

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Центр _____ Післядипломної освіти
(повна назва)

Кафедра _____ Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти _____ другий (магістерський)

_____ Використання згорткових нейронних мереж
_____ для вирішення задачі ідентифікації полум'я
(тема)

Виконав:
студент 2 курсу, групи _____ СШЗдм-20-1
_____ Пустовалов Д. О.
(прізвище, ініціали)

Спеціальність _____ 122 Комп'ютерні науки

(код і повна назва спеціальності)

Тип програми _____ освітньо-наукова

(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Системи штучного інтелекту
(повна назва спеціалізації)

Керівник _____ доц. каф. ШІ, Шевченко О.Ю.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ В.О. Філатов
(підпис) (прізвище, ініціали)

2022 р.

Харківський національний університет радіоелектроніки

Центр _____ Післядипломної освіти
(повна назва)

Кафедра _____ Штучного інтелекту
(повна назва)

Рівень вищої освіти _____ другий (магістерський)

Спеціальність _____ 122 Комп'ютерні науки
(код і повна назва)

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

Освітня програма _____ Системи штучного інтелекту (СШІ)
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові _____ Пустовалова Деніса Олександровича
(прізвище, ім'я, по батькові)

1. Тема роботи Використання згорткових нейронних мереж для вирішення задачі ідентифікації полум'я

затверджена наказом університету від _____ 20__ р. № _____

2. Термін подання студентом роботи до екзаменаційної комісії _____ 20__ р.

3. Вихідні дані до роботи Науково-технічні публікації, дані Інтернет та відомих проєктів, електронні документації, тестові набори даних

4. Перелік питань, що потрібно опрацювати в роботі аналіз предметної області та постановка задачі дослідження, аналіз традиційних автоматизованих систем виявлення полум'я, аналіз методів автоматизованого виділення полум'я на зображеннях, ідентифікація та сегментація полум'я на основі глибоких згорткових нейронних мереж, розроблення базової структури штучної нейронної мережі та аналіз існуючих нейромережевих моделей для розв'язання задачі класифікації об'єктів на зображеннях, застосування каскадного методу в штучних нейронних мережах, імітаційне моделювання та експериментальна оцінка розробленої моделі згорткової нейронної мережі для детектування полум'я

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Рисунок 1 – Архітектура згорткової нейронної мережі з двома операціями згортки на maxpool, Рисунок 2 – Проблемні ситуації під час детектування полум'я, Рисунок 3 – Структура каскадного методу детектування полум'я на основі згорткової нейронної мережі з пороговими функціями активації, Рисунок 4 – Приклад використання ROI-pooling для отримання ознак CNN з різними розмірами і опорними регіонами, Рисунок 5 – Архітектура розробленої нейронної мережі, Рисунок 6 – Екземпляри аналізованих зображень, Рисунок 7 – Результат застосування методу k-середніх, Рисунок 8 – Декілька екземплярів відеопослідовностей з тестової вибірки, Рисунок 9 – Декілька екземплярів зображень з навчальної вибірки, Рисунок 10 – Навчальна вибірка зображень, Рисунок 11 – Результати навчання базових нейронних мереж

-

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

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

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1.	Отримання завдання на кваліфікаційне проектування	22.03.2022	виконано
2.	Аналіз завдання та пошук літератури за темою	23.03-05.04	виконано
3.	Опрацювання літератури та аналіз об'єкту	06.04-12.04	виконано
4.	Вибір програмних засобів для розробки системи	13.04-19.04	виконано
5.	Розробка програмного засобу	20.04-03.05	виконано
6.	Аналіз отриманих результатів	04.04-06.05	виконано
7.	Оформлювання пояснювальної записки	06.05-11.05	виконано
8.	Оформлення презентаційних матеріалів	12.05.2022	виконано
9.	Представлення на рецензування	13.05.2022	виконано
10.	Представлення кваліфікаційної роботи		

Дата видачі завдання 22 березня 20 22 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис) _____
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 91 с., 22 рис., 4 табл., 2 дод, 58 джерел.

БАГАТОШАРОВИЙ ПЕРСЕПТРОН, ВИПРЯМОВУВАЧ, ГЛИБИННА НЕЙРОННА МЕРЕЖА, ЗГОРТКА, ПРИХОВАНИЙ ПРОСТІР, РОЗПІЗНАВАННЯ ОБРАЗІВ.

Об'єктом дослідження є процеси ідентифікації полум'я на основі цифрових відеопотоків, отриманих з камер спостереження в online режимі.

Предметом дослідження є нейромережева система, методи та засоби ідентифікації полум'я у відеопотоці в online режимі.

Методи дослідження. Методи обробки цифрових зображень (для вирішення задач фільтрації та класифікації цифрових зображень), функціонального аналізу, лінійної алгебри (для побудови математичних програмних моделей), нейронних мереж (для вирішення задач ідентифікації складних об'єктів), організації обчислювальних процесів із використанням обчислювальної техніки (для розроблення програмного забезпечення).

На сьогоднішній день актуальним завданням є розроблення систем відеоспостереження та методів аналізу відеопотоку в online режимі з метою виявлення ознак пожежі.

РЕФЕРАТ

Пояснительная записка: 91 с., 22 рис., 4 табл, 2 прил., 58 источников.

ВЫПРЯМИТЕЛЬ, ГЛУБОКАЯ НЕЙРОННАЯ СЕТЬ, МНОГОСЛОЙНЫЙ ПЕРСЕПТРОН, РАСПОЗНАВАНИЕ ОБРАЗОВ, СВЕРТКА, СКРЫТОЕ ПРОСТРАНСТВО.

Объектом исследования являются процессы идентификации пламени на основе цифровых видеопотоков, полученных из камер наблюдения в режиме online.

Предметом исследования является нейросетевая система, методы и средства идентификации пламени в видеопотоке в режиме online.

Методы исследования. Методы обработки цифровых изображений (для решения задач фильтрации и классификации цифровых изображений), функционального анализа, линейной алгебры (для построения математических программных моделей), нейронных сетей (для решения задач идентификации сложных объектов), организации вычислительных процессов с использованием вычислительной техники (для разработка программного обеспечения).

На сегодняшний день актуальной задачей является разработка систем видеонаблюдения и методов анализа видеопотока в режиме online с целью выявления признаков пожара.

ABSTRACT

Explanatory note: 91 p., 22 fig., 4 tabl., 2 ann., 58 sources.

CONVOLUTION, DEEP NEURAL NETWORK, HIDDEN SPACE, IMAGE RECOGNITION, MULTILAYER PERSEPTRON, RELU.

The object of research is the process of flame identification based on digital video streams received from surveillance cameras online.

The subject of the research is the neural network system, methods and means of flame identification in the video stream online.

Research methods. Methods of digital image processing (for solving problems of filtering and classification of digital images), functional analysis, linear algebra (for building mathematical software models), neural networks (for solving problems of identifying complex objects), organization of computational processes using computer technology (for software development).

To date, the urgent task is to develop video surveillance systems and methods of video flow analysis online to detect signs of fire.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів.....	9
Вступ.....	10
1 Аналіз предметної області та постановка задачі дослідження.....	13
1.1 Аналіз традиційних автоматизованих систем виявлення полум'я.....	13
1.2 Аналіз методів автоматизованого виділення полум'я на зображеннях.....	20
1.3 Колірна сегментація.....	21
1.4 Відокремлення рухомих об'єктів.....	27
1.5 Просторово-часовий аналіз полум'я.....	31
1.6 Постановка задачі дослідження.....	36
2 Ідентифікація та сегментація полум'я на основі глибоких згорткових нейронних мереж.....	40
2.1 Розроблення базової структури штучної нейронної мережі та аналіз існуючих нейромережових моделей для розв'язання задачі класифікації об'єктів на зображеннях.....	40
2.2 Застосування каскадного методу в штучних нейронних мережах.....	53
2.3 Тонке налаштування нейронних мереж (Fine tuning).....	57
3 Імітаційне моделювання та експериментальна оцінка розробленої моделі згорткової нейронної мережі для детектування полум'я.....	63
3.1 Ідентифікація полум'я на основі його кольорових ознак.....	63
3.2 Створення навчальної множини даних.....	66
3.3 Методи та критерії перевірки ефективності роботи детектора полум'я у відеопотоці.....	70
3.4 Оцінка ефективності розроблених методів.....	74
Висновки.....	81

Перелік джерел посилення.....	83
Додаток А Вихідний код програми.....	89
Додаток Б Відомість кваліфікаційної роботи.....	91

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ЗНМ – згорткова нейронна мережа;

ЛБШ – локальний бінарний шаблон;

РБФ – радіальна базисна функція;

РБФ-Г – радіальна базисна функція Гаусса;

ШНМ – штучна нейронна мережа;

NRLBP – Non-Redundant Local Binary Pattern – ненадлишковий локальний двійковий шаблон;

VLBP – Volume Local Binary Pattern – об’ємний локальний двійковий шаблон.

ВСТУП

Незважаючи на загальний технічний прогрес, в Україні відбувається значне зростання кількості пожеж і надзвичайних ситуацій. Щорічно один лише показник завданих матеріальних збитків збільшується в середньому на 6 %, що зумовлює пошуки більш досконаліших шляхів реагування на такі загрози. Протягом останніх років, з метою зменшення збитків, які можуть бути спричинені пожежею, чимало зусиль спрямовано на встановлення систем пожежної сигналізації, пожежних сповіщувачів (димових, теплових тощо). Проте їх вузька специфіка дії нерідко зумовлює ігнорування людьми вимог правил пожежної безпеки при їх встановленні.

Паралельно з цим, на підприємствах, в установах і житлових приміщеннях останніми роками активно впроваджують різноманітні охоронні системи. Сучасні вимоги до цих систем передбачають наявність елементів інтелектуального функціонування з метою розширення спектру їх прикладного використання. Дуже часто такі системи будуються на основі системи відеоспостереження, оскільки активний розвиток технологій комп'ютерного зору дозволяє будувати ефективні системи для функціонування в достатньо широкому спектрі умов. Проте, умови виникнення та проходження пожеж, такі наприклад як задимленість, періодичні чи аперіодичні мікрошуми, освітленість тощо, зазвичай дуже істотно впливають на точність функціонування таких систем, що призводить до неможливості вироблення ефективних зворотніх дій.

Незважаючи на загальний технічний прогрес, в Україні і загалом у світі спостерігається тенденція до зростання кількості пожеж та виникнення пов'язаних з ними надзвичайних ситуацій. Щорічно один лише показник завданих матеріальних збитків збільшується в середньому на 6%, що зумовлює пошуки шляхів вдосконалення процесів та засобів реагування на такі загрози.

Впродовж останніх років, з метою зменшення збитків, які можуть бути спричинені пожежами, чимало зусиль спрямовується на запровадження систем і пристроїв пожежної сигналізації, пожежних сповіщувачів різного типу. Проте, вузька спеціалізація таких пристроїв нерідко зумовлює ігнорування людьми вимог правил пожежної безпеки щодо їх встановлення та коректного використання. Паралельно з цим, на підприємствах, в установах і в приватних оселях активно впроваджуються різноманітні охоронні системи на основі інформаційних та комп'ютерних технологій, зокрема, на основі систем відеоспостереження.

Активний розвиток інформаційних технологій дозволяє розширювати функціонал таких систем та будувати ефективні системи для функціонування в достатньо широкому спектрі умов. Умови виникнення та розвитку пожеж, такі як задимленість, мікрошуми, освітленість тощо, зазвичай істотно впливають на якість функціонування таких систем, що призводить до неможливості вироблення ефективних та адекватних до конкретної ситуації дій. Тому актуальним є завдання розроблення інформаційної технології та методів для побудови систем відеоспостереження, які забезпечать автоматизацію процесу ідентифікації ознак пожежі у відеопотоці з метою мінімізації часу їх виявлення.

Однак, незважаючи на існуючі напрацювання в сфері використання систем технічного зору, проблема виявлення полум'я у відеопотоках в комплексі не є вирішеною. З однієї сторони є змога забезпечити ефективну ідентифікацію об'єктів у відеопотоці, проте, враховуючи складність таких операцій, робота програмного забезпечення в online режимі стає неможливою. З іншої ж сторони, завдяки інтенсивному розвитку інформаційних технологій, зокрема, методів та технологій машинного зору і можливостей сучасних відеокамер, на сьогоднішній день існує можливість побудови сучасних інтелектуальних інформаційних технологій та методів, які б дозволили проводити швидкий попередній аналіз відеопотоків з метою виявлення у них ознак виникнення пожежі в online режимі. Однак, існуючі

рішення характеризуються високим показником кількості хибних викликів, що є недопустимим в контексті роботи системи, яка відповідає за забезпечення життєдіяльності людини. Все це зумовлює актуальність розв'язання задачі виявлення полум'я у відеопотоці з урахуванням необхідності отримання високої точності ідентифікації об'єктів в широкому робочому діапазоні та стійкістю до негативного впливу різноманітного зашумлення.

На сьогоднішній день актуальним завданням є розроблення систем відеоспостереження та методів аналізу відеопотоку в online режимі з метою виявлення ознак пожежі.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Аналіз традиційних автоматизованих систем виявлення полум'я

Протягом останніх років, як в Україні, так і в усьому світі, спостерігається зростання можливих загроз для життя та здоров'я людей, які проявляються у вигляді аварій, катастроф та стихійних лих. Так впродовж 2019 року органами і формуваннями ДСНС забезпечене оперативне реагування на 146 класифікованих НС, що за масштабами розподілено на: державного рівня – 2, регіонального – 7, місцевого – 63, об'єктового – 74. Внаслідок цих НС загинули 199 осіб (з яких 23 дитини) і постраждали 1 тис. 492 особи (із них 624 дитини). В порівнянні із 2018 роком загальна кількість НС у 2019 році зросла на 14,1 %, при чому кількість НС техногенного характеру зросла на 25 % (внаслідок збільшення кількості НС спричинені пожежами і вибухами, аваріями на системах життєзабезпечення і раптового руйнування будівель і споруд), а чисельність НС природного характеру – на 5,2 % (внаслідок зростання у 4 рази кількості метеорологічних НС). За звітний період спостерігалось підвищення на 78 % кількості постраждалих (за рахунок медико-біологічних НС) і на 18,5 % кількості загиблих у НС (за рахунок НС унаслідок аварій на транспорті і пожеж, вибухів в будівлях та спорудах) [1].

З метою мінімізації можливих збитків, була проведена значна робота щодо дослідження методів та засобів, які б дозволили протидіяти виникненню або оперативно реагувати на прояви різноманітних надзвичайних ситуацій. Проте, незважаючи на значні зусилля науковців в даному напрямку та позитивну тенденцію щодо зниження кількісних показників як загиблих так і НС в цілому (у 2016 році зареєстровано найменшу кількість загиблих в НС та найменшу кількість НС техногенного характеру за період спостережень 1997–2019 років [2]), дане питання все ще

залишається актуальним, оскільки спостерігається зростання кількості матеріальних збитків спричинених надзвичайною ситуацією та травмування людей.

Аналіз характеру причин виникнення техногенних надзвичайних ситуацій показує, що переважна більшість НС (63%) спричинена внаслідок пожеж та вибухів [2]. В даних умовах виникають небезпечні фактори для життя та здоров'я людей, такі як високі температури (може призвести до опіків різного ступеню) та наявність в продуктах горіння хімічно небезпечних речовин (можуть призвести до отруєння різного ступеню). Стосовно економічних збитків, які спричинені пожежею чи вибухами, то необхідно відмітити, що відповідно до аналізу масиву карток обліку пожеж за 12 місяців 2020 р., кожною пожежею державі наносились прямі збитки на суму 21,9 тис. грн., а загальна сума матеріальних втрат склала 5 млрд. 166 млн. 872 тис. грн. (з них прямі матеріальні збитки становлять 1 млрд. 626 млн. 531 тис. грн. а побічні 3 млрд. 540 млн. 341 тис. грн.) [2].

Небезпечність виникнення і подальшого розповсюдження пожеж та вибухів обумовлена також тим фактом, що за умов відсутності будь-яких дій, спрямованих на локалізацію та ліквідацію НС, на початковому етапі розвитку (перші 10 – 20 хв.), буде проходити об'ємне розповсюдження вогню (за умов наявності горючих матеріалів та кисню), тим самим зробивши неможливим гасіння пожежі підручними засобами. Паралельно з цим, буде відбуватися заповнення приміщення димом, що в значній мірі ускладнить процеси ліквідації пожежі та евакуації людей. Відповідно, пожежа є однією з найнебезпечніших видів НС, оскільки в часовому проміжку часі, процес її локалізації та ліквідації зазвичай ускладнюється, вимагаючи дедалі більше часових та людських ресурсів. Це в свою чергу обумовлює необхідність вияву ознак пожежі на ранніх етапах та сповіщення про них відповідних органів та підрозділів, з метою прийняття ефективних рішень щодо мінімізації можливих втрат.

