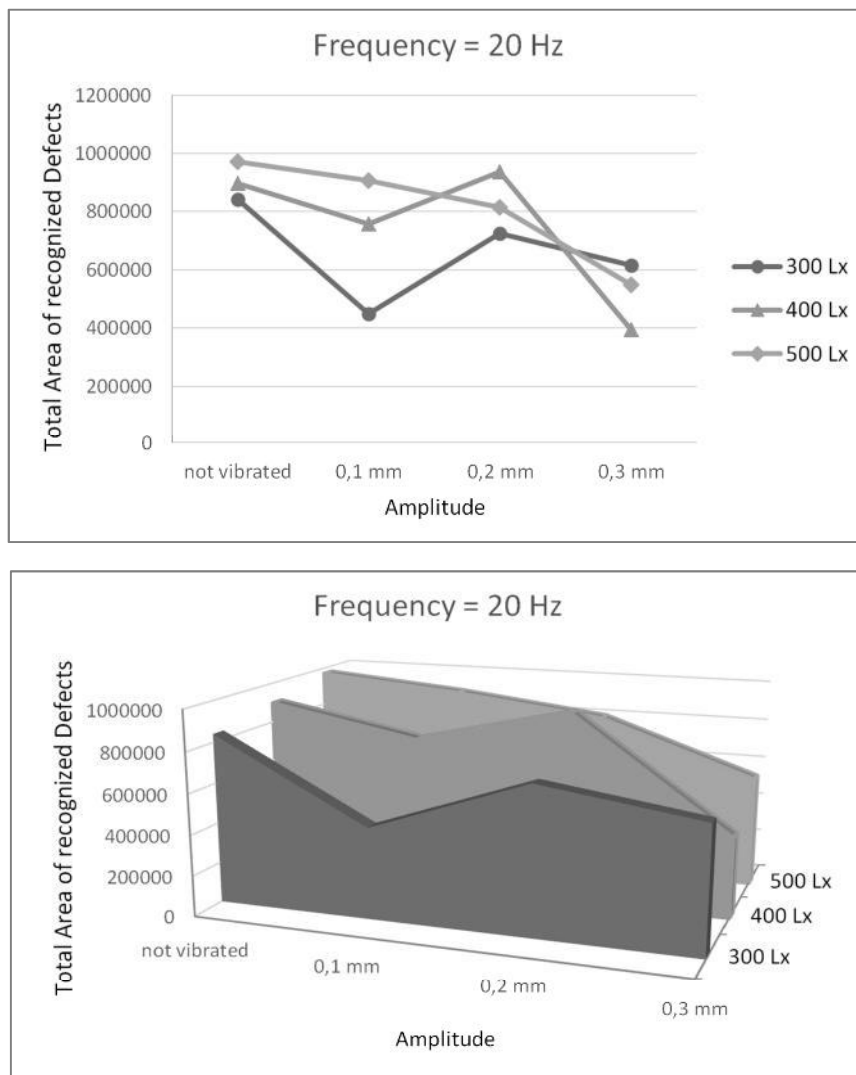


Рис. 2.28. Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної освітленості

Для всіх рівнів освітленості спостерігається зменшення розпізнаної площі дефектів при початку вібрування з найменшою амплітудою – 0,1 мм. Збільшення амплітуди ще погіршує цей показник за яскравій освітленості поверхні (500 Лк).

Цікавим є факт, що за освітленості 400 Лк та амплітуди 0,2 мм розпізнається практично уся площа поверхневих дефектів, розпізнана за різних рівнів освітленості, але без вібрації.

За показників (300 Лк; 0,1 мм) та (400 Лк; 0,3 мм) значення розпізнаної площі поверхні з дефектами є мінімальними.

Залежність кількості розпізнаних дефектів від амплітуди коливань за різної освітленості

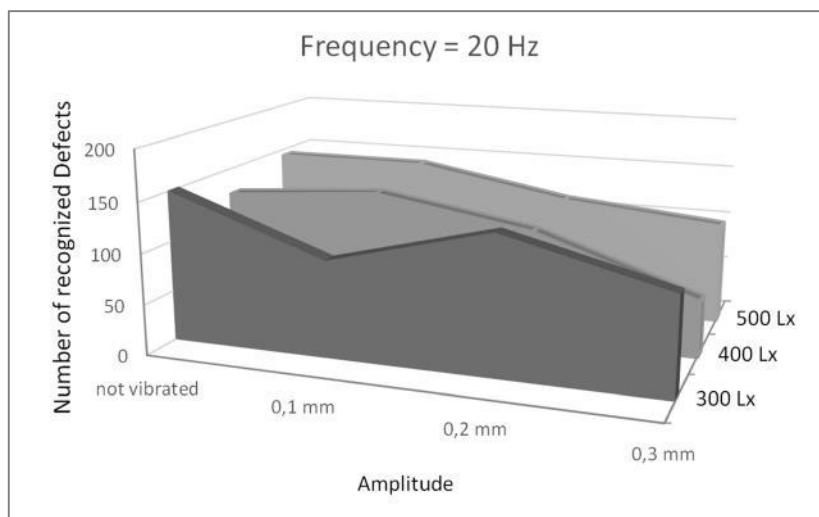
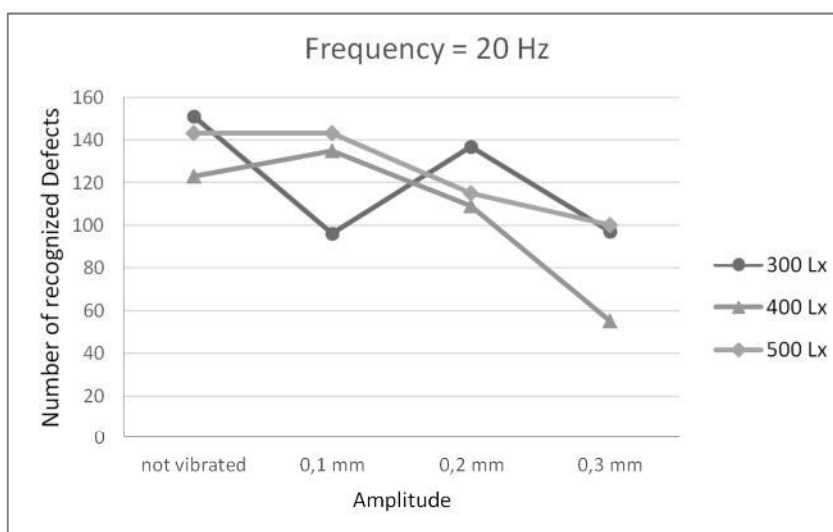


Рис. 2.29. Залежність кількості розпізнаних дефектів від амплітуди коливань за різної освітленості

Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної освітленості

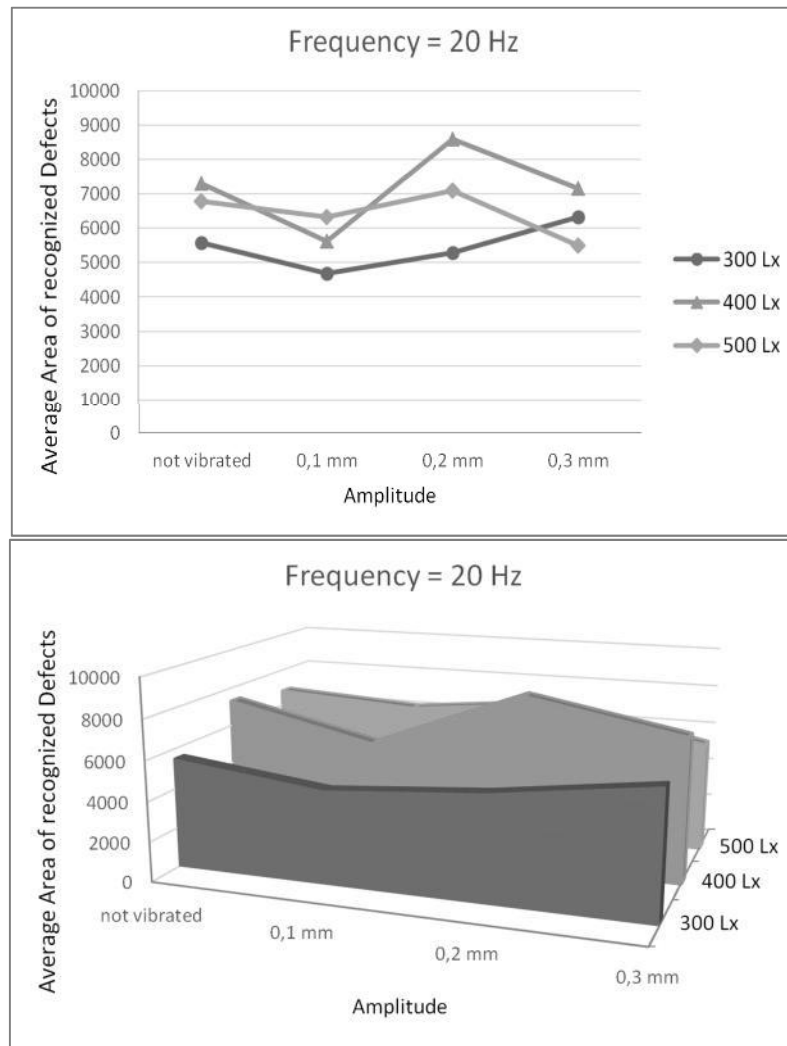


Рис. 2.30. Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної освітленості

Б.III. Частота – 30 Гц

Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної освітленості

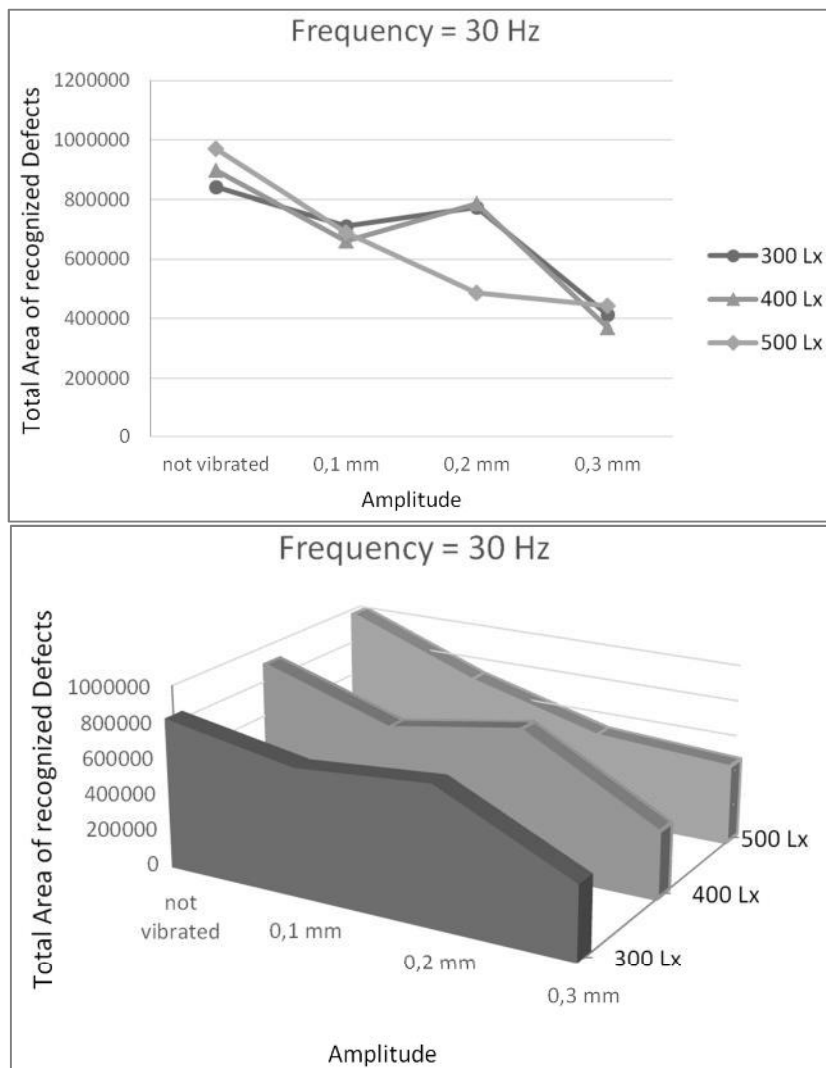


Рис. 2.31. Залежність загальної площі розпізнаних дефектів від амплітуди коливань за різної освітленості

За всіх рівнів освітленості спостерігається зменшення розпізнаної площі дефектів для всіх режимів вібрації, якщо порівнювати з показниками, отриманими в статичному стані. Слід відзначити, що за амплітуди 0,2 мм та двох рівнів освітленості (300 Лк, 400 Лк) площа розпізнаних дефектів є у півтора рази більшою, ніж за аналогічного режиму вібрації, але з яскравішим освітленням (500 Лк). За амплітуди 0,3 мм значення розпізнаної площі поверхні з дефектами є найменшими.

Залежність кількості розпізнаних дефектів від амплітуди коливань за різної освітленості

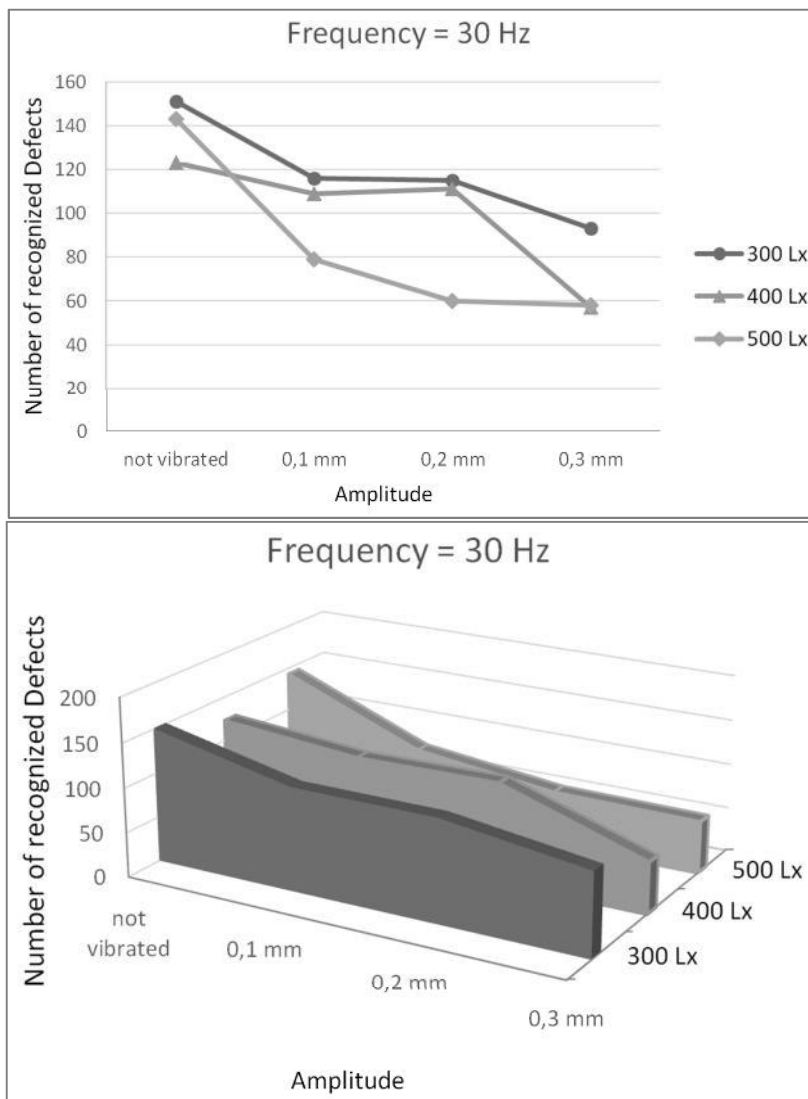


Рис. 2.32. Залежність кількості розпізнаних дефектів від амплітуди коливань за різної освітленості

Слід відзначити, що за двох рівнів освітленості (300 Лк, 400 Лк) збільшення амплітуди удвічі (від 0,1 мм до 0,2 мм) практично не впливає на кількість окремих розпізнаних дефектів.

Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної освітленості

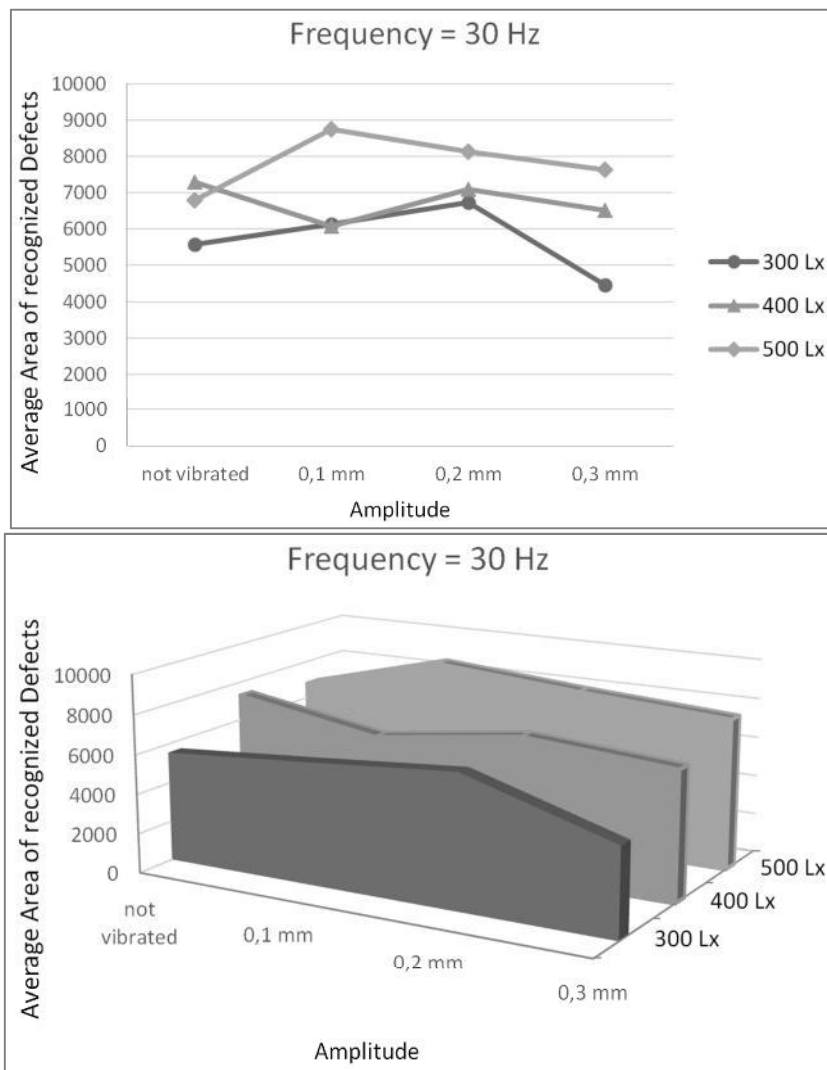


Рис. 2.33. Залежність середньої площі розпізнаних дефектів від амплітуди коливань за різної освітленості

Із появою вібрацій частотою 30 Гц значно зменшується число дефектів, розпізнаних як окремі цілісні об'єкти. При цьому середня площа розпізнаних дефектів за освітленості 300 Лк і 500 Лк дещо збільшується (наступні діаграми). Це може свідчити про «втрату» здатності виявляти окремі дрібні дефекти (камерою або неймережею) або злиття в цілісні об'єкти двох або більше, розташованих дуже близько.