Наявність або відсутність процесу горіння може бути з легкістю визначено людиною на основі таких ознак як підвищення температури, візуальний вигляд вогню та/або диму, різкого запаху. Саме тому, ряд пожежних сповіщувачів працює в ручному режимі роботи. Проте, людина нездатна здійснювати цілодобовий моніторинг усіх приміщень будівлі, оскільки в даному випадку відіграють такі елементарні людські фактори як втома, неуважність тощо. У зв'язку з цим, загальноприйнятою є практика використання саме автоматичних пожежних сповіщувачів (АПС).

Відповідно до ДСТУ 2273:2006 «Протипожежна техніка. Терміни та визначення основних понять» автоматичним пожежним сповіщувачем – є сповіщувач, який виконує функції виявлення ознак горіння та оповіщення про виникнення пожежі незалежно від зовнішніх сигналів керування і джерел живлення [3].

Згідно із статистичними даними, наведеними Українським науково–дослідним інститутом цивільного захисту, системами автоматичного протипожежного захисту у 2020 році обладнано 380 тис. 587 об'єктів, що складає 93,4% їх необхідної кількості [1]. Проте, із загальної кількості систем протипожежного захисту 7,9 % зіпсовано, 12,7 % відпрацювало свій технічний ресурс та підлягає заміні, а з працюючих установок лише 79,8% проходять планове технічне обслуговування [1]. В цілому дані показники продовжують зменшуватися, оскільки проблема якісного технічного утримання систем пожежної автоматики вимагає значної кількості як часових так і матеріальних ресурсів.

Зазначена тенденція до збільшення кількості непрацюючих АПС є неприпустимою, оскільки це в першу чергу може загрожувати життю та здоров'ю людини. Дане твердження знаходить відображення в інформаційно-аналітичній довідці асоціації «National Fire Protection Association», де зазначається, що рівень смертності на 100 зафіксованих пожежах більш ніж вдвічі вищий в будинках які не містили будь-яких

працюючих детекторів диму, аніж в будинках із функціонуючими давачами (1.18 смертей проти 0.53 на 100 чоловік) [5].

Відповідно, постає питання пошуку найбільш універсального засобу моніторингу пожежного стану, який б з однієї сторони дозволив на якісному рівні виконувати поставлені вимоги щодо виявлення вогню на контрольованій ділянці, а з іншої – усунув існуючі недоліки АПС.

Що стосується основних вимог яким повинні відповідати АПС, то вони повинні передусім:

- максимально швидко виявляти фактори пожежі;
- характеризуватися достовірністю отриманої інформації, що означає необхідність звести до мінімуму можливість хибних спрацювань;
- забезпечувати надійний та оперативний зв'язок між пожежною сигналізацією та приймально-контрольним приладом;
- попереджати про можливі несправності в роботі системи.

Стосовно їх недоліків, то в даному контексті необхідно зауважити, що незважаючи на їх однакове функціональне призначення (виявлення ознак пожежі), відокремлюють давачі різноманітного типу. Кожен з них характеризується як певними перевагами так і певними недоліками, що і зумовлює розгляд кожного типу пожежного сповіщувача окремо. Відповідно, вважається за доцільними здійснити їх умовну класифікацію за видом контрольованої ознаки пожежі на: теплові, димові, полум'я та комбіновані.

Теплові пожежні сповіщувачі можуть реагувати на перевищення певного граничного значення температури або на швидке зростання значення температури в контрольованій зоні. До їх переваг необхідно віднести високу стійкість до хибно-позитивних спрацювань в умовах із швидким коливанням температури навколишнього середовища, можливість визначити загоряння матеріалів, які не супроводжуються виділенням диму. Що стосується недоліків, то такі сповіщувачі характеризуються високою

інерційністю. Так, теплові сповіщувачі не встановлюють в приміщеннях з високими стелями.

Пожежні сповіщувачі полум'я реагують на інфрачервону та ультрафіолетову складову спектра випромінювання, що створюється при горінні різних матеріалів. Сповіщувачі даного типу дозволяють значно швидше реагувати, в порівнянні з іншими пожежними сповіщувачами. Зокрема, на протигагу тепловим сповіщувачам, відстань до пожежі істотно не впливає на час виявлення пожежі у приміщенні. З негативної сторони необхідно відмітити неможливість спрацювання сповіщувачів з ультрафіолетовими сенсорами на тліючих пожежах, оскільки виділяється значне задимлення, що тим самим призводить до розсіювання світлового випромінювання осередку пожежі. Що стосується пожежних сповіщувачів полум'я із інфрачервоними сенсорами, то у них існує проблема забезпечення перешкодозахищеності від прямого та відбитого випромінювання різних джерел освітлення, випромінювання нагрітих частин обладнання, грозових розрядів, технологічного та ремонтного устаткування.

Димові пожежні сповіщувачі дозволяють виявляти пожежу на початковій стадії, оскільки процес горіння зазвичай супроводжується виділенням диму. Зокрема, процес тління в закритих приміщеннях, за умов недостатньої кількості кисню, може тривати декілька годин, тим самим повністю заповнюючи приміщення димом, що може призвести до смертельних наслідків.

Існуючі димові сповіщувачі можна розділити на дві групи, відповідно до методу який вони використовують:

- оптично-електронний;
- радіоізотопний.

Оптично-електронний дозволяє приймати рішення про наявність диму в приміщенні на основі одного з двох критеріїв: перевищення допустимого значення світлового потоку на контрольованій ділянці або

значення світлового потоку, який пройшов через контрольовану ділянку, менше допустимого значення.

Радіоізотопні, в порівнянні з оптично-електронними димовими сповіщувачами, дозволяють реагувати на частинки диму малого розміру (0,1-1,0 мкм).

Стосовно недоліків використання даних сповіщувачів необхідно відмітити наступні: неможливість виявлення процесу горіння деяких речовин, які не виділяють дим, спостерігаються хибно-позитивні виклики від пилу чи комах, низька стійкість до електромагнітних перешкод.

З вищезазначеного можемо зробити висновок, що для ефективного реагування на пожежу засобами пожежних давачів, необхідно використовувати не лише одну їх певну різновидність, а комбіноване поєднання, оскільки лише за даних умов можливо забезпечити високі показники рівня захисту від пожеж та достовірності отримуваної інформації від сповіщувачів. Проте, за умов реалізації такого підходу, можливе виникнення наступних проблем: складність в монтажу та технічного обслуговування та висока вартість обладнання.

Зокрема, необхідно відмітити, що згідно статистики, наведеної асоціацією «National Fire Protection Association» (рисунок 1.1), на об'єктах в яких були зафіксовані пожежі і містилися димові сповіщувачі, лише в 53% відбулося коректне спрацювання сигналізації, в 12% пожежа була занадто малою для її фіксації, в 27% були відсутні димові сповіщувачі, а в 8% сповіщувач взагалі не зреагував на наявність пожежі.

Що стосується зафіксованих пожеж, які призвели до загибелі людей на об'єктах, які були оснащені димовими сповіщувачами, то показники виглядають ще трагічніше. Так, лише в 40% відбулося коректне спрацювання, в 21% димовий сповіщувач не зреагував на пожежу, а в 38% випадків сповіщувачі даного типу були взагалі відсутні.

Reported home structure fires and fire deaths by smoke alarm performance
2009-2013

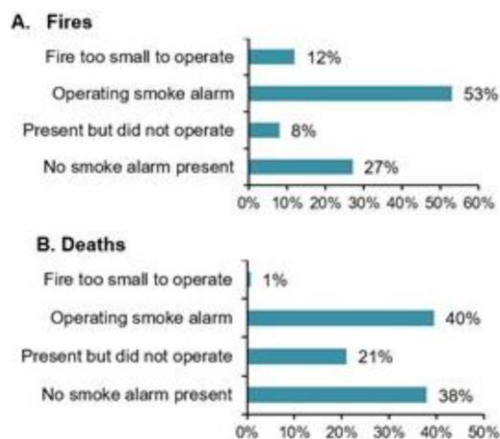


Рисунок 1.1 – Результати роботи димових сповіщувачів.

Вищенаведена статистика дозволяє припустити, що низький рівень встановлених сповіщувачів або їх підтримка в робочому стані, зумовлена їх вузькою спрямованістю і порівняно низькою можливістю виникнення пожежі. На противагу зазначеним проблемам, пропонується використання засобів відеоспостереження, які в порівнянні з традиційними автоматичними пожежними сповіщувачами, мають ряд переваг: можливість виявлення полум'я на відкритій місцевості; можливість контролю приміщень великих об'ємів мінімальною кількістю відеокамер, оскільки їх розміщення, на відміну від деяких типів традиційних сповіщувачів, не повинно знаходитися поблизу джерела займання; можливість перевірки правильності та ідентифікації безпосереднього місця загоряння; мінімально короткий час реагування на прояви пожежі; полегшення виявлення місця та причини виникнення пожежі.

Необхідно також відмітити про можливу економічну перевагу при впровадженні системи ідентифікації полум'я шляхом відеоаналітики. Передусім, даний підхід дозволяє в собі поєднати системи охоронного відеонагляду і пожежної сигналізації в єдине комплексне рішення візуального контролю об'єкту. В даному контексті необхідно згадати про той факт, що практично кожен об'єкт з масовим перебуванням людей

обладнаний системами відеонагляду, що мінімізує можливі затрати на їх монтаж.

1.2 Аналіз методів автоматизованого виділення полум'я на зображеннях

Відповідно до поставлених задач, основні зусилля передусім були спрямовані на визначення процесу горіння шляхом детектування полум'я на зображенні або на наборі послідовних кадрів за певний часовий проміжок (відеопотоку), що передбачає виокремлення візуальних характеристик, які б дозволили його описати.

Так, вогонь являє собою постійно видозмінюваний в русі об'єкт, який можливо описати лише за умов знання всіх факторів як зовнішнього середовища (напрямок та швидкість вітру, вміст кисню тощо) так і безпосередньо речовини яка горить (температура спалаху, швидкість горіння тощо). Відповідно до типу розроблюваної системи детектування полум'я, а також з урахуванням її необхідного параметру – універсальності, врахування набору усіх параметрів які б дозволили змоделювати вогонь не являється можливим.

З іншої ж сторони, вогонь як об'єкт спостереження характеризується загальними візуальними ознаками, які дозволяють відокремити його від об'єктів іншого класу.

Так, процес горіння в загальному вигляді нагадує окислювально-відновну реакцію, основною відмінністю яких є швидкість протікання даних процесів. Під час горіння, теплота виробляється настільки швидко, що не встигає розсіюватися. Це, в свою чергу, супроводжується значним зростанням температури та характерним яскравим світінням. Характерним явищем під час горіння є прояв диму, який являє собою аерозоль повного або неповного згоряння речовин. Звичайно ж необхідно відзначити, що наприклад під час тління відсутнє полум'я, проте супроводжується великою

кількістю диму, а під час горіння деяких речовин (наприклад водень) диму взагалі не спостерігається.

Відповідно, найбільш ефективніший спосіб який дозволяє описати полум'я за його візуальною характеристикою є його колір, який можна описати на основі певних граничних меж колірної моделі/набору моделей або здійснити просторовий аналіз з метою більш детальнішого опису зв'язку сусідніх пікселів. Оскільки вогонь характеризується постійною зміною форми та границь в часовому проміжку то за умови наявності відеоряду існує змога мінімізації кількості хибно сегментованих ділянок шляхом виокремлення лише рухомих об'єктів або додаткового порівняння поведінки полум'я в часі з детектованими об'єктами.

Додатково для виявлення вогню можна використати його просторово-часові характеристики такі як мерехтіння полум'я, просторова яскравість меж вогню (центральна частина вогню характеризується яскраво-білим кольором, ближче до границі полум'я колір виступає більше в жовтому, червоному та оранжевих спектрах) тощо.

Відповідно, наукові роботи присвячені даній тематиці було умовно розділено на 3 групи: колірна сегментація, виявлення рухомих об'єктів та просторово-часовий аналіз, що дає змогу більш детальніше дослідити кожену групу методів.

1.3 Колірна сегментація

Загальним підходом до виявлення вогню на зображенні є використання різноманітних колірних моделей. Для цього виконується інтерактивна сегментація області зображення, в області спостереження полум'я та, відповідно до отриманих статистичних розподілів пікселів, формуються граничні умови, в межах яких визначається належність пікселя до шуканої області.

Найбільш вживана модель, яку використовують для визначення наявності полум'я на зображенні є модель RGB, оскільки зазвичай зображення передається в даній колірній моделі. R, G та B – являють собою три різні складові каналу (червоний, зелений та синій відповідно), змішування яких надає різноманітні кольори.

На зображеннях які містять вогонь складова R повинна набувати більшого значення в порівнянні з іншими, що можна описати наступним чином $R > G$ і $R > B$. В свою чергу, компонента R повинна перевищувати певне усереднене значення (R_{mean}). Комбінація даних правил може бути відображена наступними виразами (1.1):

$$R > G > B \quad (1.1)$$

$$R > R_{mean}$$

Даний підхід знайшов своє відображення у роботі [6], де автори визначили наступні правила для перевірки приналежності об'єкту до категорії полум'я (1.2):

$$R_1(x, y) = 1, \text{ if } R(x, y) > G(x, y) \&\& G(x, y) > B(x, y) \\ 0, \text{ інакше} \quad (1.2)$$

$$R_1(x, y) = 1, \text{ if } R(x, y) > 190 \&\& G(x, y) > 100 \&\& B(x, y) > 140 \\ 0, \text{ інакше}$$

Користуючись вищенаведеним набором правил, у роботі [7] були введені значення в яких може коливатися полум'я відповідно до умов спостереження. Результати виявлених значень наведені у таблиці 1.1.

Таблиця 1.1 – Граничні значення для виявлення полум'я на зображенні в залежності від умов спостереження

Освітлення	R	G	B
Ранок	>230	>160	>50
Обід	>180	>70	>100
Ніч	>230	>100	>40

У роботі [6] автори запропонували дещо інший набір правил для визначення пікселів вогню. Оскільки освітлення заднього фону може спричинити коливання значення насиченості пікселів полум'я, було запропоновано введення додаткової величини S_{mean} , яка відповідає за насиченість фону і буде зменшуватися із зростанням компоненти R . Повний набір правил можна викласти в наступному вигляді (1.3):

$$R_1(x, y) = 1, \text{ if } R(x, y) > R_{mean} \\ 0, \text{ інакше}$$

$$R_2(x, y) = 1, \text{ if } R(x, y) > G(x, y) \&\& G(x, y) > B(x, y) \\ 0, \text{ інакше} \quad (1.3)$$

$$R_3(x, y) = 1, \text{ if } S(x, y) \geq ((255 - R(x, y)) \times S_{mean}/R_{mean}) \\ 0, \text{ інакше}$$

В роботі [7], Т. Selik на заміну фіксованим значенням граничних меж червоного кольору використовує статистичну колірну модель. Базова колірна модель пожежі описується на основі статистичного аналізу ділянок зображення які містять вогонь (1.4).

$$R_{\text{mean}} = \frac{1}{k} \sum_{i=1}^k R(x_i, y_i), \quad (1.4)$$

де K – загальна кількість пікселів зображення.

В даній роботі автор знову ж звертає увагу на проблему різного освітлення і пропонує використовувати набір правил, який ґрунтується на основі співвідношення каналів R-G, R-B та G-B (1.5):

$$\begin{aligned} 0,25 &\leq G(x, y)/(R(x, y) + 1) \leq 0,65 \\ 0,05 &\leq B(x, y)/(R(x, y) + 1) \leq 0,45 \\ 0,20 &\leq B(x, y)/(G(x, y) + 1) \leq 0,60 \end{aligned} \quad (1.5)$$

Як недолік даного метода необхідно відмітити використання шість евристичних меж для заміни значення компоненти яскравості колірному простору RGB певним співвідношенням колірному каналу, що не робить його повністю самостійним.

У роботі [8] автор поєднує колірну модель RGB разом з каналом насиченості (S) моделі HSV. Дана процедура ґрунтується на закономірному припущенні, що пікселі полум'я мають більш інтенсивніший червоний колір аніж зелений і, відповідно, зелений інтенсивніший аніж синій (1.6):

$$S \geq \left((255 - R) \frac{S_T}{R_T} \right), \quad (1.6)$$

де порогові значення для S_T та R_T коливаються в районі від 55 до 65 та від 115 до 135 відповідно.

З метою визначення пікселів полум'я також часто використовують комбінацію з колірною моделлю YCbCr [52], [9], [10], де Y – компонента яскравості, а Cb та Cr – синя та червона колірні компоненти відповідно.

Вважається за цікавою комбінація колірних моделей L*a*b та YCbCr у роботі [11]. Оскільки модель L*a*b містить в собі кольори які не можуть бути сприйняті людиною, то на початку здійснюється конвертація моделі RGB в CIE XYZ, а з CIE XYZ уже в L*a*b. Після чого, до отриманої колірної моделі застосовується метод К-середніх (K-means) квадратів для прийняття рішення щодо віднесення пікселів до категорії «полум'я».