3. СПЕЦІАЛЬНА ЧАСТИНА

3.1. Алгоритм локалізації та класифікації поверхневих дефектів на сталевому листі

У металургії, однією з головних проблем є виявлення дефектних поверхонь металопрокату. Але якщо їх не виявити, це може зумовити виникнення браку готової продукції та втрату репутації компанії.

Тому, важливо допомогти інженерам удосконалити алгоритм дефектоскопії та дефектометрії дефектів шляхом локалізації та класифікації поверхневих пошкоджень на сталевому листі. У разі успіху це допоможе підтримувати високі стандарти виробництва сталі. Розглянемо приклад алгоритму з виявлення дефектів на основі аналізу набору даних який надано одним із провідних світових виробників сталі. Набір даних містить 3 функції – ImageId, ClassId і Encoded Pixels.

- ClassId містить чотири унікальних класи, які є 4 типами дефектів і можуть бути окремо або одночасно присутні на зображенні. У наборі даних є 7095 зображень. Сегментація кожного класу дефектів кодується в один рядок, навіть якщо на зображенні є кілька несуміжних дефектів. Сегменти мають форму закодованих пікселів.

Перевірка дисбалансу

По-перше, визначали аналізовані дані збалансовані чи ні. Нижче наведено простий графік залежності ідентифікатора класу від кількості.

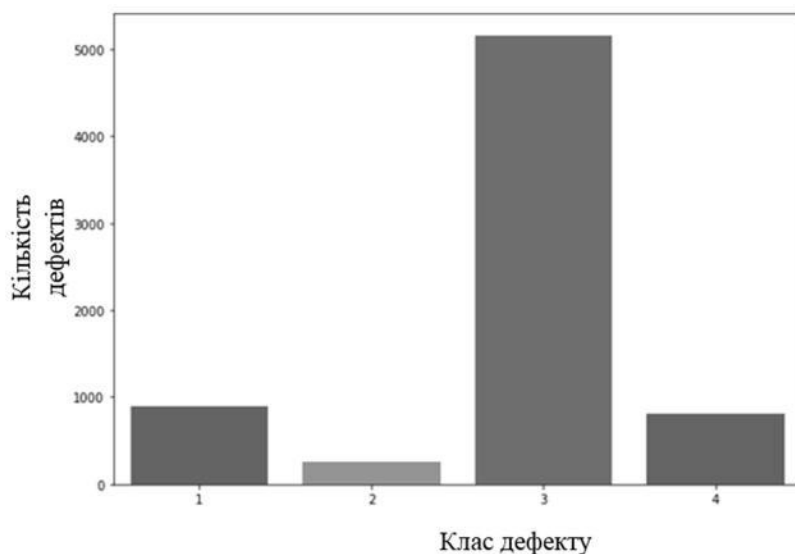


Рис. 3.1. Графік залежності ідентифікатора класу дефекту від їх кількості у вибірці

Помітно, що дефект типу 3 є більш переважаючим, ніж будь-який інший тип дефекту. Дефект 2 є найменш поширеним дефектом. Отже, виявлено класовий дисбаланс.

Наявність дефектів що накладаються

Перевіряли, чи містить вхідне зображення більше одного дефекту одночасно.

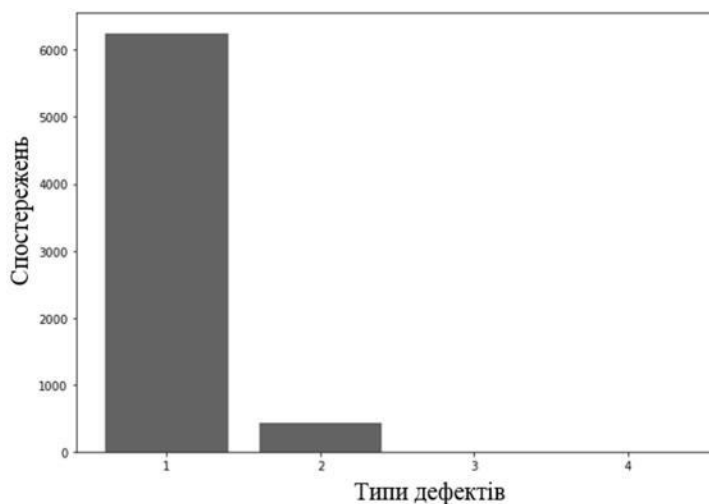


Рис. 3.2. Типи дефектів

Більшість зображень містять лише один тип дефекту. Окремі, мають два дефекти одночасно. Зображень з трьома і більшою кількістю дефектів у вибірці не було.

Приналежність пікселів до масок

Закодовані пікселі - це інформація про пікселі, які належать дефектам. Він містить шаблон - піксельний індекс, за яким слідує кількість. Це інтерпретується як індекс пікселя до індексу пікселя + значення підрахунку — це всі пікселі з дефектами. Цей шаблон повторюється, доки не будуть закодовані всі пікселі з дефектами.

Наступна функція перетворює закодовані пікселі на маски:

```
def masks(encoded_pixels):
```

```
    counts=[]
```

```
    mask=np.zeros((256*1600), dtype=np.int8)
```

```
    pre_mask=np.asarray([int(point) for point in encoded_pixels.split()])
```

```
    for index,count in enumerate(pre_mask):
```

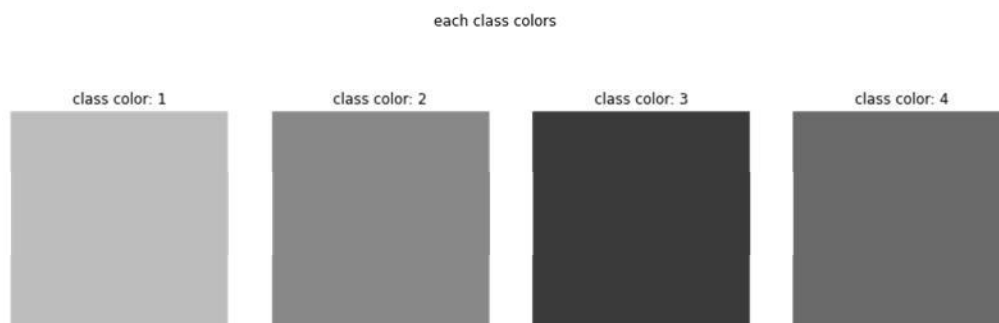
```

if(index%2!=0):
    counts.append(count)
i=0
for index,pixel in enumerate(pre_mask):
    if(index%2==0):
        if(i==len(counts)):
            break
        mask[pixel:pixel+counts[i]]=1
        i+=1
mask=np.reshape(mask,(1600,256))
mask=cv2.resize(mask,(256,1600)).T
return mask

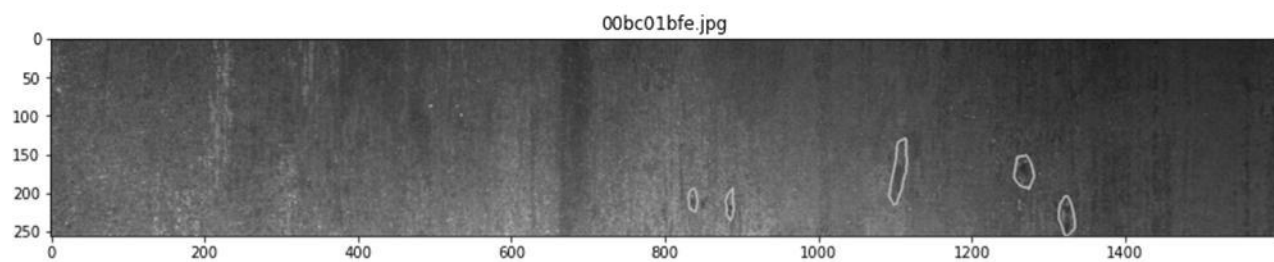
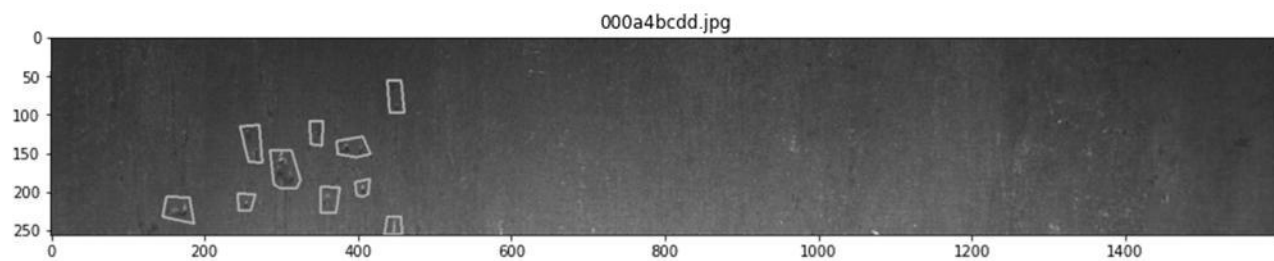
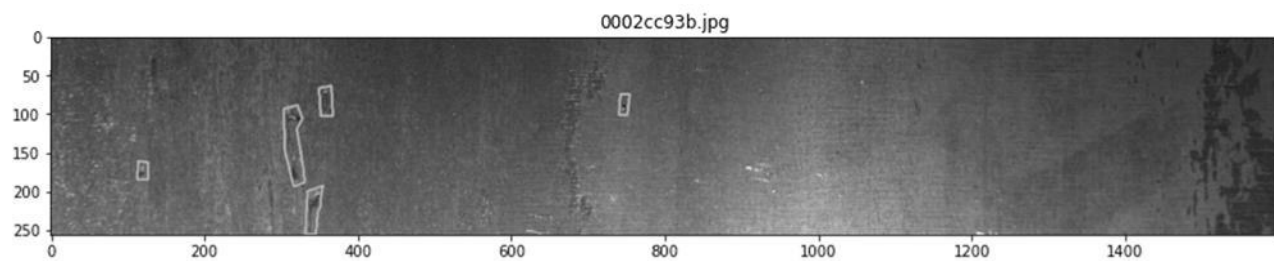
```

Вхідне зображення та цільові маски

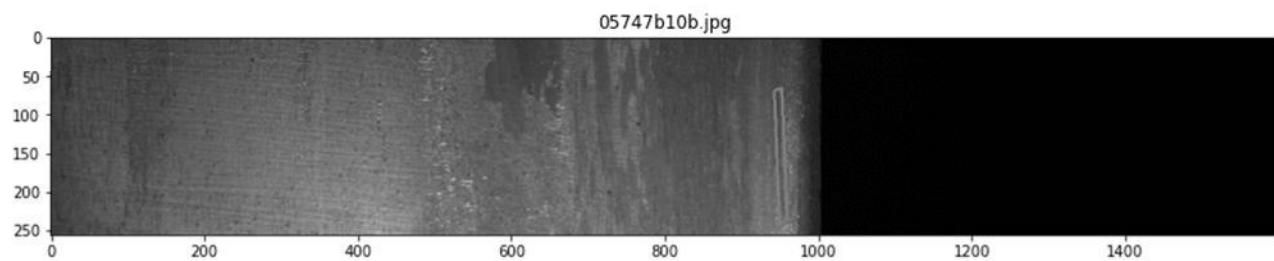
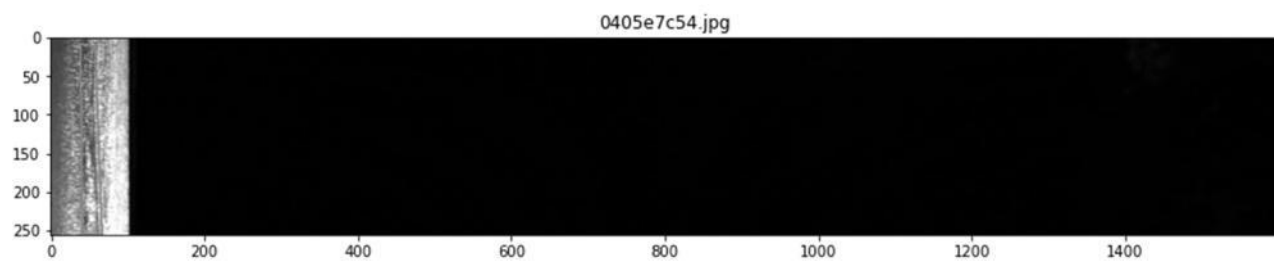
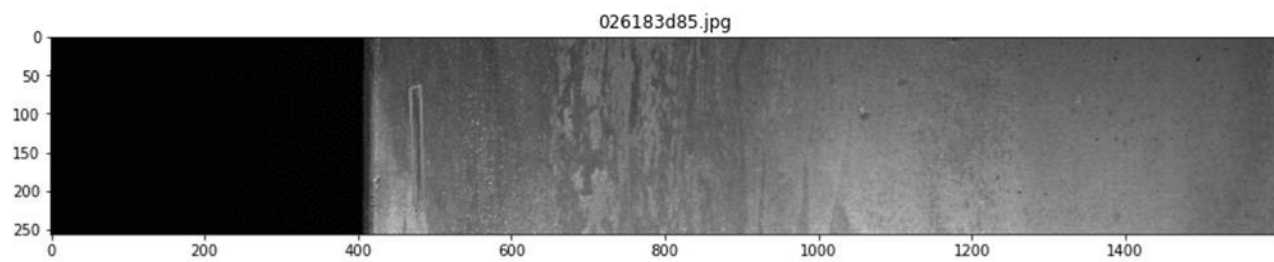
Тут ми візуалізуємо, як виглядають вхідне зображення та цільові маски. Я надав кожному типу дефекту різний колір, щоб було легше розрізняти дефекти.



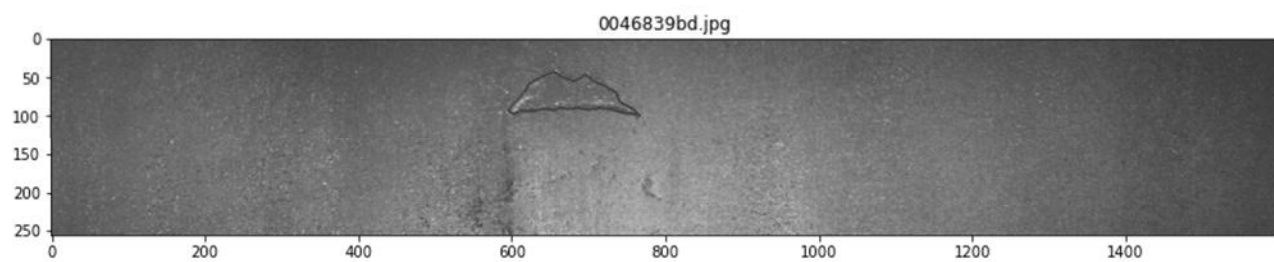
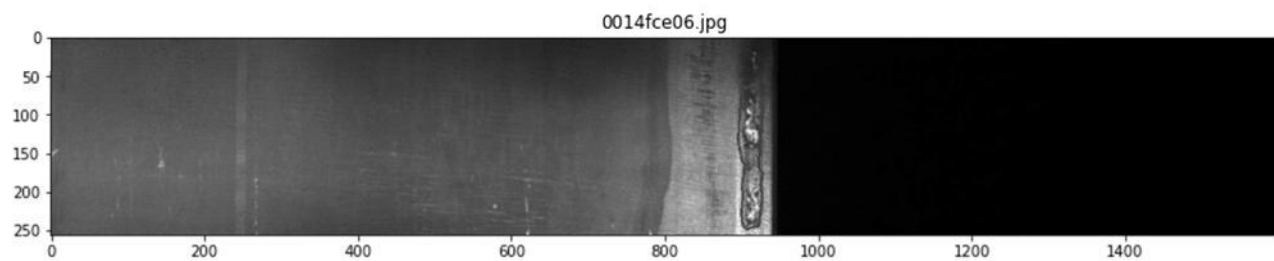
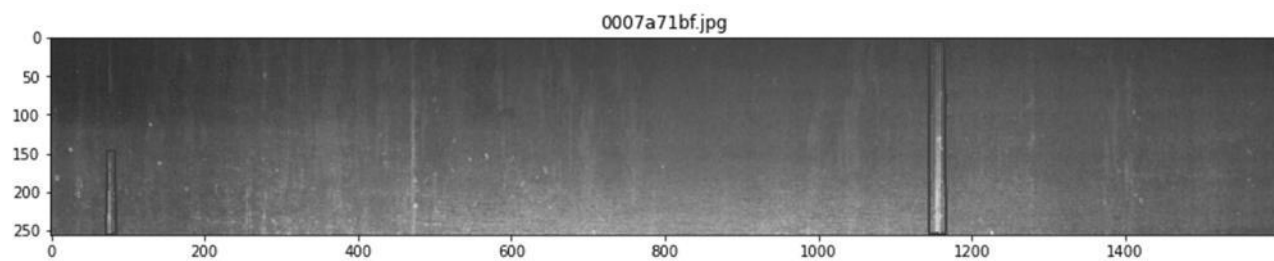
Дефект 1 типу