Що стосується моделі YCbCr, то в даному випадку було використано метод статистичного розподілу, який дозволив виокремити наступне правило (1.7):

$$\begin{cases} 1, \text{ якщо } f(\text{Cr}(x, y), \text{Cb}(x, y)) > T \\ 0, \quad \text{в протилежному випадку} \end{cases}$$

$$f(\text{Cr}, \text{Cb}) = \frac{1}{2\pi S_{\text{Cr}} S_{\text{Cb}}} \times \exp \left\{ \frac{1}{2} \left[\left(\frac{\text{Cr} - \text{Cr}_m}{S_{\text{Cr}}} \right)^2 - 2 \frac{(\text{Cr} - \text{Cr}_m)(\text{Cb} - \text{Cb}_m)}{S_{\text{Cr}} S_{\text{Cb}}} + \left(\frac{\text{Cb} - \text{Cb}_m}{S_{\text{Cb}}} \right)^2 \right] \right\} \quad (1.7)$$

де Cr_m , Cb_m , S_{Cr} , S_{Cb} – розходження та відхилення відповідно до чергового червоного та синього каналів.

Що стосується значень Cr_m , Cb_m та T , то статистично визначено що вони повинні дорівнювати 0.246431, 0.682332 та $3e-5$ відповідно.

У роботі [12] для визначення пікселів, які можуть візуально нагадувати полум'я було запропоновано використовувати колірну модель HSI. Трансформація може бути здійснена за формулою (1.8):

$$H = \begin{cases} \theta & \text{якщо } B \leq G \\ 360 - \theta & \text{якщо } B > G \end{cases}$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{[(R - G)^2 + (R - G)(R - B)]^{\frac{1}{2}}} \right\}$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

$$I = \frac{1}{3}(R + G + B)$$
(1.8)

За цікавим видається підхід до сегментації ділянок зображення на яких міститься полум'я. Для цього використовується алгоритм, який дозволяє порівняти дистанцію між пікселями з граничними значеннями (1.9).

$$D(x, m) = \|x - m\| = \{(x - m)^T(x - m)\}^{\frac{1}{2}} =$$

$$= [(x_R - m_R)^2 + (x_G - m_G)^2 + (x_B - m_B)^2]^{\frac{1}{2}}$$
(1.9)

Відповідно, якщо $D(x, m) \leq$ граничного значення, то ділянка належить до категорії «полум'я», інакше – вона рахується як задній фон. Регіони, які необхідно сегментувати детектуються на основі результату роботи оператора Собеля.

У роботах [13], [14] для сегментації зображення використовується колірна модель YUV. Базова методика знаходження пікселів, які б дозволили описати пожежу, базується на аналізі компоненти U або порівнянні компонент U та V , що може бути описане наступним чином (1.10):

$$\begin{cases} 1, & \text{якщо } |U(x, y) - 128| > T_U \text{ або } |U(x, y) - V(x, y)| > T_{UV} \\ 0, & \text{в іншому випадку} \end{cases} \quad (1.10)$$

1.4 Відокремлення рухомих об'єктів

Для надання коректної оцінки щодо наявності полум'я на зображенні, за доцільним буде аналіз його руху, оскільки вогонь є не статичним об'єктом. На ранньому етапі процесу горіння полум'я буде постійно зростати за умов наявності горючих речовин та достатньої концентрації окисника. Відповідно до цього, ми можемо описати процес зміни полум'я в часі наступною формулою (1.11) [15]:

$$P_i = \frac{[A_{i+1} - A_i]}{A_i} \quad (1.11)$$

де A_i – можлива ділянка з полум'ям на кадрі i ;

A_{i+1} – можлива ділянка з полум'ям на кадрі $i+1$.

З метою виявлення рухомих об'єктів у відеопотоці широко використовують різноманітні методи, які б дозволили виявити можливі зміни на основі просторових або часових ознак. Умовно, серед існуючих методів можна виділити чотири найбільш використовуваних та ефективних підходів: вилучення заднього фону, статистичний аналіз, різниця кадрів та оптичний потік.

Метод різниці кадрів є одним із найбільш вживаних методів оскільки характеризується простотою реалізації та високою швидкістю обчислення. Алгоритм даного методу базується на використанні двох або трьох послідовних кадрів з відеопослідовності, на основі чого розраховується їх різниця відповідно до кожного пікселя. Стосовно недоліків такого підходу необхідно відзначити низьку ефективність виділення повної форми об'єкта, особливо його внутрішньої частини.

Ключовим моментом під час використання методу вилучення заднього фону є моделювання безпосереднього фону, що звичайно ж можливе лише за умов використання стаціонарних відеокамер. В зазначеному методі фоном буде вважатися піксель або група пікселів які протягом певного проміжку часу (проміжок часу може бути довільним) не характеризувалися будь-якою зміною в часі та просторі. Після того як фон буде сформований, наявність рухомих об'єктів визначається на основі піксельної різниці між попереднім та теперішнім кадрами відповідно до заданого порогу (T), що дозволяє сформувати двійкову маску (1.12):

$$m_t(x, y) = \begin{cases} 0, & d_t(x, y) < T \\ 1, & d_t(x, y) > T \end{cases} \quad (1.12)$$

де $d_t(x, y)$ – отримана різниця між теперішнім та попереднім кадром.

Серед основних недоліків застосування вищезазначеного методу необхідно відмітити наступні: вразливість до незначних динамічних змін фону (зміна освітлення, періодичні мікрорухи, зашумлення тощо), що ускладнює його використання особливо поза приміщеннями і вимагає постійного оновлення моделі фону; необхідність використання лише стаціонарних камер відеоспостереження.

З метою вирішення проблеми необхідності врахування динамічних змін фону широко використовують методи статистичного опису, які дозволяють описати статистичні характеристики індивідуальних пікселів з метою врахування їх особливостей зміни. Виявлення рухомих пікселів відбувається на основі порівняння статистичних даних кожного пікселя з статистично вдосконаленою моделлю заднього фону.

Незважаючи на можливість вирішення однієї із зазначених проблем методу заднього фону, проблема необхідності використання лише стаціонарних камер відеоспостереження все ще лишається актуальною.

Зокрема, до недоліків даного методу можна віднести складність врахування та опису можливих періодичних змін заднього фону.

Метод оптичного потоку дозволяє описати зображення видимого руху об'єктів на основі переміщення кожної точки між двома площинами зображеннями. Дана функція дозволяє оцінити не тільки вектор руху пікселів зображення, а й визначити швидкість кожного пікселя зображення між суміжними зображеннями для більш досконалішого аналізу.

Даний метод базується на знаходженні зміщень (dx , dy) для кожної точки зображення $I_1(x, y)$ щоб вихідній точці відповідала точка на іншому зображенні $I_2(x + dx, y + dy)$. Для знаходження таких відповідностей зазвичай використовують концепцію стабільності яскравості, яка передбачає, що за умов порівняно невеликого руху та рівномірних умов освітлення, точка зберігатиме стає значення яскравості протягом певного періоду.

Метод оптичного потоку, в порівнянні з іншими методами виявлення рухомих об'єктів, характеризується значною перевагою, яка полягає в можливості його використання навіть за умов зміни руху камери чи заднього фону. Однак, за своє обчислювальною спроможністю, вказаний метод є надзвичайно складним і вимагає використання значної кількості обчислювальних ресурсів, що в умовах реального відеоспостереження є не завжди можливим.

Вищевказані методи широко використовуються дослідниками при розробленні систем детектування полум'я. Так, наприклад у [18], метод ідентифікації передбачає використання оптичного потоку та нейронної мережі зворотного поширення для обчислення особливостей руху полум'я на основі таких статистичних характеристик: середнє значення швидкості та швидкість зміни оптичного потоку, а також його напрямку.

Однак, використання методу оптичного потоку, як уже було вищезазначено, є процесом, який потребує використання значних

комп'ютерних ресурсів, що певною мірою ускладнить процес інтеграції таких систем на реальних об'єктах.

Дещо інший підхід запропоновано авторами у [18]. В даній роботі передбачається ідентифікація полум'я за допомогою трьох послідовних етапів: виділення регіону для аналізу, класифікація вогню на основі колірно-текстурних ознак та часової ідентифікації. Проте, знову ж враховуючи різноманітність процесу горіння, на наш погляд, урахування усіх можливостей може призвести до перенавчання класифікатора та, як наслідок, значного зниження ефективності.

Для виявлення руху у відеопотоці широко використовують методи різниці кадрів або виділення заднього фону. Як недолік необхідно відзначити, що дані методи обмежені необхідністю використовувати лише стаціонарні камери.

Так, у роботі [19] пропонується поєднати дані методи. Для цього на початковому етапі розраховується різниця кадрів на основі двох вхідних кадрів (i та $i-1$). Паралельно з цим, генерується задній фон шляхом порівняння поточного кадру із фоновим зображенням, яке зберігається в буфері. Метод виділення заднього фону використовується як спосіб отримання основної інформації про рухомі пікселі на зображенні. На наступному етапі, на основі інформації отриманої шляхом порівняння попередніх кадрів, формується висновок про належність тих пікселів, які протягом певного проміжку часу не здійснювали ніякого переміщення як таких, які можна віднести до заднього фону. На прикінцевому етапі використовується сукупна інформація отримана з двох методів і формується висновок щодо рухомих пікселів зображення.

Різниця кадрів між поточним кадром $L^*(x, y, t)$ в часовому проміжку t та попереднім кадром $L^*(x, y, t-1)$ може бути описана наступним чином (1.13):

$$FD(x, y, t) = \begin{cases} 1, & \text{якщо } |L^*(x, y, t) - L^*(x, y, t - 1)| \geq T_{FD} \\ 0, & \text{в іншому випадку} \end{cases} \quad (1.13)$$

де T_{FD} – граничне значення.

Схожий підхід використовується при визначенні заднього фону, що передбачає порівняння поточного кадру $L^*(x, y, t)$ із зображенням заднього фону $BG(x, y, t-1)$ (1.14).

$$BD(x, y, t) = \begin{cases} 1, & \text{якщо } |L^*(x, y, t) - BG(x, y, t - 1)| \geq T_{BD} \\ 0, & \text{в іншому випадку} \end{cases} \quad (1.14)$$

Розглянуті методи є базовими для виявлення рухомих об'єктів у відеопотоці. І хоча існує ряд більш складніших методів які можуть використовуватися з метою надання оцінки щодо напрямку руху полум'я, їх використання не є доцільним, оскільки за важливим вважається передусім виявлення рухомих пікселів в режимі реального часу.

Подальший аналіз отриманої інформації здійснюватиметься на основі інших методів, які не вимагають додаткової інформації про рух об'єкта.

1.5 Просторово-часовий аналіз полум'я

Однією із ключових характеристик, яка дозволяє розрізнити полум'я від інших візуально схожих об'єктів є його мерехтіння, наявність якого може бути визначене на основі його просторово-часової поведінки.

Так, у роботі [22] пропонується аналіз граничних меж вогню на частоту появлення/зникання пікселів, колір яких може нагадувати полум'я. Зокрема, для визначення наявності мерехтіння у відеоряді, використовують 2 базових принципи:

- вогонь має здатність до мерехтіння з частотою від 1 до 10 Гц;
- значення яскравості у полум'я наближене до максимуму.

Відповідно до вищенаведеного, визначити наявність мерехтіння у полум'ї можна на основі алгоритму, який представлений у таблиці 1.2:

Таблиця 1.2 – Алгоритм знаходження мерехтіння полум'я

Крок	Стислий опис
1. Перетворення моделі RGB до моделі YUV	Конвертація кожного кадру $F_i (1 \leq i \leq N)$ відеопотоку в кольорову модель YUV, де: N – загальна кількість кадрів; $Y_i(x, y)$ – яскравість кольорової компоненти пікселю.
2. Для кожної координати (x, y) $F_i (1 \leq i \leq N)$: – визначається значення похідної яскравості;	$D_i(x, y) = Y_i(x, y) - Y_{i-1}(x, y) $
– визначається значення кумулятивної похідної часу	$CT_i(x, y) = \alpha CT_{i-1}(x, y) + (1 - \alpha)w_i(x, y)D_i$ де α – кумулятивний ефект, який прирівнюється до $(N - 1)/N$; w_i – вагове значення яскравості i може обчислюватись за формулою: $w_i(x, y) = \begin{cases} Y_i(x, y), & \text{якщо } Y_i(x, y) \geq \delta \\ 0, & \text{у всіх інших випадках} \end{cases}$ $CT_i(x, y)$ – апіорі дорівнює 0.
– крок нормалізації даних	$CT_i(x, y)$ перетворюється у 8-бітний діапазон яскравості ($0 \leq CT_i(x, y) \leq 255$) та знаходиться середнє значення (Ave).
– створення маски	В масці знаходяться всі пікселі після кроку нормалізації $CT_i(x, y) > Ave$.

Стосовно сегментації зображення на основі кольору, то в даній роботі окрім використання базових співвідношень між червоною, синьою та зеленою компонентами передбачається використання значень інтенсивності та насиченості в колірній моделі HSV.

Так, на початковому етапі передбачається переведення кожного кадру відеопотоку з колірної моделі RGB в HSV, де S і V – значення насиченості та інтенсивності відповідно.

Після чого, до кожної позиції (x, y) кадру F ($2 \leq i \leq N$) проводяться i наступні операції:

– створюється колірна маска полум'я (FCM) (1.15):

$$FCM_i(x, y) \begin{cases} 1, \text{ якщо } R_i(x, y) > \eta \cup R_i > G_i(x, y) \cup G_i(x, y) > B_i(x, y) \\ 0, \quad \text{в іншому випадку,} \end{cases} \quad (1.15)$$

де η експериментально було вирішено прирівняти до 180;

– створити нову матрицю насиченості (1.16):

$$S_i'(x, y) = \begin{cases} S_i(x, y), & \text{якщо } S_level \geq 0.5 \\ 1 - S_i(x, y), & \text{якщо } S_level < 0.5 \cup R_i(x, y) > 180 \\ S_i(x, y), & \text{якщо } S_level < 0.5 \cup R_i(x, y) \leq 180 \end{cases} \quad (1.16)$$

– розрахувати рівень насиченості $SLevel$ шляхом усереднення всіх ненульових значень в S_i' ;

– розрахувати можливість наявності полум'я шляхом усереднення значення інтенсивності $IntensityAve$ в V (1.17)

$$FV_i(x, y) = \begin{cases} V_i(x, y), & \text{якщо } V_i(x, y) > \max(0.51 IntensityAve); \\ 0, & \text{в іншому випадку} \end{cases}; \quad (1.17)$$

– визначити присутність полум'я на основі (1.18):

$$FCSV_i(x, y) = FS_i(x, y) \times FV_i(x, y); \quad (1.18)$$

– розрахувати загальний колір полум'я (1.19):

$$CC_i(x, y) = \alpha CC_{i-1}(x, y) + (1 - \alpha) FSCV_i(x, y), \quad (1.19)$$

де α – кумулятивний ефект і прирівнюється до $(N-1)/N$.

Усі пікселі в CC_N , де значення є меншими ніж усереднені ненульові значення в CC_N прирівнюються до нуля. В результаті, регіони які мають візуальну схожість з вогнем та характеризуються мерехтінням виділяються на основі рівняння (1.20):

$$\text{Fire}(x, y) = CT_N(x, y) \times CC_N(x, y). \quad (1.20)$$

У роботі [23], з метою виявлення мерехтіння пікселів полум'я, пропонується використання моделі Маркова, оскільки ділянки вогню просторово виявляють аналогічну періодичну поведінку. Уздовж лінії, що проходить через область полум'я, розташування пікселів у кольоровому просторі різняться так само, як і пікселі полум'я, що спостерігаються протягом певного періоду часу. У даній роботі запропонована модель Маркова, яка натренована для пікселів які характеризують як полум'я, так і його відсутність. Імовірності переміщення між станами пікселя оцінюються протягом заздалегідь визначеного періоду часу навколо границь полум'я. Таким чином, модель не тільки дізнається, як протягом певного періоду відбувається мерехтіння меж полум'я, а також підлаштовує його параметри для імітації просторових характеристик районів полум'я, що дозволяє знизити кількість хибних помилок. Наприклад, рухомий об'єкт, що оголюється кольором, не відобразатиме коливань пікселів у вигляді справжнього полум'я, і це призведе до іншої моделі Маркова.

Вищевказана модель Маркова для виявлення мерехтіння у відеопотоці базується на використанні інформації про колір пікселів. Так, можна виокремити три основні умови на основі яких формується висновок про приналежність пікселя до шуканої категорії (формула 1.21):

$$\begin{aligned} R &> R_T \\ R &> G > B \\ S &> (255 - R) \frac{S_T}{R_T} \end{aligned} \quad (1.21)$$

де R , G , B – відповідно червоний, зелений та синій колірні канали зображення;

R_T – гранична межа для колірного каналу R , в межах якої коливається полум'я;

S_T – гранична межа для насиченості (S), яка буде зменшуватися за умов зростання значення R .