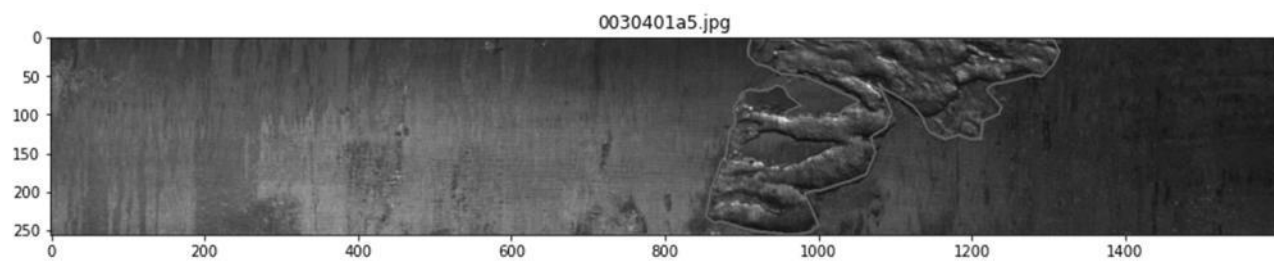
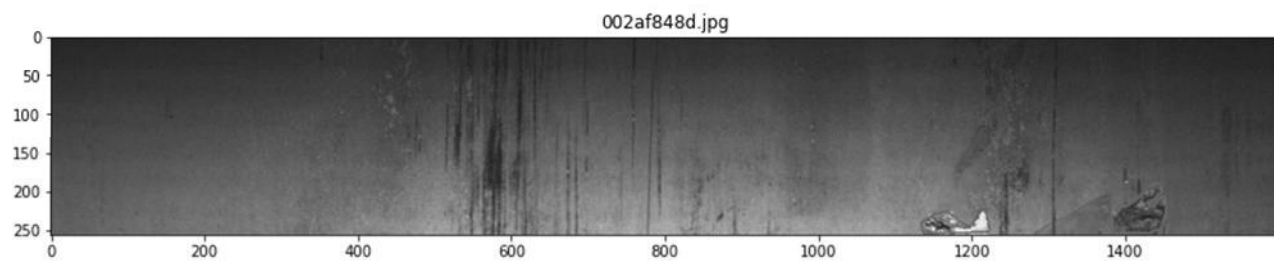
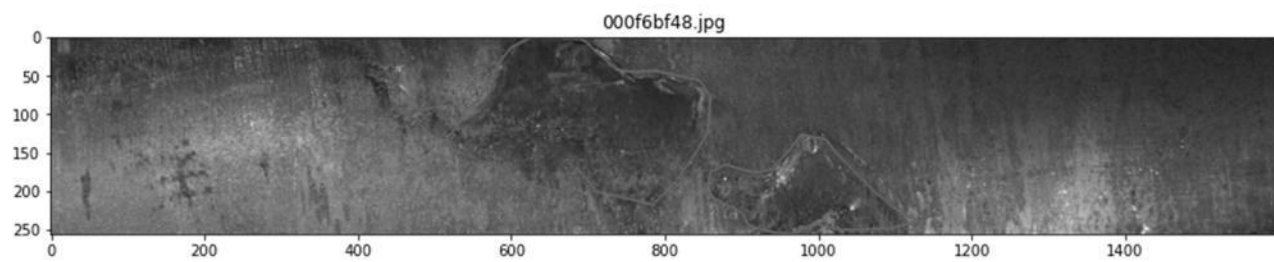
Дефект 2 типу



Дефект 3 типу



Дефект 4 типу



Два дефекти одночасно

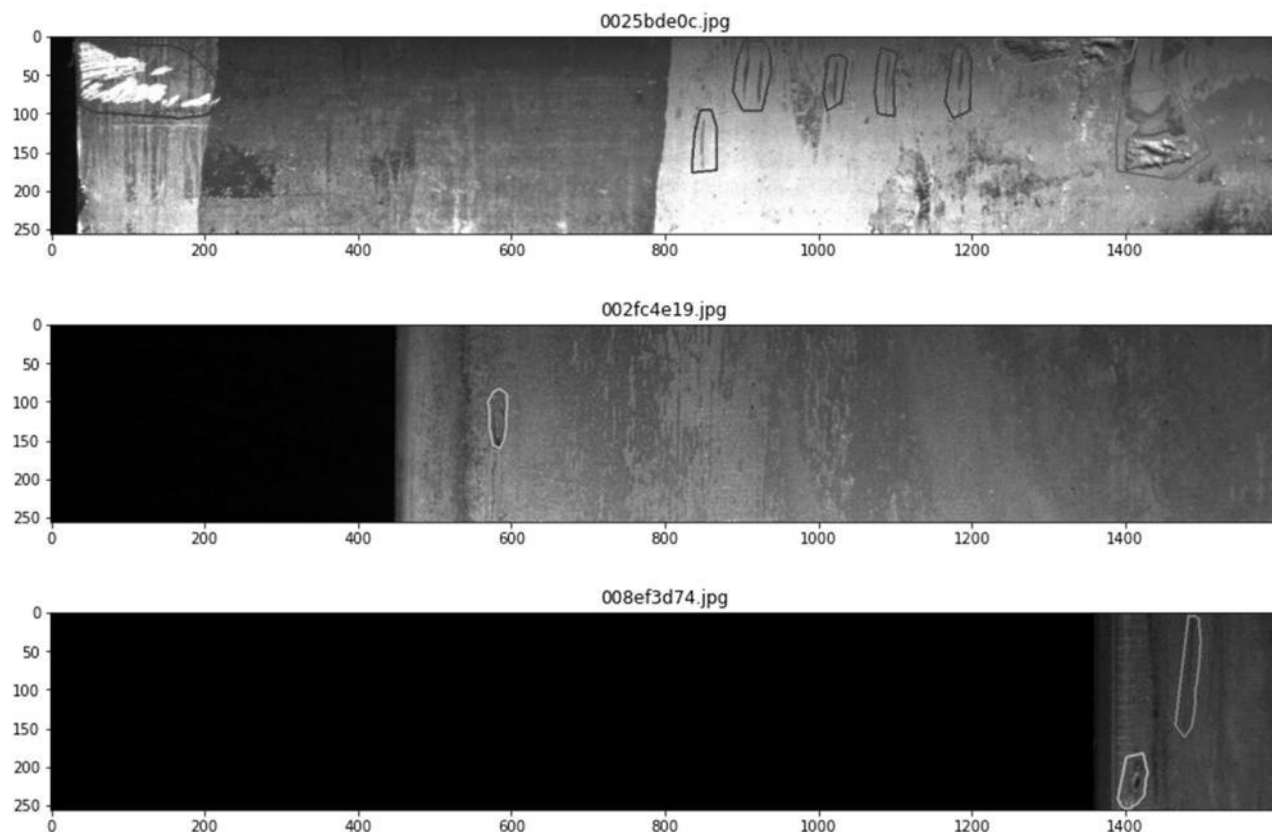


Рис. 3.3. Експлуатаційні дефекти та їх розпізнавання

З наведених вище візуалізацій можна помітити, що дефект типу 4 є дуже чітким, ніж будь-які інші дефекти. Дефект 1 типу важко побачити. Дефект 2 і 3 типу виглядає схожим. Через дисбаланс дефекти типу 2 можуть бути класифіковані моделлю як тип 3, якщо дані не збалансовані.

Розробка функцій

Виявлено, що класи дефектів сильно розбалансовані. Крім того, вибірка діагностичних зображень є невеликою. Тому зображення було доповнено таким чином, щоб це компенсувало дисбаланс класів, а також збільшило обсяг навчальних даних. Клас 3 не було доповнено, оскільки він містить достатньо даних. Класи 1 і 2 були доповнені таким чином, що їх обсяги зросли в 5 разів. Оскільки для 4-го класу було найменше доступних даних, його збільшено у 8 разів, щоб компенсувати дисбаланс.

На цьому кроці доповнено обидва зображення та їх маски, а потім зображення збережено. Значення масок були перетворені в піксельне кодування за допомогою наведеної нижче функції.

```
def rle(img):  
    pixels= img.T.flatten()  
    pixels = np.concatenate([[0], pixels, [0]])  
    runs = np.where(pixels[1:] != pixels[:-1])[0]  
    runs[1::2] -= runs[:-1]  
    return ' '.join(str(x) for x in runs)
```

Загальну кількість зображень збільшено з 7000 до 13000.

Метрика ефективності

Дайс-коефіцієнт обрано як показник ефективності, оскільки маски результатів повинні мати хорошу точність, так і добре запам'ятовування.

Функція втрат

Використано бінарну крос-ентропію як функцію втрат, оскільки розуміння елемента, що належить до певного класу, не повинно впливати на рішення щодо іншого класу, оскільки певні зображення можуть містити більше ніж 1 клас.

ResUNet - це модель семантичної сегментації, що поєднує глибоке залишкове навчання і UNet. Архітектура використовує переваги обох моделей Residual і UNet.

Ця комбінація забезпечує дві переваги:

- 1) залишковий блок полегшить навчання мережі;
- 2) пропуск зав'язків у залишковій одиниці та між низькими та високими рівнями мережі сприятиме розповсюдженню інформації без деградації, дозволяючи розробити нейронну мережу з набагато меншою кількістю параметрів, однак з порівняно кращою продуктивністю семантичної сегментації.