На основі експериментальних досліджень граничні значення для R_T та S_T , коливаються на рівні від 40 до 60 та від 170.

У роботі [24] автори досліджують можливість опису поведінку полум'я в просторі шляхом аналізу напрямків переміщення ділянок зображення, які можуть нагадувати полум'я. Це досягається шляхом визначення нульових перетинів коефіцієнтів вейвлетових перетворень, що і будуть характеризувати показник активності в досліджуваному регіоні.

Імплементация вказаних методів вважається не доцільною, оскільки для виявлення мерехтіння або коливань у пікселях, які можуть дозволити описати пожежу, швидкість відеозйомки повинна бути на достатньо високому рівні, щоб зафіксувати мерехтіння у вогні. Так, для фіксації мерехтіння з частотою 10 Гц, відеозапис повинен фіксувати принаймні 20 кадрів на секунду [24].

1.6 Постановка задачі дослідження

Враховуючи аналіз предметної області можна сказати, що задача розробки інтелектуальних систем відеоспостереження для автоматизованої ідентифікації полум'я у відеопотоці з метою запобігання виникнення та

розвитку пожеж є актуальною та необхідною як в межах нашої країни так і у світі в цілому.

Реалізація поставленої задачі вимагає значних зусиль, спрямованих на вирішення певних протиріч, які можуть виникати між теорією та практикою.

Так, з практичної сторони необхідно передусім досягнути високих показників правильно розпізнаних шуканих об'єктів (True positive (TP)) та правильно проігнорованих об'єктів іншого типу (True negative (TN)), що можливо за допомогою мінімізації показників хибно детектованих об'єктів (False positive (FP)) або якщо шуканий об'єкт був проігнорований детектором (False negative (FN)).

Одночасно з цим, на сучасному етапі розвитку алгоритмів машинного навчання, задача детектування полум'я у відеопотоці не є вирішеною в повній мірі, оскільки досягнення ефективності роботи, достатнього для системи безпеки життєдіяльності, можливо лише за умов значного збільшення кількості обчислювальних ресурсів, що в сучасних реаліях є доволі складно. За відсутності ж необхідних ресурсів можуть спричинятися часові затримки в роботі системи детектування полум'я, що є недопустимим, враховуючи її специфіку.

Зокрема, ряд зовнішніх та внутрішніх факторів може впливати на якість детектування. Відповідно, розроблювана система повинна враховувати:

а) тип горючої речовини:

- тип джерела запалювання;
- тип горючої речовини по якій буде розповсюджуватися полум'я;

б) умови середовища:

- загальні погодні умови (туман, дощ тощо);
- візуальна видимість;
- час дня;

в) розміщення камери спостереження:

- кут нахилу;
- висота;
- зона спостереження.

Основні обмеження з якими ми плануємося справитися полягають в наступних випадках:

– невідомі умови спостереження. Для забезпечення універсальності роботи системи детектування полум'я необхідно передусім врахувати що середовище в яких проводиться спостереження можуть бути абсолютно різноманітними (коридор, вулиця, автостоянка тощо). Одночасно з цим будуть відрізнятися і умови спостереження на різних об'єктах (якість цифрового потоку, кут нахилу та висота розміщення камери тощо);

– особливості горіння полум'я. В даному випадку за доцільним вважається врахувати наявність горючих речовин та можливі температурні режими, що можуть призвести до зміни забарвлення полум'я. Перекриття об'єктами також може спричинити ігнорування детектором випадків загоряння, відповідно необхідно звернути увагу на нестійку форму вогню, особливо мінливість розмірів;

– можливість донавчання. Враховуючи вище наведені обмеження, робота детектора, особливо на початковому етапі може супроводжуватися неправильними спрацюваннями. Саме тому вважається за важливе проводити постійне донавчання, враховуючи попередні помилки, тим самим уникаючи їх повторення в майбутньому.

Використання традиційних методів машинного навчання не дозволить в повній мірі вирішити вищезазначені обмеження, саме тому ми плануємо використовувати згорткові нейронні мережі, які на сьогоднішній день показують найкращі результати особливо в сфері розпізнавання зображень.

Звичайно ж, незважаючи на ряд переваг щодо застосування нейронних мереж для класифікації наявності полум'я на зображенні,

необхідно згадати про їх недоліки, які необхідно врахувати під час розроблення класифікатора вогню. Основна проблема полягає в формуванні навчальної вибірки, оскільки з однієї сторони вона повинна бути репрезентативною та включати в себе всі можливі сценарії зйомки об'єкту (розширення та кут нахилу камери, рівень освітлення, віддаленість до об'єкту). З іншої ж сторони, даний підхід може призвести до перенавчання моделі та, відповідно, супроводжуватиметься зниженням її ефективності.

Іншою проблемою вважається низька швидкість не тільки навчання згорткових нейронних мереж, а й, що найбільш критично, низька швидкість їх роботи, що не дозволяє в багатьох випадках використовувати їх в режимі наближеному до реального часу.

Відповідно, можна відзначити наступні напрямки дослідження:

- розроблення алгоритму виявлення схожих за своїми характеристиками (колір, рух, текстура) об'єктів за допомогою традиційних методів комп'ютерного зору (колірна сегментація, виявлення рухомих об'єктів тощо) з метою формування навчальної вибірки для подальшого навчання нейронної мережі;

- розроблення та модифікація архітектури штучної нейронної мережі, яка б дозволила здійснювати більш точнішу верифікацію шляхом усунення хибних викликів отримуваних з попереднього етапу;

- розроблення програмного забезпечення для виявлення полум'я у відеоряді, яке б ґрунтувалось на використанні двох вищезазначених методів.

Об'єктом дослідження є процеси ідентифікації полум'я на основі цифрових відеопотоків, отриманих з камер спостереження в online режимі.

Предметом дослідження є нейромережева система, методи та засоби ідентифікації полум'я у відеопотоці в online режимі.

Методи дослідження. Методи обробки цифрових зображень (для вирішення задач фільтрації та класифікації цифрових зображень), функціонального аналізу, лінійної алгебри (для побудови математичних

програмних моделей), нейронних мереж (для вирішення задач ідентифікації складних об'єктів), організації обчислювальних процесів із використанням обчислювальної техніки (для розроблення програмного забезпечення).

2 ІДЕНТИФІКАЦІЯ ТА СЕГМЕНТАЦІЯ ПОЛУМ'Я НА ОСНОВІ ГЛИБОКИХ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

2.1 Розроблення базової структури штучної нейронної мережі та аналіз існуючих нейромережових моделей для розв'язання задачі класифікації об'єктів на зображеннях

При побудові сучасних моделей, як для розпізнавання зображень, так і в будь-яких інших задачах пов'язаних з аналізом та обробкою інформації, все частіше використовують штучні нейронні мережі (ШНМ).

ШНМ являють собою технологію, принципи якої побудовані на основі стику таких дисциплін як математика, біологія та інформатика, що дозволило реалізувати алгоритм, який є максимально близький до роботи людської нервової системи (яка вважається найбільш досконалим класифікатором).

В контексті даного дослідження, використання нейромережових технологій вважається за доцільним передусім завдяки їх адаптивної властивості. Так, завдяки здатності донавчання мережі, розроблюваний детектор отримає змогу правильно класифікувати вхідну інформацію навіть за умов зміни параметрів середовища спостереження. Стосовно безпосередньої задачі розпізнавання зображень, то необхідно також відмітити, що ШНМ дозволяють не лише врахувати особливості окремого класу спостереження, а й таких, які можливо помилково віднести до нього.

Ефективність використання ШНМ для побудови систем штучного інтелекту підтверджена рядом наукових публікацій та їх широким розповсюдженням для вирішення ряду як академічних так і практичних проблем. В контексті поставлених задач дисертаційного дослідження, нас цікавлять саме згорткові нейронні мережі, які служать для розпізнавання візуальних патернів безпосередньо з пікселів зображення. Мережі даного типу являють собою розширення нейронних мереж прямого

розповсюдження (feed-forward neural network), які також добре відомі як багатошаровий перцептрон натренований з методом зворотнього розповсюдження помилки.

Вперше, використання згорткових нейронних мереж запропоноване Яном Лекуном в 1990 році [24], яке було отримане внаслідок навчання багатошарової нейронної мережі з алгоритмом зворотнього розповсюдження помилки. Запропонований підхід став базовим під час побудови архітектури ЗНМ і проявляється упорядкованому використанні згорткових шарів (англ. convolution layers) та субдискритизаційних шарів (англ. subsampling layers), або інших слоїв підвибірки (англ. pooling layers). В кінці мережі використовуються повнозв'язні шари (англ. fully-connected layers).

Наступною віхою в етапі розвитку ЗНМ, можна вважати 2012 р., коли Алексу Крижевському, навчивши згорткову мережу, яка складалася з 8 шарів, вдалося досягнути показника 16 % помилок (найближча команда, для порівняння, має показник 26 %) [25]. Для її навчання було використано базу зображень ImageNet, яка включала в себе більш ніж 15 мільйонів зображень розподілених на приблизно 22,000 категорій. Усі зображення були переведені до одного розширення 256x256.

Для того щоб зробити мережу нелінійною була використана функція активації ReLU, яка реалізує простий пороговий перехід в нулі і може бути описана формулою (3.1) [25]:

$$f(x) = \max(0, x) \quad (2.1)$$

Це дозволило в значній мірі підняти швидкість сходження стохастичного градієнтного спуску та зменшити затрати, необхідні на її вирахування, в порівнянні з іншими функціями активації.

Сама мережа характеризувалася двома паралельними наборами шарів, обоє з яких по чергово використовували згортковий шар та пулінг по максимуму (max-pooling). На перший згортковий шар подається розмір

224x224x3, який поступово зменшується від входу в мережу до самого виходу за допомогою фільтрів. Фільтри починаються з просторових розмірів 5x5 та 55 в глибину (усього 1375 параметрів), які після кожного проходження згорткового шару поступово зменшують свій об'єм до 3x3, а глибина фільтра, навпаки, зростає до 192 (усього 1728 параметрів) в кінці мережі. В кінці мережі присутні два повнозв'язних шари з 2048 вузлами та вихідним шаром на 1000 вузлів (що відповідає 1000 класів об'єктів).

В даному контексті необхідно відзначити, що після 2012 р., в усіх подальших конкурсах по розпізнавання зображень чи детектування об'єктів у відеопотоці, перемогу здобували лише ті методи, які базувалися на використанні нейромережових моделей, що і обумовлює необхідність їх використання під час розроблення системи детектування полум'я у відеопотоці.

В 2013 р. перемогла мережа ZFNet [26], яка за своєю суттю є модифікованою версією мережі AlexNet, зі збільшеним розміром згорткових шарів та зменшеним фільтром для першого шару.

Більш цікавішою виявилася модель VGG, яка отримала друге місце в конкурсі класифікації ImageNet, але перше в локалізації об'єкту. Основна ідея полягає у використанні менших фільтрів, розмірами 3x3, та добавлені більшої кількості шарів (16-19 в залежності від модифікації, в порівнянні з 8 шарами у AlexNet).

GoogleNet, яка стала переможцем в 2014 р., була ще глибшою (22 шари) і пропонувала використання ефективного модуля – Inception, що дозволило в значній мірі зменшити кількість параметрів мережі (з 60 мільйонів, в порівнянні з AlexNet, до 5 мільйонів) за рахунок менших згорток, але більшої кількості шарів та повної відмови від використання прихованих повнозв'язних шарів, за рахунок використання global average pool у вихідному шарі. Графічно даний процес відображений на рисунку 2.1.

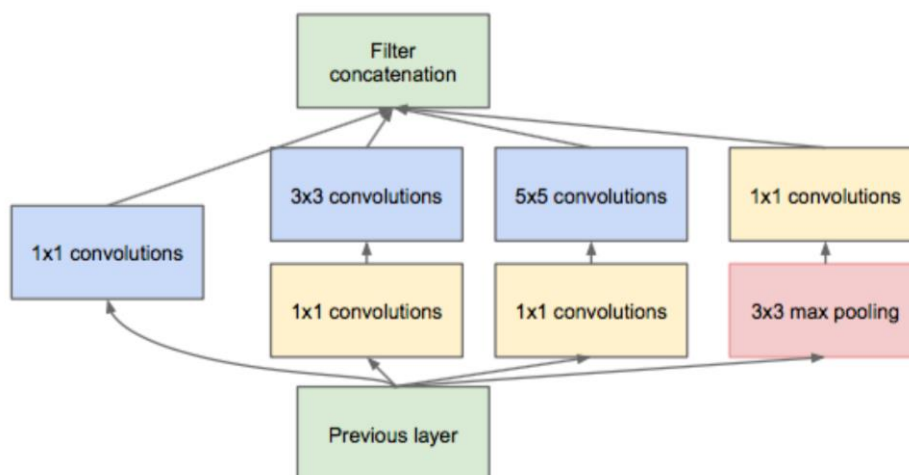


Рисунок 2.1 – Принцип роботи модуля Inception

Звідси можемо спостерігати, що використовується згортки розміром 1x1 із метою зменшення розмірності, зокрема, що є доволі ефективним процесом враховуючи можливості його донавчання. Повна архітектура мережі включає в себе 9 таких повторних модулів. Графічно даний процес відображений на рисунку 2.2.

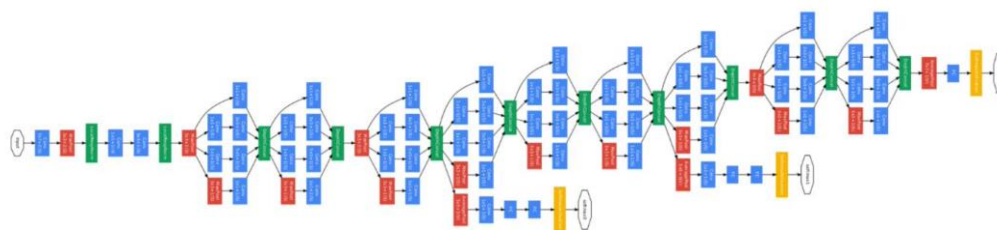


Рисунок 2.2 – Архітектура нейронної мережі GoogleNet

Уже в 2015 р. компанії Microsoft, за допомогою глибоких згорткових мереж, із коефіцієнтом помилки 3,6%, вдалося отримати не тільки перемогу на конкурсі ImageNet, а й показати результати кращі ніж в людини, які складають 5,1% [26].

Даний показник було досягнуто завдяки тренуванню дуже глибоких нейронних мереж (152 шари). Проте, звичайне додавання шарів до нейронної мережі призведе лише до збільшення помилки як на навчальній, так і на тестовій вибірці. Для вирішення даної проблеми було запропоновано використання концепції остаточного блоку (residual block), принцип якого відображений на рисунок 2.3.

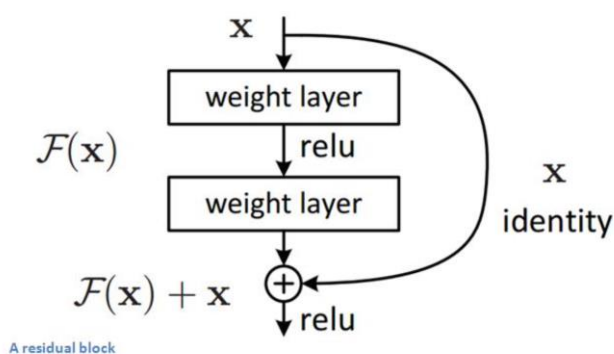


Рисунок 2.3 – Графічне відображення принципу роботи остаточного блоку

З вищезазначеного рисунку можемо спостерігати, що вхідний сигнал проходить через два шари з вагами, посередині яких знаходиться функція активації, тим самим отримуючи певний результат $F(x)$, який на виході додається до вхідного сигналу.

Отже, узагальнюючи вищесказане, можемо констатувати, що існуючі нейромереві моделі дозволяють ідентифікувати об'єкти на зображеннях з доволі високим показником точності (від 80% правильності реагування за умов належних налаштувань), проте, в контексті поставленого перед нами завдання, необхідно забезпечити максимальне наближення до ідеальної точності.

Відповідно, вважається за доцільним, використання б нейромеревих моделей, проте лише за умов їх модифікації з метою покращення показника точності під час детектування полум'я на зображеннях.

Для початку спробуємо відобразити архітектуру звичайної нейронної мережі, яка б дозволяла визначати наявність полум'я на зображенні. З метою спрощення числових розрахунків, на вхід будемо подавати зображення розмірами 28x28 в трьох кольірних каналах (R, G, B). Графічно даний процес відображено на рисунок 2.4.

Як видно із вищезазначеної схеми, кожен з нейронів з'єднаний з відповідним пікселем зображення. На виході отримуємо шар з двома нейронами, які відповідають за відсутність та наявність полум'я на зображенні відповідно. Загальну кількість параметрів нейронної мережі можемо вирахувати за наступною формулою (2.2):

$$y_j = \sum_{j=0}^{2352} w_{ij}x_j + b_i = 2352 \times 2 + 2 = 4702, \quad (2.2)$$

де $i = 0,1$ – виходи нейронної мережі.

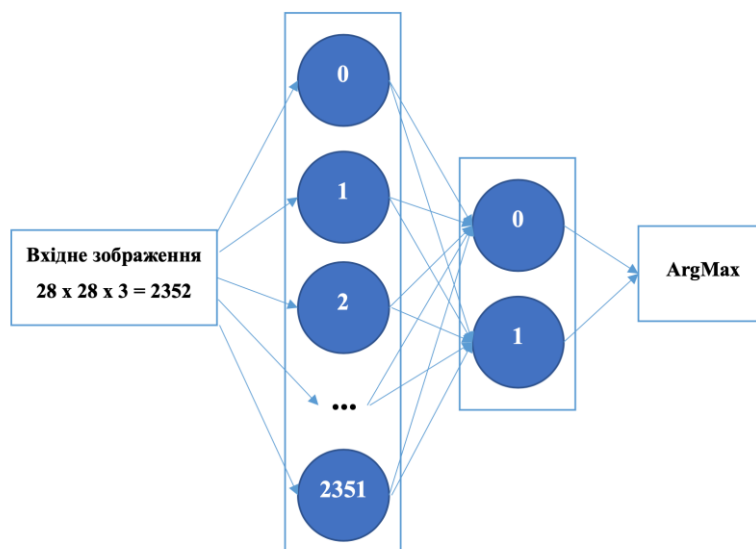


Рисунок 2.4 – Архітектура простої нейронної мережі для бінарної класифікації

Для покращення здатності нашої ШНМ до навчання й узагальнення, пропонується додати додатковий прихований шар з певною функцією активації. Найбільш розповсюджені функції активації перераховані в таблиці 2.1.

Таблиця 2.1 – Список найбільш розповсюджених функцій активації нейронної мережі

Функція	Формула	Область значення
Лінійна	$\varphi(x) = ax + b$	$(-\infty; +\infty)$
Логістична	$\varphi(x) = \log(1 + e^x)$	$(0; 1)$
Сигмоїд	$\varphi(x) = 1 / (1 + e^x)$	$(0; 1)$
Гіперболічний тангенс	$\varphi(x) = a \operatorname{th} bx$	$(-1; 1)$
«Випрямовувач» (ReLU)	$\varphi(x) = \max(0, x)$	$(0; 1)$

З метою приближення нейронною мережею складної функції, за умов дотримання умови достатньої кількості шарів, використовують нелінійну функцію активації.

В даному контексті доволі часто зустрічається використання сигмоїдної функції, проте вона може призвести до збільшення градієнтів або затухання. З метою вирішення даної проблеми загальноприйнятим підходом є використання більш спрощеної нелінійної функції активації – ReLU, яку і було вирішено застосовувати під час побудови ШНМ.

Відповідно, модифікована архітектура розробленої ШНМ набуває наступного вигляду (рисунок 2.5). Кількість нейронів в обраній функції активації буде набувати значень (2.3):

$$N = 128, 256, 512, 1024, 2046 \quad (2.3)$$

$$h_i = \text{ReLU} \left(\sum_{j=0}^{2352} W_{i,j}^{(h)} \times x_j + b_j^{(h)} \right), i = 0, 1, \dots, (H - 1)$$

$$y_i = \sum_{j=0}^{2352} W_{i,j}^{(h)} \times x_j + b_j^{(h)}, i = 0, 1$$

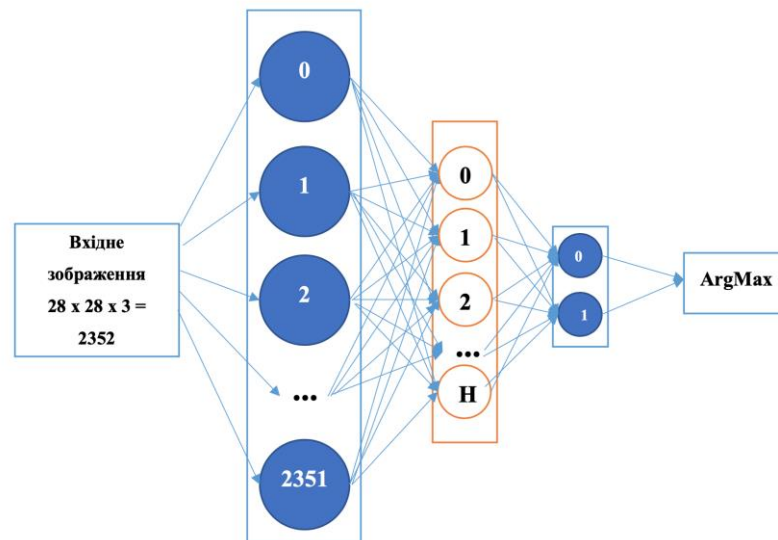


Рисунок 2.5 – Архітектура простої нейронної мережі для бінарної класифікації з функцією активації ReLU

Сумарна кількість параметрів ШНМ буде залежати від кількості нейронів які розміщуються в прихованому шарі (2.4):

$$2352 \times H + H + H \times 2 + 2. \quad (2.4)$$

Що стосується згорткових нейронних мереж, то вони за своєю архітектурою, складаються з трьох основних частин: згорткові шари (convolution layers), шари підвибірки (pooling layers) та повнозв'язні шари (fully connected layers). Кожна з частин даної нейронної мережі виконує свою певну особливу роль, об'єднуючись разом в єдину ієрархічну модель, де кожен з шарів поступово іде один за одним.

На початку кожної згорткової нейронної мережі розміщується безпосередньо згортковий шар, який дозволяє репрезентувати усе зображення як набір ознак, шляхом перемноження значень фільтрів на значення пікселів зображення, тим самим генеруючи різноманітні карти ознак. Значення кожного пікселя карти ознак може бути отримане за наступною формулою (2.5):

$$V = \left| \frac{\sum_{i=1}^q \left(\sum_{j=1}^q f_{ij} d_{ij} \right)}{F} \right|, \quad (2.5)$$

де f_{ij} – коефіцієнт згорткового ядра на позиції i, j ;

d_{ij} – значення вхідного пікселя, який відповідає позиції f_{ij} ;

q – розмірність ядра;

F – сума коефіцієнтів ядра (якщо сума коефіцієнтів прирівнюється до нуля, то 1).

Що стосується безпосередньо процесу згортки, то дану операцію математично можна описати наступним чином (2.6):

$$c[m, n] = \alpha[m, n] \otimes h[m, n] = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} h(m, n) \times \alpha \left(m + j - \frac{N}{2}, n + k - \frac{M}{2} \right), \quad (2.6)$$

де $h[m, n]$ – ядро з розмірами $m \times n$, де $m = 0, \dots, (J - 1), n = 0, \dots, (K - 1)$;

$\alpha[m, n]$ – вхідне зображення з розмірами $m \times n$ (в нашому випадку матриця розмірами 28x28).

Оскільки у нас кольорове зображення, то буде використовуватися трьохмірне ядро згортки $h(m, n, y)$ і одразу ж додаємо ReLU як функцію активації (2.7):

$$c[m, n, z] = \sum_y \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} h(m, n, z) \times \alpha\left(m + j - \frac{N}{2}, n - \frac{M}{2}, z\right) + b(z) \quad (2.7)$$

Щодо переваг використання шару такого типу, то необхідно відмітити три основні положення:

- інваріативність до розташування об'єкта;
- невелика кількість параметрів, яка встановлюється при навчанні;
- здатність помітити кореляцію між сусідніми пікселями [30].

Шар підвибірки іде одразу ж за згортковим шаром, та може бути використаний для зменшення розмірності карти ознак та параметрів мережі.

З цією метою зазвичай використовують операції *max pooling* та *average pooling*. Пулінг по максимуму приймає найбільше значення серед ділянки зображення, яку на даний момент покриває ядро згортки. Середньозважений пулінг натомість приймає до уваги всі значення у вікні.

Так, для прикладу на рисунку 2.6, відображений процес зменшення розмірності карти ознак розміром 4 x 4 до розміру 2 x 2, шляхом використання оператора *max pooling* з розміром 2 x 2.

Говорячи про доцільність вибору *max pooling* чи *average pooling*, то у роботах [31], [32] був проведений детальний аналіз їх ефективності, де було виявлено, що в порівнянні з двома операціями пулінгу, *max pooling* має ряд переваг: покращена генералізація та вибір ключових характеристик, здатність до більш швидшої конвергенції (сходження), покращена стійкість до зрушень.

Про це, зокрема, свідчить переважне використання даної концепції в сучасних варіаціях ЗНМ. Отже, спробуємо його використати в нашій ШНМ.

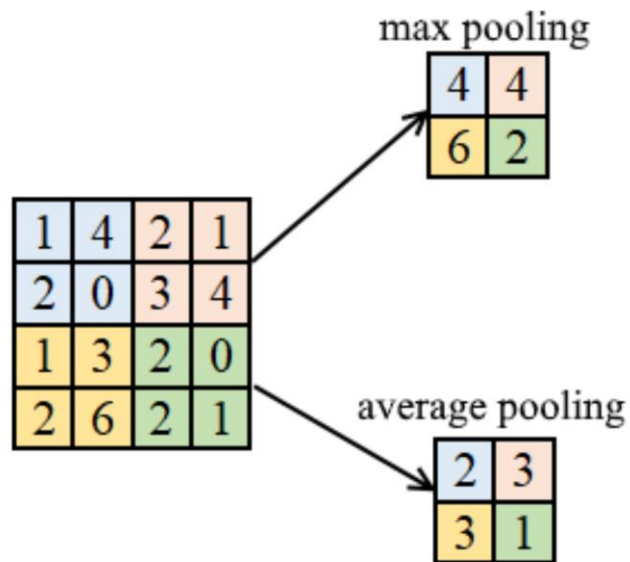


Рисунок 2.6 – Приклад зменшення розмірності карти ознак за допомогою max pooling та average pooling

Математично max pooling може бути представлений наступним чином (2.8):

$$M(m, n, z) = \max I(i, j, z);$$

$$i = D_x \times m, \dots, (D_x \times m + J); \quad (2.8)$$

$$j = D_j \times n, \dots, (D_j \times n + K).$$

Графічно процес побудови ШНМ відображений на рисунку 2.7. Як можемо спостерігати, на вході подається зображення розмірами 28x28x3 до якого застосовують фільтр з розміром 5x5x3. В результаті отримуємо 6 карт ознак, кожна з яких має розміри 24x24. Після застосування maxpool розмірність карт впаде до 12x12. Повторне використання згортки з ядром 5x5x3 і операції maxpool дозволить для кінцевого шару отримати $4 \cdot 4 \cdot 6 = 96$ параметрів, які лінійно з'єднані з вихідним шаром на 2 нейрони.

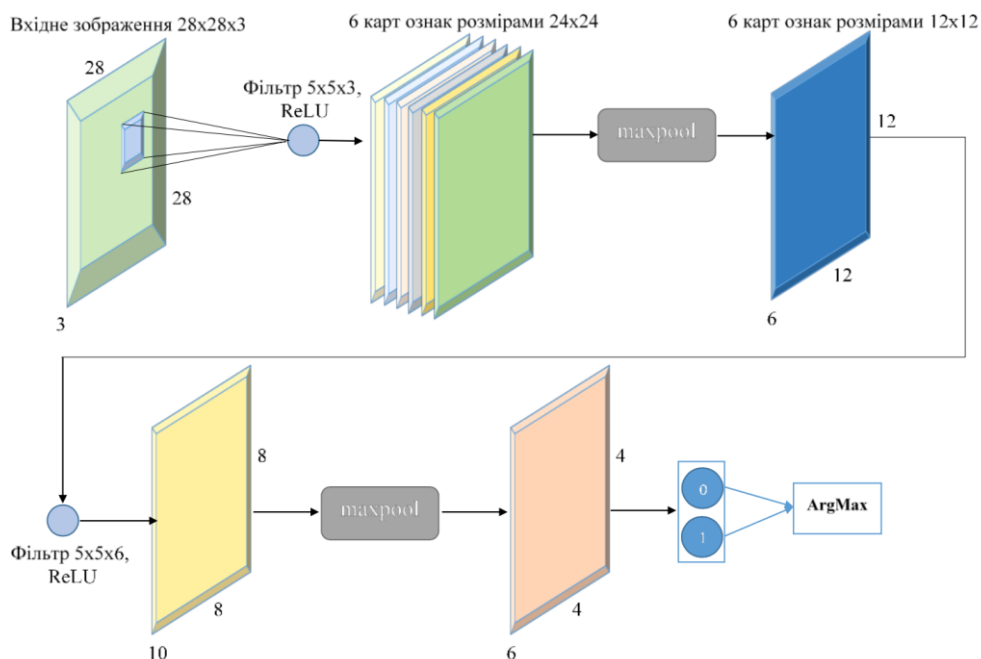


Рисунок 2.7 – Архітектура згорткової нейронної мережі з двома операціями згортки на maxpool

Окрім вищезгаданих підходів, використовуваних у шарах підвибірки, існують ще три добре відомих пулінги, кожен з яких використовується в залежності від його призначення:

- стохастичний пулінг – ще один метод пулінга, який дозволяє знизити дискретизацію, шляхом випадкового виділення значень в кожному регіоні пула відповідно до мультиномінального розподілу [82], що певною мірою нагадує пулінг по максимуму. Використання даного підходу дозволяє запобігти проблемі перенавчання нейронної мережі;

- пулінг просторової піраміди (Spatial pyramid pooling) – метод пулінгу, який дозволяє зняти обмеження щодо необхідності подачі на вхід зображень лише одних розмірів, з метою забезпечення належного функціонування відповідних методів. Цілком зрозуміло, що за умов передачі зображень з різними розширеннями, може знизитися точність розпізнавання, і щоб запобігти уникненню даної проблеми, у роботі [33] пропонується замінити останній шар пулінгу на шар пулінгу просторової

піраміди. Зазначений підхід дозволить виділити ознаки сталих розмірів з набору зображень, тим самим створюючи універсальний метод представлення даних у різних масштабах, розмірах чи пропорціях;

– def-pooling дозволяє справитися з надзвичайно складною проблемою під розв'язання задачі комп'ютерного зору – необхідності врахування можливості деформації об'єкту [34]. Даний метод може замінити max-pooling на будь-якому етапі проектування ЗНМ, що дозволить врахувати можливість незначної зміни частини досліджуваного об'єкту.

Вищезазначені пулінги проектуються відповідно до поставлених перед архітектором мережі задач. Вони можуть як об'єднуватися в певну групу, так і можливе використання лише одного пулінгу.

Що стосується повнозв'язного шару, то він дозволяє перетворити усю отриману інформацію на попередніх етапах в вектор з певною довжиною, в залежності від поставленої перед нами задачі.

У випадку, що розглядається в кваліфікаційній роботі, повнозв'язний шар набуватиме вигляду $y_i \in [0,1]$, оскільки $\sum_i y_i = 1$ може бути інтерпретована як можливість приналежності x_i до певного класу (в цьому випадку – присутність або відсутність полум'я).

Сама архітектура повнозв'язних шарів модифікується лише в окремих випадках, проте доволі часто використовують підхід трансферного навчання (transfer learning), який дозволяє враховувати параметри мережі, яка, для прикладу, була натренована на базі зображень ImageNet, з додаванням нових шарів, які б дозволили адаптувати її під поставлену задачу.

Проте, в багатьох випадках дані шари навпаки, видаляють, щоб зменшити розмірність самої мережі. Більш детально даний процес досліджується в наступних підрозділах.

2.2 Застосування каскадного методу в штучних нейронних мережах

Аналізуючи набір зображень на яких попередньо розроблені детектори полум'я здійснювали помилкові спрацювання, або не спрацювали за умов наявності вогню, було виділено дві ключові проблеми, за умов вирішення яких лише можливо забезпечення надійної ефективності роботи даного детектору.

Перша проблема полягає у низькій якості як усього зображення так і самого осередку полум'я, яке може знаходитися на значній відстані від місця розташування камери відеоспостереження і відповідно характеризуватиметься абсолютно інакшими ознаками (рисунок 2.8 а). Друга проблема – неможливість врахування повної картини зображення традиційними методами детектування (рисунок 2.8 б).

З метою вирішення вищезазначених проблем було запропоновано використовувати додаткову CNN з метою перевірки правильності інформації наданої генератором гіпотез про можливість наявності об'єкту. Мережа такого типу в науковій літературі називається білінійною [35], [36] і використовується зазвичай для моделювання двохфакторних варіантів таких як «стиль» та «зміст».

В нашому ж випадку, структура мережі (рисунок 2.9), з метою розв'язання проблеми низького розширення зображень, поділяється на дві під-мережі. Для оцінки розмірів полум'я, з метою вибору НМ, яка повинна використовуватися, аналізується так звана гіпотеза, яка була отримана в результаті генерування ділянок, які можуть візуально нагадувати полум'я.