Архітектура CNN

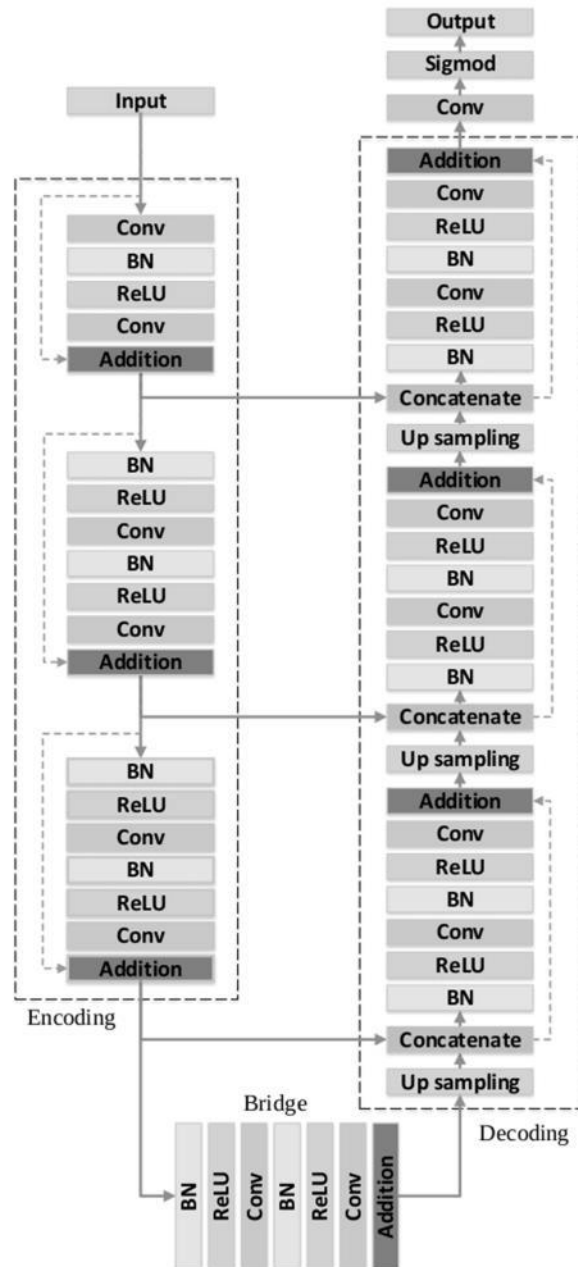


Рис. 3.4. Мережа містить 3 частини: кодування, мосту та декодування

1 Перша частина кодує вхідне зображення в компактні представлення. Остання частина відновлює представлення до піксельної категоризації, тобто семантичної сегментації. Середня частина є мостом, що з'єднує шляхи кодування і декодування. Усі три частини побудовані із залишковими одиницями, які містять два блоки згортки 3×3 та ідентифікаційного відображення. Кожен блок згортки має шар BN, рівень

активації ReLU і згортковий рівень. Відображення ідентичності з'єднує вхід і вихід пристрою.

Архітектура містить два шляхи.

Перший шлях - це шлях скорочення (також званим кодувальником), який використовується для захоплення контексту в зображенні. Кодер - це лише традиційний стек згорткових і максимальних шарів об'єднання.

Другий шлях - це симетричний шлях розширення (також званим декодером), який використовують для забезпечення точної локалізації за допомогою транспонованих згорток.

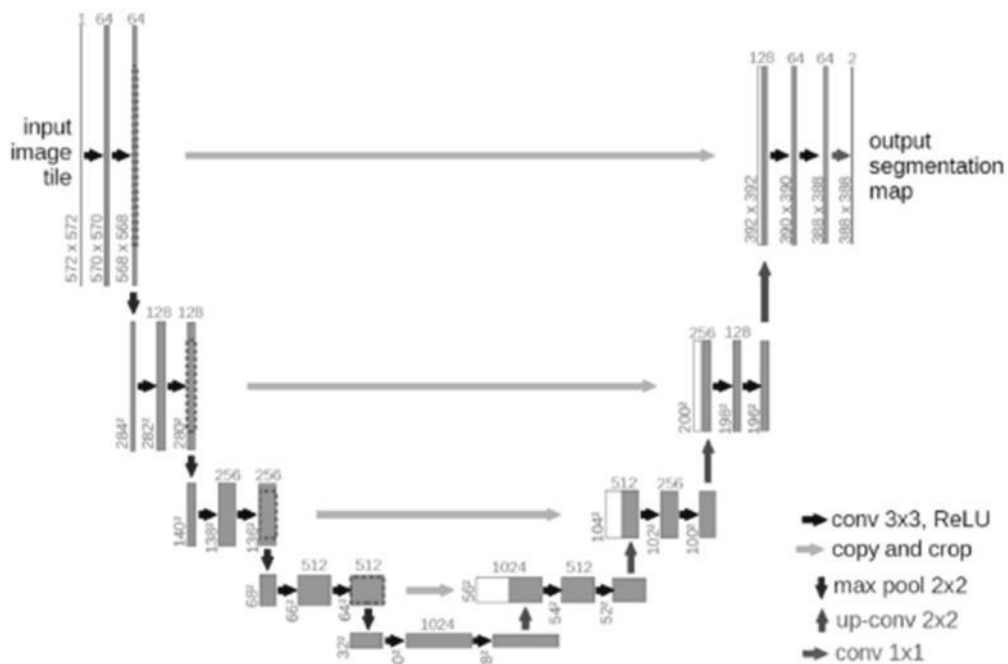


Рис. 3.5. U-net архітектура

Таким чином, це наскрізна повністю згортка (FCN), тобто вона містить лише згорткові шари і не має жодного щільного шару, через який вона може приймати зображення будь-якого розміру. Втрата перевірки досягла мінімуму на 0,0125 після 23 епох. Нижче наведено реальні та прогнозовані маски.

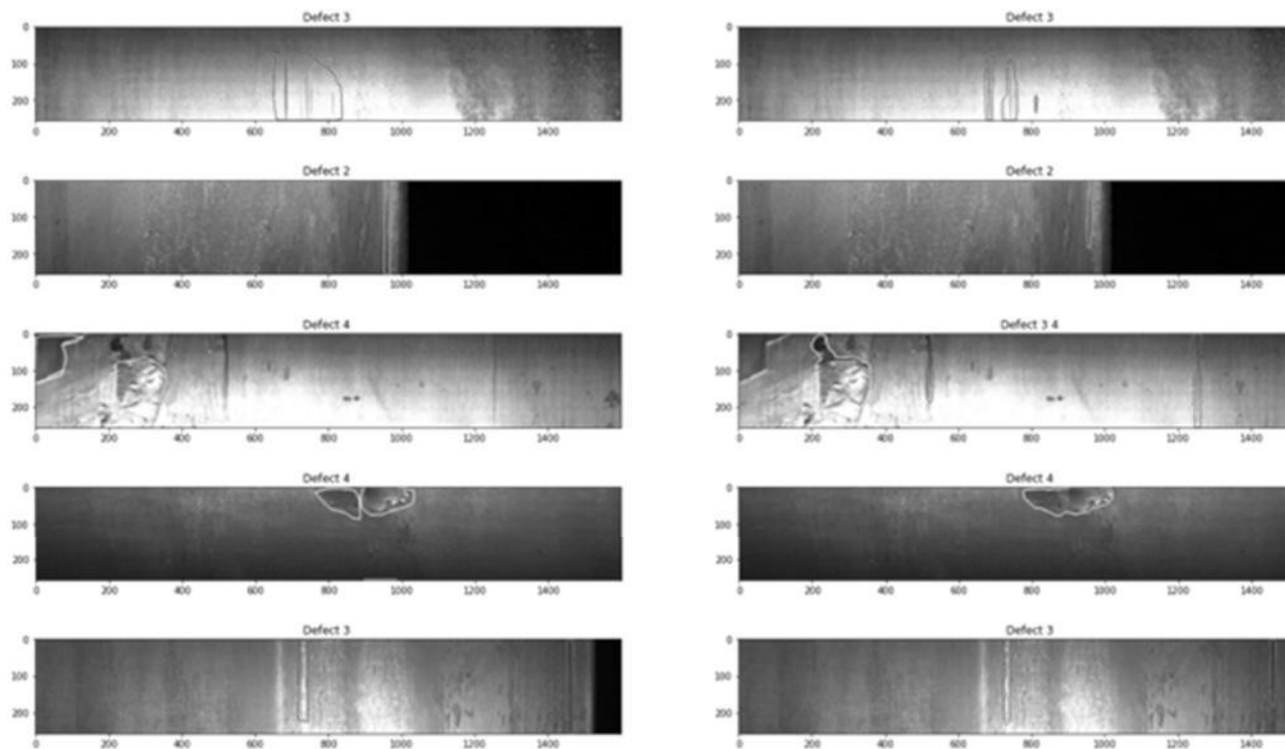


Рис. 3.6. Маски передбачення точніші, ніж маски, помічені вручну

Втрата перевірки досягла мінімуму на 0,0125 після 23 епох. Нижче наведено реальні та прогнозовані маски.

4. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1. Актуальність охорони праці

Розміщення і планування комплексу для роботизованого зварювання вимагає до себе підвищеної уваги, зокрема:

- слід передбачити спеціальні буферні зони для виробів після зварювання;
- вибираючи місце для розташування зварювального комплексу, важливо пам'ятати, що стандартні вимоги до території включають в себе якісну бетонну підлогу, товщина якої не повинна бути меншою 300 мм, з перепадами, що не перевищують 5 мм на 1000 мм.
- на території розташування роботизованого зварювального комплексу бажано спроектувати підводку сухого повітря, а при проектуванні електроживлення необхідно передбачити використання стабілізаторів.

Одним з головних завдань при розробці технічних проектів є вирішення питань по охороні праці та техніці безпеки, розробка конкретних заходів, які виключають травматизм і усувають вплив на організм робітників небезпечних і шкідливих факторів.

Згідно Закону України “Про охорону праці” [29], що був прийнятий 14 жовтня 1992 року, на будь – якому об’єкті, де працюють люди, повинні бути створені здорові і безпечні умови праці. Будівлі, обладнання, технологічні процеси не повинні створювати загрози працюючим, негативно впливати на їхній стан здоров’я та самопочуття. Для забезпечення цих вимог заходи по охороні праці повинні здійснюватися на базі нової технології і наукової організації виробництва [29]. Особливо важливими факторами полегшення і оздоровлення умов праці, підвищення його продуктивності є комплексна механізація і автоматизація робіт і технологічних процесів, застосування засобів обчислювальної техніки в наукових дослідженнях і на виробництві.

Охорона праці робітників і службовців є одним із головних обов’язків адміністрації будь-якого підприємства. Покращення умов праці, розробка і здійснення заходів по зниженню виробничого травматизму і професійної

захворюваності, крім великого соціального ефекту, дають економічні результати, які виражаються у збільшенні періоду професійної активності працюючих, зростанні продуктивності праці, зменшенні втрат пов'язаних з травматизмом, професійною і виробничою захворюваністю, зменшення текучості кадрів і зменшення затрат на пільги і компенсації. Раціональний комплекс заходів, направлених на покращення умов праці, може забезпечити приріст продуктивності праці на 15 – 20 % і більше [30].

4.2. Правила техніки безпеки при роботі в лабораторіях

У лабораторному приміщенні мають місце шкідливі і небезпечні чинники першою і четвертої групи.

Фізичні небезпечні й шкідливі виробничі чинники.

- 1) Підвищений рівень шуму, відповідно до ГОСТу 12.1.003-83 [30], джерелами якого є друкуючі пристрої і установки кондиціонування. Вплив шуму на організм людини призводить до зниження гостроти зору, зорової адаптації, порушує сприйняття зорової інформації та знижує продуктивності праці на 30-60%.
- 2) Підвищена температура довкілля, виникає й унаслідок одночасної роботи обчислювальної техніки і освітлювальних установок, які значну частину споживаної електроенергії перетворюють в теплову. Вплив цього чинника на організм людини призводить до зайвим затратам енергії нормалізацію теплового балансу, наслідком є відчуття дискомфорту й відповідне зниження працездатності.
- 3) Підвищене значення напруги в електричному ланцюгу, замикання якого може статися через людину. Це може призвести до електроудару, який впливає на організм людини біологічно (роздратування і порушення нервових і м'язових тканин) і механічно (розриви шкірного покриву, вивихи).
- 4) Недостатня освітленість робочої зони, причиною якої може бути широкий діапазон зміни і мінливість природного світла, неправильний вибір штучного освітлення і їхнього розташування. Цей чинник викликає стомлення очей,

зниження працездатності, можуть призвести до патологічному погіршення зору людини.

5) Відсутність чи недолік природного світла, причиною його можуть стати такі джерела: площа і орієнтація вікон, ступінь чистоти скла вікон, забарвлення стін приміщення. Природне освітлення має важливе значення для працюючих. З іншого боку, природне освітлення має і психологічну дію, створюючи для працюючих відчуття безпосереднього зв'язок з його оточенням.

6) Підвищений рівень електромагнітних випромінювань, джерелом якого є електронно-променеві трубки моніторів ЕОМ. Вплив цього чинника на організм людини призводить до захворювань нервової системи, раку.

7) Психофізіологічні шкідливі чинники. Джерелами фізіопсихологічного чинника є нерациональна організація робочого місця, напруженість праці.

Усі ці фактори сприяють зміни в людини функціонального стану центральної нервової системи [30], нервово-м'язового апарату рук (під час роботи з клавіатурою введення інформації):

- розумова перенапруга;
- перенапруження зорових аналізаторів;
- монотонність праці;
- емоційні перевантаження;
- статичні перевантаження.

4.3. Санітарно-гігієнічні вимоги у лабораторному приміщенні

Виробнича санітарія – система організаційних, гігієнічних і санітарно-технічних заходів і засобів [32], які запобігають дії на робітників виробничих факторів.

Виробнича санітарія включає оздоровлення повітряного середовища і нормалізацію параметрів мікроклімату в робочій зоні, захист працюючих від шуму, вібрації, ультразвуку і електромагнітних випромінювань; забезпечення потрібних нормативів природнього і штучного освітлення; підтримання у

відповідності з санітарними вимогами території підприємства, основних виробничих і допоміжних приміщень [32]. Основними санітарно-технічними системами забезпечення потрібних умов праці у даному випадку є: система вентиляції і кондиціонування повітря, система опалення, освітлення виробничих приміщень. В розливочному цеху встановлена система приточно-втяжної вентиляції. Опалення цеху парове. Подача тепла здійснюється від загальнозаводської котельні.

Вимоги до освітлення

Основні вимоги охорони праці до освітленості такі. Освітленість повинна бути достатньою і відповідати характеру зорових робіт. Освітленість повинна бути рівномірною без різких тіней. Між об'єктом розрізнення і фоном, на якому розглядається об'єкт, повинна бути деяка контрастність [32]. Джерело світла не повинно осліпляти працюючого. Рівень освітленості робочих поверхонь повинен бути постійним в часі і мати оптимальний спектральний склад світла. Електроосвітлювальні пристрої повинні бути безпечними при експлуатації.

$$P = P_{num} \cdot S_n / n = 15 \cdot 900 / 150 = 90 \text{ Вт} .$$

Протипожежні заходи

Будівлі та їх частини у яких розміщені ЕОМ, повинні мати не нижче II ступінь вогнестійкості. Пожежі помешкань з ЕОМ представляють собою особливо небезпечні ситуації, так як пов'язані з великими матеріальними втратами. Пожежа може виникнути при взаємодії горючих речовин [31].

Можливі причини виникнення пожежі у приміщенні роботи програмістів і операторів ЕОМ:

- а) недотримання правил експлуатації електронно-обчислювальної техніки;
- б) недотримання правил пожежної безпеки;
- в) перегрів розеток та частин обладнання перехідного опору у місцях сполук;
- р) несправність загального чи місцевого освітлення робочих місць.

Горючими компонентами у приміщенні є:

- а) віконні рами, двері, робочі столи;
- б) книжки, довідники-роздруковки програм, інші робочі матеріали;

в) розчинники, мастила та інші пожежонебезпечні речовини, які періодично йдуть на проведення ремонтних і профілактичних робіт на встановленому у приміщенні устаткуванні.

Джерелами виникнення пожежі є:

- а) елементи електропроводки, що потенційно можуть перегріватись внаслідок несправності устаткування чи порушення правил експлуатації устаткування;
- б) устаткування, встановлене у приміщенні.

Щоб запобігти пожежі необхідно проводити ряд технічних і організаційних заходів у відповідність до ГОСТ 12.1.004 – 91 і ГОСТ 12.4.009 – 83.

У системі запобігання пожежі передбачити:

Встановлюють 1 вогнегасник на $40 - 50 \text{ м}^2$ [31], але не менше двох у приміщенні. У приміщенні де, проводяться роботи, пов'язані зі створенням автоматизованої системи збирання та опрацювання інформації необхідно встановити 2 вогнегасника ЗУ – 5;

- а) ящик з піском обсягом $0,5 \text{ м}^3$;
- б) систему з автоматичною пожежною сигналізацією з датчиками, що реагують на поява диму ФНП – 1 з розрахунку 1 на 10 м^2 ;
- в) телефон, встановлений у легкодоступному місці.

Організаційно-технічні заходи щодо пожежної безпеки містять у собі такі положення:

- а) включення питань пожежної профілактики в усі інструктажі технічної безпеки;
- б) заборона паління в неналежному місці;
- в) призначення відповідального за пожежну безпеку;
- г) контроль ізоляції і реальний стан електропроводки з періодичністю 1 раз на рік.

ВИСНОВКИ

У кваліфікаційній роботі бакалавра представляємо модель CNN та вирішено завдання інтелектуалізованої класифікації дефектів на поверхні металопрокату сталевих листів. На відміну від існуючих методів, запропонований підхід досягає подвійної мети: вилучення ознак дефектів та дефектоскопії.

З експериментами на наборі даних поверхневих дефектів сталевих листів продемонстровано працездатність запропонованих технічних рішень. Результати експерименту показують, що з невеликим набором даних і маленькою моделлю запропонований підхід дозволяє досягти високої точності в класифікації поверхневих дефектів сталевих листів.

Користувач може вибрати певні параметри аналізу, наприклад кількість шарів,

розмір зображення тощо, залежно від компромісу між продуктивністю класифікації та обчислювальною складністю. Хоча результати були продемонстровані на гарячекатаному сталевому листі, вони запропоновані підходи можуть підійти для інших текстурованих матеріалів, таких як дерево, папір, пластик чи тканина.