а)



б)

Рисунок 2.8 – Проблемні ситуації під час детектування полум'я:

а) низька роздільна здатність (червоним прямокутником відзначені місця розташування полум'я);

б) текстура полум'я як частина певного об'єкту

Якщо її розмір не перевищує $1/15$ висоти усього зображення, то використовується підмережа два (n_l). В інших випадках – підмережа один (n_h) (2.9):

$$n_l \text{ якщо } b_b \leq \frac{1}{15} \text{ image_height}, \quad (2.9)$$

n_h в іншому випадку.

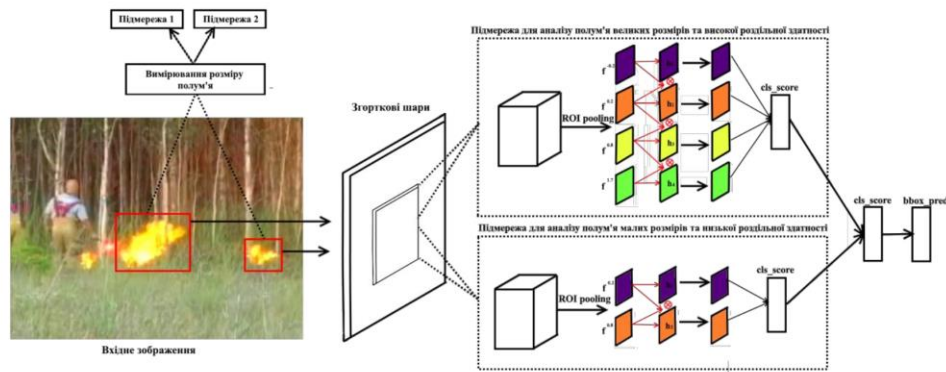


Рисунок 2.9 – Структура каскадного методу детектування полум'я на основі згорткової нейронної мережі з пороговими функціями активації

Перша мережа визначає наявність полум'я враховуючи його особливості на зображеннях з високою роздільною здатністю. В свою чергу, друга мережа враховує особливості полум'я на зображеннях з низькою роздільною здатністю. Кожна з мереж приймає в якості вхідних даних карту ознак отриманих з попередніх шарів мережі на етапі генерації гіпотез, та закінчуються вихідним шаром який оцінює можливість наявності полум'я на зображенні, та, за умов наявності якого, виробляється два вихідних вектори для кожного пропонованого об'єкту:

- перший вектор виводить оцінку класифікації відповідно до класу об'єкту та заднього фону;
- другий використовується для виведення обмежувальної рамки щодо розміщення об'єкту;
- шар підвибірки ROI (ROI pooling layer) використовується для поєднання карти ознак у вектор фіксованої довжини.

З метою розв'язання проблеми неможливості врахування усієї картини зображення пропонується введення так званих порогових функцій на основі пропонованих ділянок різних розмірів та опорними регіонами ($f_{-0.2}, f_{0.2}, f_{0.8}, f_{1.7}$). Схематично процес отримання таких ознак мережею відображено на рисунку 2.10.

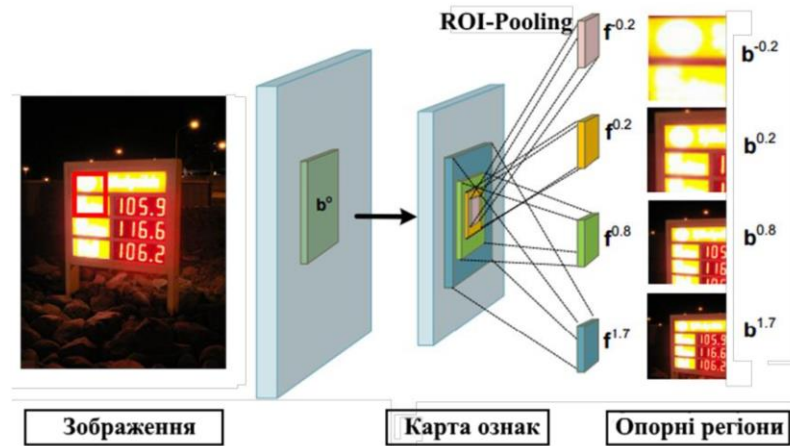


Рисунок 2.10 – Приклад використання ROI-pooling для отримання ознак CNN з різними розмірами і опорними регіонами

Червоним прямокутником позначено можливе розміщення полум'я надане генератором гіпотез. ROI-pooling надає чотири опорних ділянки для $\{b^P\}$. Для передачі повідомлень між функціями було використано двосторонню структуру з пороговими функціями [36].

Після чого всі ознаки h^3 для $i = 1, 2, 3, 4$ проходять через ряд слоїв i CNN із загальними параметрами для отримання кінцевого результату, що дозволяє здійснити прогнозування щодо належності об'єкту до класу y , де $y \in Y = \{0,1\}^C$.

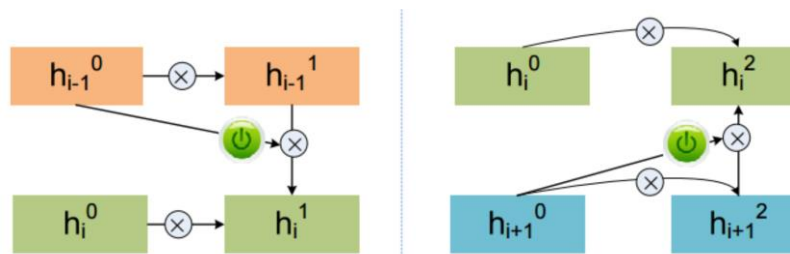


Рисунок 2.11 – Двостороння структура з пороговими функціями. \otimes – передбачає процес згортання, кнопка перемикавання відповідає за порогову функцію

2.3 Тонке налаштування нейронних мереж (Fine tuning)

Навчання ЗНМ шляхом присвоєння випадкових вагових величин вимагає значних часових та обчислювальних ресурсів, особливо якщо зображення мають широке розширення. Окрім того, без достатньої кількості зображень в навчальній вибірці, яка повинна бути попередньо розмічена на окремі класи об'єкту, неможливо досягнути якісних показників ефективності. З іншої ж сторони, незважаючи на велику кількість навчальних екземплярів, якщо вибірка є не репрезентативною, то це може призвести до перенавчання мережі.

З метою уникнення вищезазначених проблем, загальноприйнятою технікою є використання трансферного навчання. Ключова ідея такого підходу полягає у тому, що використовується уже готова архітектура ЗНМ, побудована на основі доволі великого та репрезентативного датасету (для прикладу ImageNet, який поділяється на 1000 категорій, сумарною кількістю 1.2 млн. зображень), для донавчання на окремому датасеті, відповідно до поставлених завдань.

Враховуючи поставлені перед нами вимоги щодо побудови детектора полум'я (висока ефективність детектування та можливість роботи в режимі наближеному до реального часу), використання загальноприйнятих архітектур нейронних мереж (GoogleNet, VGG-16 тощо) є проблематичним, оскільки вони характеризуються великими розмірами та кількістю параметрів, що може спричинити її неефективність на машинах із невеликими обчислювальними ресурсами.

Відповідно, з метою розв'язання поставленої проблеми, було вирішено використати мережу SqueezeNet [37]. Даний вибір обумовлений тим, що дана мережа дозволяє працювати з однаковою ефективністю глибокої мережі AlexNet, проте має в 50 раз менше параметрів. Основними стратегіями до побудови такої архітектури є:

- заміна фільтрів 3x3 на 1x1 фільтри, що дозволяє зменшити кількість параметрів в 9 разів;
- зменшення кількості каналів більш стиснутими модулями;
- пониження дискретизації останніх шарів мережі. Стратегії 1 та 2 націлені на зменшення кількості параметрів мережі, третя стратегія на покращення її ефективності, незважаючи на зменшення параметрів.

З метою покращення ефективності детектування було вирішено використовувати модифіковану архітектуру мережі SqueezeNet шляхом впровадження залишкових (residual) блоків (рисунку 2.12).

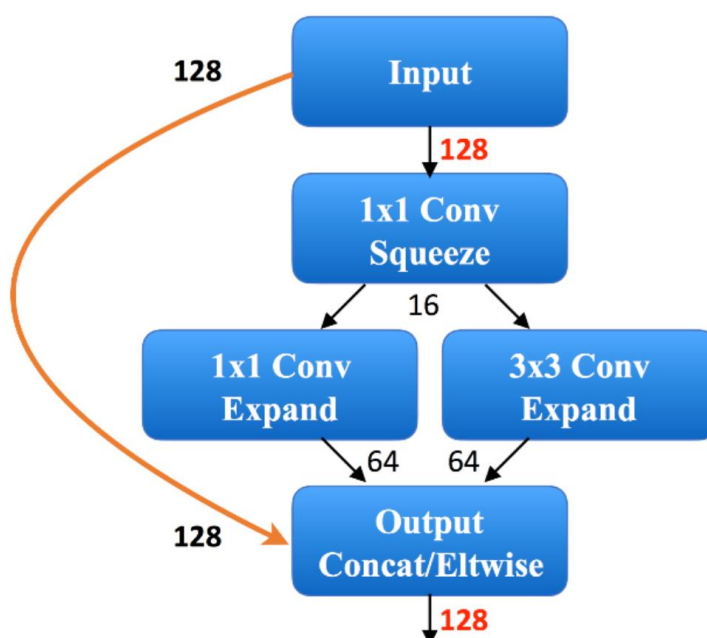


Рисунок 2.12 – Residual блок

Кінцева архітектура мережі зображена на рисунку 2.13. Під час тонкого налаштування ваг нейронної мережі, ми починаємо з заздалегідь підготовлених моделей, які навчені на оригінальному наборі даних (тобто зображення без будь-яких спотворень). Ми здійснили тонке налаштування перших N шарів моделі на спотвореному наборі даних, а в інших шарах використовували фіксовані параметри. Основною причиною такої фіксації

є спостереження, що під час розмиття і шуму останні шари набувають більшого впливу на низькорівневі властивості, такі як колір, границі та текстурні ознаки. Проте, ці спотворення мають незначний вплив на інформацію вищого рівня, таку як семантичні значення зображень.

Таким чином, під час тонкого налаштування ваг нейронної мережі, ми орієнтуємося на вихідні шари SqueezeNet, які містять в собі інформацію більш низького рівня.

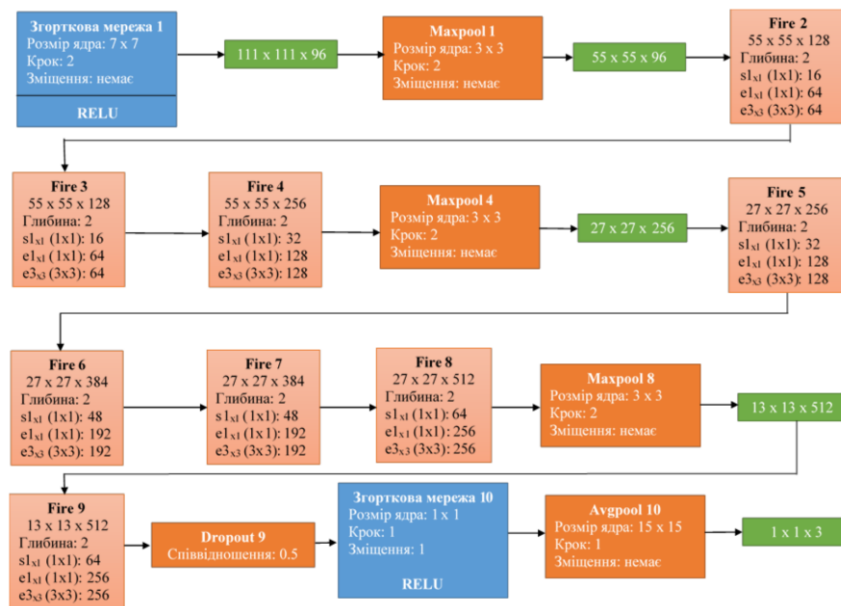


Рисунок 2.13 – Архітектура розробленої нейронної мережі

У зв'язку з обмеженим набором даних у нашому датасеті (лише 5000 зображень на яких міститься полум'я) пропонується використовувати: поворот зображень, передискретизацію та кадрування.

На початковому етапі було вирішено навчати мережу з випадковими поворотами 0° , 75° , 285° , проте це не дало будь-яких покращень. Лише за умов використання усередненого значення отриманого від класифікатора під час повороту зображень в зазначених напрямках (значення були підібрані експериментальним шляхом) була отримана змога покращити ефективність роботи детектора.

Що стосується передискретизації та кадрування зображень, то процес їх використання нагадує операцію з поворотами зображення. Так, вхідне зображення поділяється на 5 частин (чотири кутових та одне центральне), до яких і застосовується операція передискретизації. Отримуваним результатам з кожної частини зображення знову ж усереднюються, внаслідок чого отримується кінцеве передбачення.

Необхідно також відзначити про те, що усі моделі на певному етапі навчання зіткнулися з проблемою перенавчання (overfitting), що було виявлено у зв'язку з початком зростання втрат на валідаційному наборі даних (250 зображень). З метою вирішення даної проблеми, починаючи з даного етапу, показник швидкості навчання був зменшений в 10 раз (в порівнянні з початковим).

Прикінцевим етапом доналаштування детектора полум'я стало реалізація ансамблю з нейронних мереж. Для цього була навчена модель, яка навчалася лише на базі даних зображень із полум'ям (без застосування підходу трансферного навчання).

Незважаючи на те, що така модель характеризувалася нижчою точністю, поєднавши її з іншими тонко налаштованими моделями вдалося покращити загальний показник ефективності детектора полум'я.

На нашу думку даний факт зумовлений тим, що мережа натренована лише на датасеті який містить з собі зображення з вогнем, врахувала інші ознаки об'єкту аніж тонко налаштована мережа.

Додатково до ансамблю ввійшли ще дві нейронні моделі, які містили уже в собі ознаки отримувани в результаті трансферного навчання:

- перша модель донавчалася внаслідок штучного збільшення існуючого датасету шляхом повороту зображень;
- друга – на датасеті збільшеного шляхом кадрування та передискретизації.

Розроблена модель ідентифікації полум'я у відеопотоці представлена в вигляді структурно-функціональної схеми на рисунок 2.14.

Процес виявлення полум'я можна розділити на два послідовних етапи: швидка фільтрація кадрів відеопотоку на наявність або відсутність можливих ознак горіння та, у разі наявності ознак об'єкту, подальша верифікація з використанням глибоких згорткових мереж. Такий підхід дозволяє зменшити як кількість ресурсів, які необхідні для опрацювання отриманих зображень, так і загальний час який необхідний для надання оцінки про наявність чи відсутність полум'я.

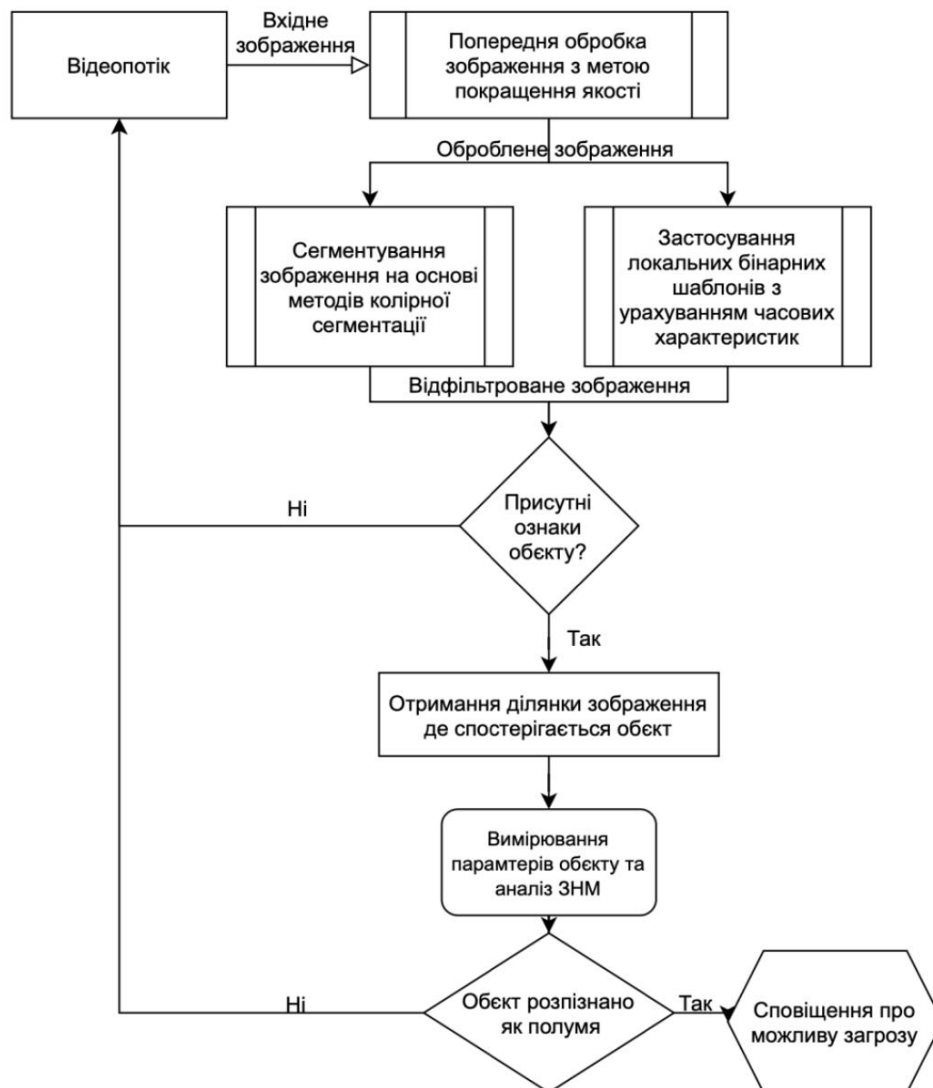


Рисунок 2.14 – Структурно-функціональна схема системи ідентифікації полум'я

В кваліфікаційній роботі пропонується використання згорткових нейронних мереж для врахування певних особливостей об'єкту спостереження та зменшення кількості хибно-позитивних спрацювань. Досліджено архітектури GoogleNet, AlexNet, ResNet та ZFNet, які сприяють розвитку ЗНМ та дозволяють покращити ефективність їх роботи. Модифіковано нейромережеві моделі в напрямку покращення показника точності. З цією метою була побудована елементарна архітектура ЗНМ, яка дозволяє визначати наявність полум'я на зображенні.

Для покращення здатності штучної нейронної мережі до навчання й узагальнення, було добавлено додатковий прихований шар з функцією активації, а на противагу широко застосовуваним нелінійній та сигмоїдній функціям, в розроблюваній архітектурі було вирішено використати більш спрощену нелінійну функцію активації – ReLU.

Запропонована архітектура згорткової нейронної мережі з двома операціями згортки на maxpool та структура каскадного методу детектування полум'я на основі згорткових нейронних мереж з використанням порогових функцій.

3 ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ ТА ЕКСПЕРИМЕНТАЛЬНА ОЦІНКА РОЗРОБЛЕНОЇ МОДЕЛІ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ДЕТЕКТУВАННЯ ПОЛУМ'Я

3.1 Ідентифікація полум'я на основі його кольорових ознак

Базовою характеристикою яка дозволяє здійснити класифікацію полум'я на зображенні є його колір. Як уже було зазначено вище, з цією метою використовують різноманітні кольірні моделі. Однак, говорити про ефективність певної окремої моделі доволі складно, оскільки передусім все залежить від набору правил до яких проводиться віднесення пікселів до тієї чи іншої категорії та набору зображень на яких проводилось тестування ефективності роботи алгоритмів.

Так, у роботі [40], для виявлення полум'я на зображенні була здійснена його сегментація на основі кольірної моделі RGB. Ефективність запропонованого методу (із використанням адаптивної моделі заднього фону на основі Гауссового розподілу) складає 98.89%.

Проте, переважна більшість авторів зазначає, що використання однієї кольірної моделі RGB є недостатнім для проведення ефективної сегментації зображення. Так, у науковому дослідженні [41] було здійснено порівняння кольірних моделей HSI, YCbCr та RGB. Зокрема, в даній роботі було запропоновано та експериментально доказано ефективність використання декількох кольірних моделей (таблиця 2.1).

Для вирішення даної проблеми було вирішено самостійно проаналізувати ефективність повноти сегментації за допомогою 4 кольірних моделей: RGB, HSV, CIE L*a*b, та YCbCr.

Для цього було відібрано 150 зображень на яких спостерігається пожежа (рисунок 3.1). Дані зображення є повністю варіативні, тобто характеризується різним розширенням, рівнем зашумленості, освітленості, умовами зйомки тощо.

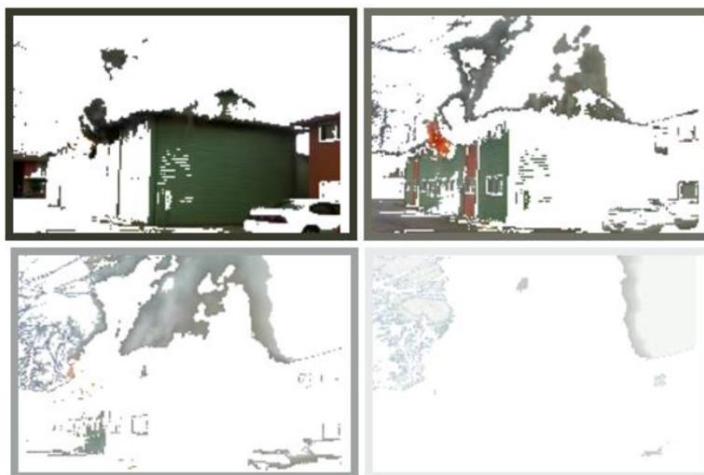
Для виділення границь в яких може знаходитися полум'я було використано метод кластеризації k -середніх (k -means) (рисунок 3.2). В загальному випадку кластеризація дозволяє розділити множину вхідних векторів на певну задану кількість груп відповідно до їх степені схожості. Використання методу k -середніх спрямовано на мінімізацію середньоквадратичного відхилення на точках кожного кластеру (рисунок 3.3).



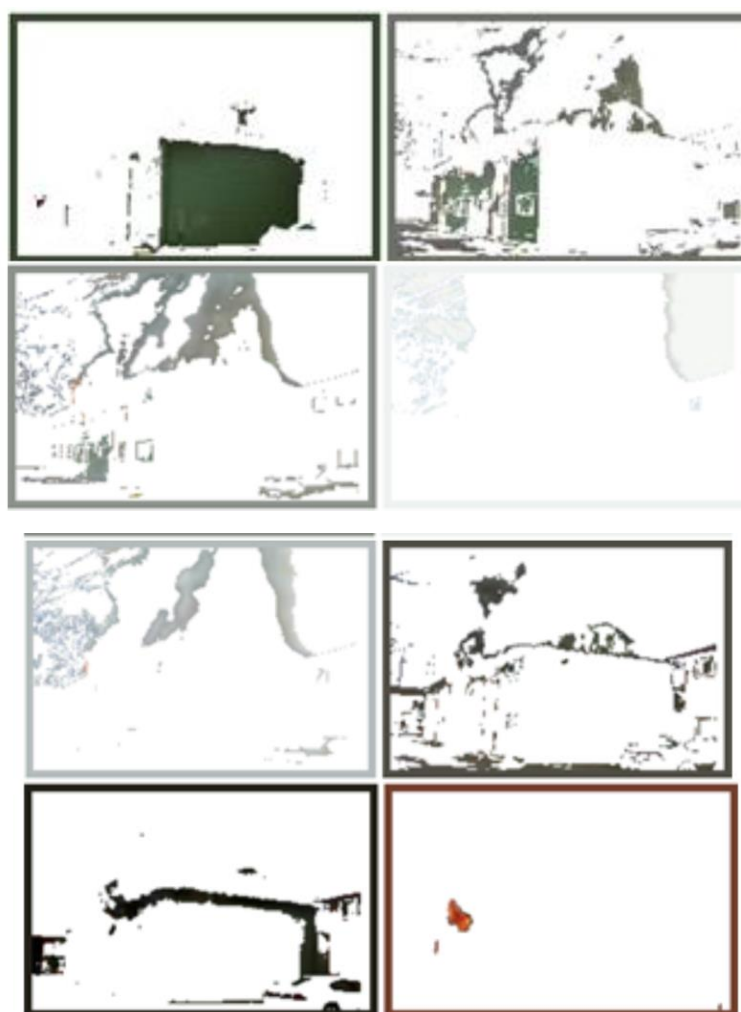
Рисунок 3.1 – Екземпляри аналізованих зображень



Рисунок 3.2 – Результат застосування методу k -середніх (чотири кластера)



а)



б)

Рисунок 3.3 – Результат застосування методу k-середніх:

а) кластерування на чотири кластера;

б) кластерування на вісім кластерів

Шляхом застосування методу k -середніх було отримано ділянки зображення в якій знаходиться полум'я (на рисунках 3.2 та 3.3б це останні зображення). Отримані результати дозволили дослідити кореляцію між різними колірними каналами, визначити межі в яких може коливатися полум'я та, як наслідок, вивести правила, за умов дотримання яких, пікселі зображення необхідно верифікувати на приналежність до категорії «полум'я», оскільки за колірними характеристиками вони можуть мати певну схожість з іншими об'єктами (червона футболка, проміння сонця тощо).

3.2 Створення навчальної множини даних

Побудова якісної та репрезентативної навчальної вибірки відіграє ключове значення для навчання нейронної мережі з високим показником точності. Навіть за умов використання сучасних моделей ЗНМ неможливо досягнути будь-яких якісних результатів якщо дана вибірка буде неправильно сформована. Відповідно, за важливим вважається дослідження даного процесу, з метою виявлення додаткових умов та параметрів, які можуть позитивно або негативно вплинути на роботу класифікатора.

Базовим підходом щодо вирішення проблеми навчання класифікатора зображень є використання готових наборів зображень, таких як ImageNet [24], Pascal VOC [26], CIFAR10 / CIFAR100 [42] тощо. Навчальні вибірки такого типу містять в собі велику кількість зображень, які розподілені по декілька груп різноманітних категорій. Натомість такі навчальні вибірки як MNIST, NIST [43], The Street View House Numbers (SVHN) Dataset, Labelled Faces in the Wild тощо, містять набори зображень які відносяться лише до однієї категорії. Незважаючи на велику кількість можливих навчальних вибірок та кількості зображень, які вони в собі містять (для прикладу ImageNet станом на кінець 2017 року містить 14, 197, 122 зображення поділених на 21,841 категорію), для вирішення

прикладних задач нерідко вимагається формування власної навчальної вибірки.

При формуванні нової навчальної вибірки необхідно передусім забезпечити її репрезентативність, що передбачає врахування усіх можливих параметрів шляхом додавання зображень отриманих в різноманітних умовах зйомки. Даний підхід дозволяє уникнути проблему, яка може призвести до перенавчання класифікатора та, як наслідок його неправильна робота внаслідок спрацювання на об'єктах, які не відносяться до заданої категорії пошуку.

Найбільш яскраво про дану проблему свідчить історія щодо знаходження ШНМ прихованих танків на зображенні [44].

Так, дослідники спробували навчити нейронну мережу з метою бінарної класифікації щодо наявності або відсутності танків серед дерев. Незважаючи, на високі показники ефективності, отримувані під час верифікації роботи класифікатора, навчальна вибірка була сформульована неправильно, оскільки не розрізнялася репрезентативністю умов отримання зображень. Так, нейронна мережа навчилася розрізняти сонячні дні від пахмурних, оскільки усі зображення з прихованими танками були отримані в пахмурні дні, а зображення без них – в сонячні.

Незважаючи на широку знаменитість даної історії серед дослідників комп'ютерного зору та її цитованість, вона цілком може бути апокрифічною, проте цілком вірно ілюструю наявність такої проблеми, існування якої зокрема було підтверджене в результаті власних експериментальних досліджень.

З іншої ж сторони, існує проблема коли одні і ті ж самі дані можуть бути помилково віднесені до іншої категорії. Для прикладу, під час формування власної навчальної вибірки для ШНМ, яка б дозволила виявляти полум'я на зображеннях, було виявлено, що за умов недостатньої кількості навчальних даних, помилково можуть класифікуватися візуально

схожі об'єкти (засоби освітлення, сонячне проміння, екран телевізора чи комп'ютера в роботі тощо).

Отже, для побудови результативного класифікатора необхідно забезпечити формування високоякісної навчальної вибірки. Процес побудови такої вибірки повинен здійснюватися з врахування усіх можливих параметрів, які можуть певним чином вплинути на об'єкт дослідження. Так, вказані параметри умовно можна розділити на декілька категорій:

– цільові параметри. Параметри, які дозволяють візуально описати об'єкт класифікації та визначити його приналежність до заданого класу. Враховуючи поставлені задачі дисертаційного дослідження, то в даному випадку, цільовими параметрами для полум'я є форма, яскравість тощо;

– внутрішньокласові параметри. Параметри, які дозволяють визначають візуальну різноманітність одного і того самого класу об'єкта. В вищезазначеному контексті це може бути колір полум'я, який змінюється в залежності від типу горючої речовини та температури горіння;

– параметри спостереження. Параметри, на основі яких формується умови отримання зображення (розміщення камери, кут нахилу, якість отриманого зображення тощо). Необхідно відзначити що параметри спостереження навіть для однієї категорії об'єктів можуть дуже сильно відрізнятися, що може призвести до хорошої роботи класифікатора з одними умовами спостереження та поганої роботи за інших умов. Так, при використанні навчальних вибірок з обличчями людей, таких як LFW та FERET, можемо спостерігати цікавий факт: при навчанні одного і того самого класифікатора на одній із навчальних вибірок, класифікатор буде показувати незадовільні результати на іншій, оскільки дані вибірки формувалися на основі інших параметрів спостереження;

– випадкові параметри. Тип параметрів, який ґрунтується на врахуванні можливості виникнення на зображенні випадкових мікрошумів чи перетворень. Врахування даних параметрів вважається за важливим передусім завдяки дослідженню описаному у [45], де було виявлено, що

завдяки зміни значень певних пікселів можливо порушити правильність роботи ШНМ, тим самим заставивши її відносити один і той самий об'єкт до іншої категорії навіть за таких умов, де людина візуально не може помітити будь-які зміни.

Сам процес створення навчальної вибірки для навчання класифікатора можна розділити на декілька етапів:

- безпосередній збір графічних даних (на основі існуючих наборів зображень та шляхом синтезу власних зображень як із об'єктом спостереження так і без нього);

- фільтрація – перевірка зображень на відповідність вищезазначеним параметрам;

- штучна генерація – застосування підходів, які б дозволили розширити існуючий набір зображень за допомогою штучних перетворень.

На етапі безпосереднього збору графічних даних була використана база зображень ImageNet. Синтез власних зображень був проведений на основі сервісу пошуку зображень Google.

Додатково, з метою врахування можливих візуально схожих об'єктів було використано онлайн-відеокамери із вільним доступом. Висновок щодо їх візуальної схожості проводився на основі розробленого методу генерації кандидатів-регіонів, описаному у розділі 2.

Під час фільтрації здійснювався поділ усіх отриманих зображень на дві задані категорії – полум'я та його відсутність. Зокрема, з усього набору обиралися лише ті зображення, які є наближеними до реальних умов відеоспостереження та характеризуються достатньою якістю для здійснення пошуку об'єкту на ньому.

Штучна генерація дозволяє розширити отриманий з попередніх етапів набір зображень та врахувати можливість незначних зміщень, масштабування та інших перетворень об'єкта. З цією метою було використано поворот зображень, кадрівання та передискретизація.

3.3 Методи та критерії перевірки ефективності роботи детектора полум'я у відеопотоці

Для надання об'єктивної оцінки щодо ефективності роботи детектора полум'я, було вирішено розділити отримані результати на дві групи: ефективність роботи детектора на етапі генерування регіонів інтересу та загальна ефективність, що передбачає додаткове використання розробленої штучної ЗНМ.

Що стосується набору відеопослідовностей, безпосередньо на яких і відбувалося проведення експериментальних досліджень, то вони були умовно розподілені на 3 групи:

- відеопослідовності з високою якістю зображення та/або полум'ям великих розмірів;
- відеопослідовності з низькою якістю зображення, можливим зашумленням та/або полум'ям невеликих розмірів;
- відеопослідовності на яких відсутнє полум'я, але містять інші рухомі об'єкти, які, зокрема, візуально можуть нагадувати полум'я (захід сонця, людина в червоній футболці, світло від фар машини тощо).

Загальна кількість відеопослідовностей складає 30 шт., по 10 на кожен вищезазначену групу. Тестові екземпляри були відібрані із бази UltimateChase [46] та Youtube. Декілька екземплярів з тестової вибірки наведено на рисунку 3.4.

Висновок щодо якості роботи детектора формується на основі трьох показників:

- true positive – детектор визначив ділянку зображення, де знаходилось полум'я і відніс її до відповідної категорії;
- false positive – детектор помилково визначив ділянку зображення, де були відсутні будь-які прояви вогню як категорія «полум'я»;
- false negative – детектор не сповістив про наявність існуючого на зображенні полум'я.