ПЕРЕЛІК ПОСИЛАНЬ

1. Coates, A., Lee, H., Ng, A.Y.: An analysis of single-layer networks in unsupervised feature learning. In: Proc. of International Conference on Artificial Intelligence and Statistics (AISTATS) (2011)
2. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11), 2278–2324 (1998)
3. Fukushima, K.: Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics* 36(4), 193–202 (1980)
4. Masci, J., Meier, U., Ciresan, D., Schmidhuber, J., Fricout, G.: Steel defect classification with max-pooling convolutional neural networks. In: Proc. of International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (June 2012)
5. Westin, S.H., Li, H., Torrance, K.E.: A comparison of four BRDF models. In: Jensen, H.W., Keller, A. (eds.) Proc. of Eurographics Symposium on Rendering, pp. 1–10 (2004)
6. Herwig, J., Leßmann, S., Bürger, F., Pauli, J.: Adaptive anomaly detection within near-regular milling textures. In: Proc. International Symposium on Image and Signal Processing and Analysis, Trieste, Italy, pp. 106–111 (2013)
7. Pang, G.K.H., Chu, M.-H.: Automated optical inspection of solder paste based on 2.5D visual images. In: Proc. of International Conference on Mechatronics and Automation, pp. 982–987 (2009)
8. Hoßfeld, M., Chu, W., Adameck, M., Eich, M.: Fast 3D-vision system to classify metallic coins by their embossed topography. *Electronic Letters on Computer Vision and Image Analysis* 5(4), 47–63 (2006)
9. Ciresan, D.C., Masci, J., Meier, U., Schmidhuber, J.: Stacked convolutional autoencoders for hierarchical feature extraction. In: Proc. of International Conference on Artificial Neural Networks, ICANN (2011)
10. Bookstein, F.L.: Principal warps: Thin plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* 11(6), 567–585 (1989)
11. Ciresan, D., Meier, U., Schmidhuber, J.: Multi-column deep neural networks for image classification. In: Proc. of Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3642–3649 (June 2012)
12. Ciresan, D., Meier, U., Masci, J., Schmidhuber, J.: A committee of neural networks for traffic sign classification. In: Proc. of International Joint Conference on Neural Networks (IJCNN), pp. 1918–1921 (July 2011)

13. Goodfellow, I.J., Bulatov, Y., Ibarz, J., Arnoud, S., Shet, V.: Multi-digit number recognition from street view imagery using deep convolutional neural networks. In: Proc. of International Conference on Learning Representations (ICLR) April
14. Doherty, A., Clark, S., Care, R., Dembowsky, M.: Why rails crack. *Ingenia* (23), 23–28 (2005)
15. Huber-Mörk, R., Nölle, M., Oberhauser, A., Fischmeister, E.: Statistical rail surface classification based on 2D image analysis. In: Blanc-Talon, J., Bone, D., Philips, W., Popescu, D., Scheunders, P. (eds.) *ACIVS 2010, Part I. LNCS*, vol. 6474, pp. 50–61. Springer, Heidelberg (2010)
16. Soukup, D., Huber-Mörk, R.: Cross-channel co-occurrence matrices for robust characterization of surface disruptions in 2D rail image analysis. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P., Zemčák, P. (eds.) *ACIVS 2012. LNCS*, vol. 7517, pp. 167–177. Springer, Heidelberg (2012)
17. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. *Optical Engineering* 19(1), 139–144 (1980)
18. Basri, R., Jacobs, D., Kemelmacher, I.: Photometric stereo with general, unknown lighting. *International Journal of Computer Vision* 72(3), 239–257 (2007) CNNs for Steel Surface Defect Detection from Photometric Stereo Images
19. Arel, I., Rose, D.C., Karnowski, T.P.: Deep machine learning - a new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine* 5(4), 13–18 (2010)
20. Simard, P.Y., Steinkraus, D., Platt, J.C.: Best practices for convolutional neural networks applied to visual document analysis. In: Proc. of International Conference on Document Analysis and Recognition (ICDAR), pp. 958–963 (2003)
21. Amini, A.; Kanfoud, J.; Gan, T.-H. An Artificial-Intelligence-Driven Predictive Model for Surface Defect Detections in Medical MEMS. *Sensors* 2021, 21, 6141. <https://doi.org/10.3390/s21186141>
22. Song, L.; Li, X.; Yang, Y.; Zhu, X.; Guo, Q.; Yang, H. Detection of Micro-Defects on Metal Screw Surfaces Based on Deep Convolutional Neural Networks. *Sensors* 2018, 18, 3709. <https://doi.org/10.3390/s18113709>
23. D’Antuono, R. (2022). Basic Digital Image Acquisition, Design, Processing, Analysis, Management, and Presentation. In: Nechyporuk-Zloy, V. (eds) *Principles of Light Microscopy: From Basic to Advanced*. Springer, Cham. https://doi.org/10.1007/978-3-031-04477-9_4
24. Patar, M. N. A. A., Ayub, M. A., Zainal, N. A., Rosly, M. A., Lee, H., & Hanafusa, A. (2022). Detection of Micro-defects on Metal Screw Surfaces Based on Faster Region-Based Convolutional Neural Network. In A. N. Reddy, D. Marla, M. N. Favorskaya, & S. C. Satapathy (Eds.), *Intelligent Manufacturing and Energy Sustainability - Proceedings of ICIMES 2021* (pp.

- 587-597). (Smart Innovation, Systems and Technologies; Vol. 265). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-16-6482-3_58
25. Aldunin, A. Development of method for calculation of structure parameters of hot-rolled steel strip for sheet stamping. *J. Chem. Technol. Metall.* 2017, 52, 737–740.
 26. Xu, Z.W.; Liu, X.M.; Zhang, K. Mechanical properties prediction for hot rolled alloy steel using convolutional neural network. *IEEE Access* 2019, 7, 47068–47078.
 27. Ren, Q.; Geng, J.; Li, J. Slighter Faster R-CNN for real-time detection of steel strip surface defects. In *Proceedings of the IEEE 2018 Chinese Automation Congress (CAC), Xi'an, China, 30 November–2 December 2018*; pp. 2173–2178. [Google Scholar]
 28. He, D.; Xu, K.; Zhou, P. Defect detection of hot rolled steels with a new object detection framework called classification priority network. *Comput. Ind. Eng.* 2019, 128, 290–297. [Google Scholar] [CrossRef]
 29. Jeon, M.; Jeong, Y.S. Compact and accurate scene text detector. *Appl. Sci.* 2020, 10, 2096.
 30. Vu, T.; Van Nguyen, C.; Pham, T.X.; Luu, T.M.; Yoo, C.D. Fast and efficient image quality enhancement via desubpixel convolutional neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018*.
 31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016*; pp. 770-778.
 32. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* 2016, arXiv:1602.07360.
 33. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, UT, USA, 18–23 June 2018*; pp. 6848–6856.
 34. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* 2017, arXiv:1704.04861.
 35. Wang, Z.; Lu, W.; He, Y.; Xiong, N.; Wei, J. Re-CNN: A robust convolutional neural networks for image recognition. In *Proceedings of the International Conference on Neural Information Processing, Siem Reap, Cambodia, 13–16 December 2018*; Springer: Cham, Switzerland, 2018; pp. 385–393.
 36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 2017, 30.

37. Liu, Y.; Zhang, Y.; Wang, Y.; Hou, F.; Yuan, J.; Tian, J.; Zhang, Y.; Shi, Z.; Fan, J.; He, Z. A Survey of Visual Transformers. arXiv 2021, arXiv:2111.06091.
38. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv 2020, arXiv:2010.11929.
39. Vannocci, M.; Ritacco, A.; Castellano, A.; Galli, F.; Vannucci, M.; Iannino, V.; Colla, V. Flatness defect detection and classification in hot rolled steel strips using convolutional neural networks. In Proceedings of the International Work-Conference on Artificial Neural Networks, Gran Canaria, Spain, 12–14 June 2019; Springer: Cham, Switzerland, 2019; pp. 220–234.
40. Gao, Y.; Xiang, X.; Xiong, N.; Huang, B.; Lee, H.J.; Alrifai, R.; Jiang, X.; Fang, Z. Human action monitoring for healthcare based on deep learning. IEEE Access 2018, 6, 52277–52285.
41. Wu, C.; Ju, B.; Wu, Y.; Lin, X.; Xiong, N.; Xu, G.; Li, H.; Liang, X. UAV autonomous target search based on deep reinforcement learning in complex disaster scene. IEEE Access 2019, 7, 117227–117245.
42. Luo, Q.; He, Y. A cost-effective and automatic surface defect inspection system for hot-rolled flat steel. Robot. Comput.-Integr. Manuf. 2016, 38, 16–30.
43. Ashour, M.W.; Khalid, F.; Abdul Halin, A.; Abdullah, L.N.; Darwish, S.H. Surface defects classification of hot-rolled steel strips using multi-directional shearlet features. Arab. J. Sci. Eng. 2019, 44, 2925–2932.
44. Gong, R.; Wu, C.; Chu, M. Steel surface defect classification using multiple hyper-spheres support vector machine with additional information. Chemom. Intell. Lab. Syst. 2018, 172, 109–117.