а)



б)



в)

Рисунок 3.4 – Декілька екземплярів відеопослідовностей з тестової вибірки: а) відеоряд першої групи; б) відеоряд другої групи; в) відеоряд третьої групи

Для виведення кінцевої оцінки, на основі вищезазначених показників, отримуються значення метрик точності та повноти (3.1, 3.2):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (3.1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (3.2)$$

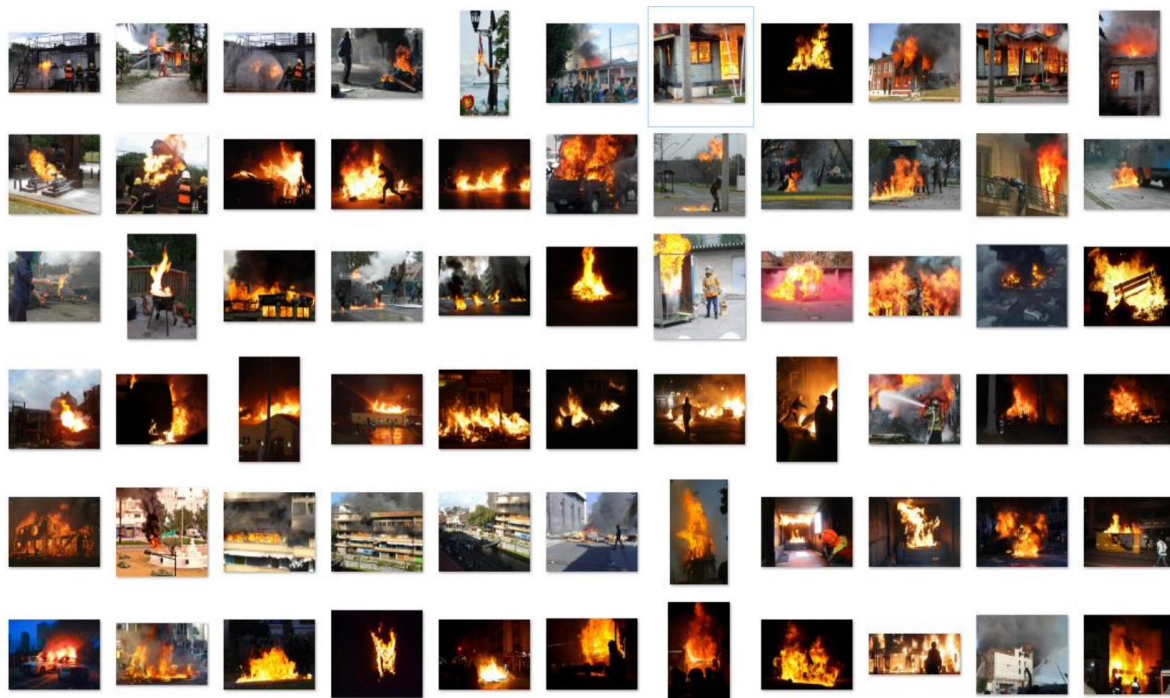
Для навчання нейронної мережі було використано власний набір зображень, процес формування якої більш детально описано в розділі 3.2. Загальна кількість фотографій складає 5000 екземплярів (рисунок 3.5).

Для оцінки розміру вибірки, який дозволить забезпечити статистично достовірні результати було використано метод розрахунку довірчого інтервалу [47].

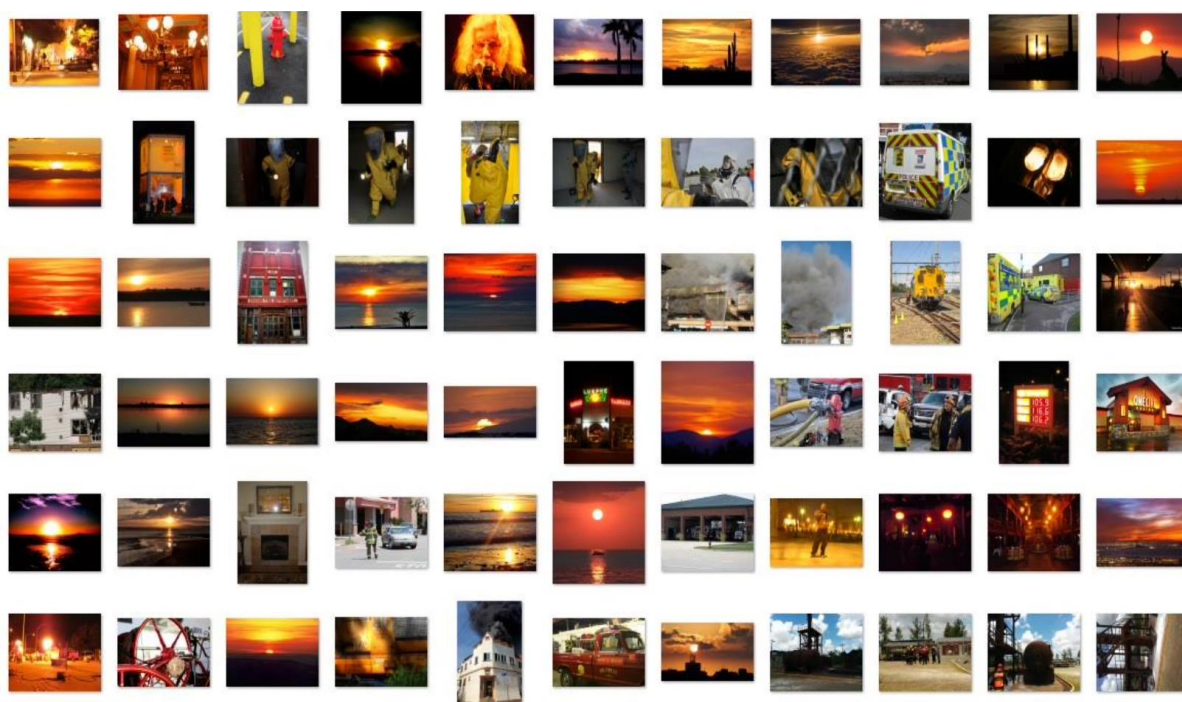
Для параметру θ – числа зображень які входять у вибірку, що являє собою випадкову величину X із рівнем довіри p , таку щоб $X = (x_1, x_2, \dots, x_n)$ довірчий інтервал представлений за допомогою таких граничних меж $l(x_1, x_2, \dots, x_n)$ та $u(x_1, x_2, \dots, x_n)$, які являються реалізаціями випадкових величин $L(X_1, X_2, \dots, X_n)$, $U(X_1, X_2, \dots, X_n)$, таких що:

$$P(L \leq \theta \leq U), \quad (3.3)$$

при цьому граничні точки довірчого інтервалу l та u являє собою довірчі межі. Рівень довіри p вибирається на основі специфіки поставленого завдання та було обрано рівним 0.95.



а)



б)

Рисунок 3.5 – Декілька екземплярів зображень з навчальної вибірки:

а) Екземпляри зображень в категорії «полум'я»

б) екземпляри зображень, які помилково можуть бути віднесені до категорії полум'я

3.4 Оцінка ефективності розроблених методів

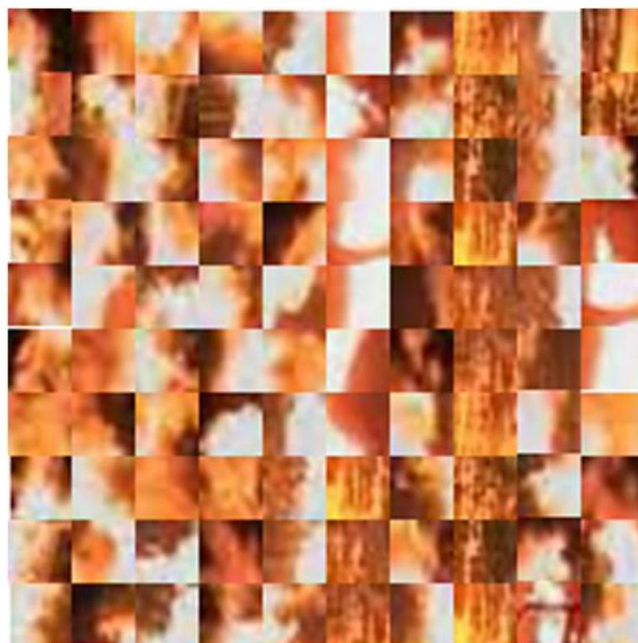
Відповідно до запропонованої архітектури метода виявлення кандидатів-регіонів, на початковому етапі відбувається колірною сегментація відповідно до виведених правил. Її реалізація спрямована на відсіювання ділянок зображення, які, з максимально можливою впевненістю, за своїми колірними характеристиками, не можуть бути віднесені до категорії «полум'я».

Відповідно до вищевказаних метрик, значення повноти під час реалізації даного метода складатиме максимально можливе – 1 (усі можливі ділянки з полум'ям були ідентифіковані). Значення точності буде знаходитися в районі 0.5-0.6, оскільки буде виникати ряд хибних викликів внаслідок можливої колірної схожості різноманітних об'єктів.

В результаті застосування метода виявлення рухомих об'єктів, значення повноти залишається не змінним, оскільки полум'я завжди супроводжується певними мікрорухами. Показник точності піднялося до 0.6-0.7, завдяки врахуванню лише рухомих об'єктів, та в середньому на 0.05 завдяки врахуванню напрямку руху вогню.

Для навчання ЛБШ була використана власно сформована група зображень із навчальної вибірки. Основною метою навчання було підвищення показника точності за умови дотримання максимально можливого значення повноти. Відповідно до поставленої мети, використовувалися зображення з різноманітними проявами вогню, полум'я під різноманітним кутом та інтенсивністю світла. Декілька екземплярів даної вибірки відображені на рисунку 3.6.

Використання локальних бінарних шаблонів знову ж дозволило зберегти значення повноти максимально можливим. Показник точності в результаті використання даного методу піднявся на 0.05.



а)



б)

Рисунок 3.6 – Навчальна вибірка зображень:

а) де міститься полум'я;

б) де полум'я відсутнє

Відповідно, завдяки комплексному використанню запропонованих підходів до генерування кандидатів-регіонів вдалося зберегти максимальний показник повноти з однієї сторони, а з іншої – досягнути високого значення точності на відеопотоках з хорошою якістю зображення – 0.85, дещо меншою у відеопотоках з можливим зашумленням/низькою якістю – 0.75.

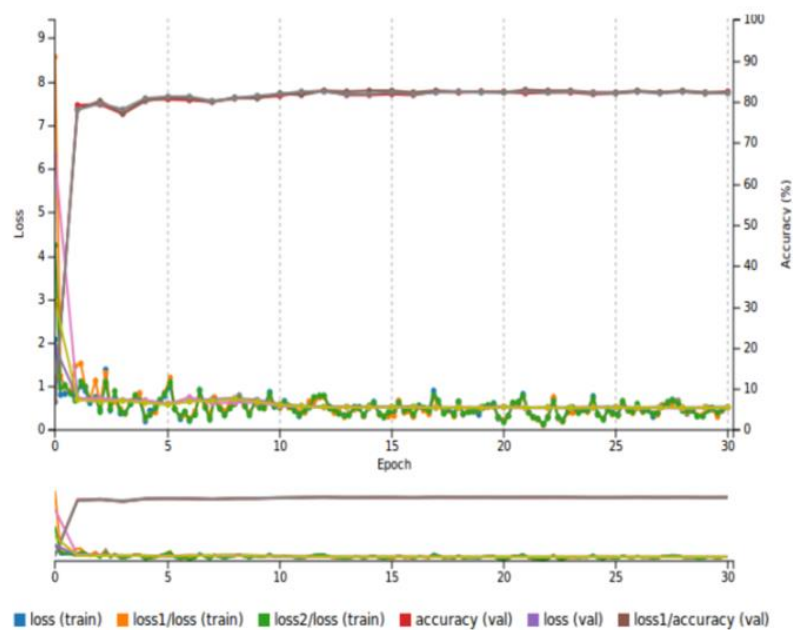
У відеопотоках, де знаходяться об'єкти, які є візуально схожі до полум'я спостерігається значна кількість хибних викликів, що і зумовлює подальше використання згорткових нейронних мереж.

В рамках експерименту, окрім розробленої архітектури ЗНМ, здійснено співставлення отриманих результатів шляхом використання популярних нейронних мереж таких як AlexNet та GoogleNet.

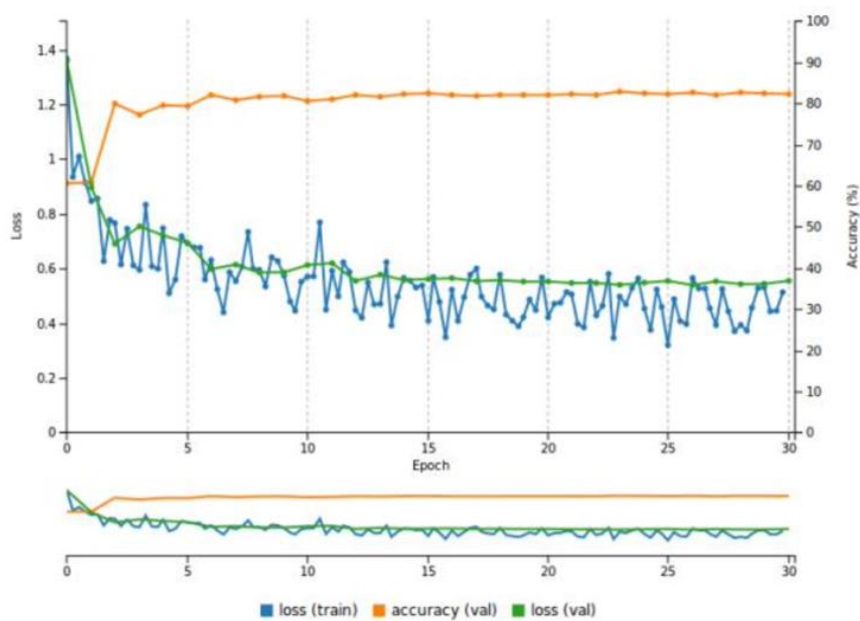
Результати навчання зазначених нейронних мереж наведено на рисунку 3.7.

Можна спостерігати, що в середньому показник точності, на основі використанню базових нейронних моделей, в середньому складає 83%, що є доволі високим результатом. В результаті використання модифікованої моделі Squeezenet вдалося не тільки зменшити кількість параметрів без втрати показників точності, а й розмір самої моделі ~7 Mb. Для порівняння розмір моделі GoogLeNet складає 41 MB, а моделі VGG-16 цілих 528 MB.

Загальний показник ефективності роботи нейронної мережі SqueezeNet без тонких налаштувань складає 83.5 %. Завдяки використанню трансферного підходу було проведено донавчання мережі шляхом додавання додаткових ваг на виходах з мережі, що в значній мірі дозволило відсіяти візуально схожі об'єкти. Для цього використовувалася попередньо натренована нейронна модель на базі зображень ImageNet, що дозволило врахувати близько 1000 додаткових категорій. Загальна ефективність після застосування трансферного підходу складала 89 %.



а)



б)

Рисунок 3.7 – Результати навчання базових нейронних мереж:

- а) мережа на основі AlexNet;
- б) мережа на основі GoogleNet

З метою підвищити узагальнюючу здатність мережі щодо детектування полум'я до навчальної вибірки застосовувався ряд трансформаційних операцій: поворот зображень, передискретизація та кадрування. Відповідно, з допомогою даних операцій, даний показник ефективності додатково на 2,4 % (таблиця 3.1).

Таблиця 3.1 – Ефективність окремих етапів тонкого налаштування нейронної мережі

Етап тонкого налаштування мережі (Fine tuning)	Відсоток покращення роботи системи ідентифікації полум'я
Поворот усіх зображень	0,7
Кадрування та попередня дискретизація	0,4
Зменшення кроку навчання	0,5
Використання ансамблевого підходу	0,8

На початковому етапі було вирішено навчати мережу з випадковими поворотами 0° , 75° , 285° , проте це не дало будь-яких покращень. Лише за умов використання усередненого значення отриманого від класифікатора під час повороту зображень в зазначених напрямках (значення були підібрані експериментальним шляхом) була отримана змога покращити ефективність роботи системи ідентифікації полум'я (таблиця 3.1).

Що стосується передискретизації та кадрування зображень, то процес їх використання нагадує операцію з поворотами зображення.

Так, вхідне зображення поділяється на 5 частин (чотири кутових та одне центральне), до яких і застосовується операція передискретизації. Отримуванні результати з кожної частини зображення знову ж усереднюються, внаслідок чого отримується кінцеве передбачення.

Необхідно також відзначити про те, що усі моделі на певному етапі навчання зіткнулися з проблемою перенавчання (overfitting), що було виявлено у зв'язку з початком зростання втрат на валідаційному наборі даних (250 зображень). З метою вирішення даної проблеми, починаючи з даного етапу, показник швидкості навчання був зменшений в 10 раз (в порівнянні з початковим).

Прикінцевим етапом доналаштування детектора полум'я стало реалізація ансамблю з нейронних мереж. Для цього була навчена модель, яка навчалася лише на базі даних зображень із полум'ям (без застосування підходу трансферного навчання).

Незважаючи на те, що така модель характеризувалася нижчою точністю, поєднавши її з іншими тонко налаштованими моделями вдалося покращити загальний показник ефективності системи ідентифікації полум'я. На мою думку даний факт зумовлений тим, що мережа натренована лише на датасеті який містить з собі зображення з вогнем, врахувала інші ознаки об'єкту аніж тонко налаштована мережа.

Додатково до ансамблю ввійшли ще дві нейронні моделі, які містили уже в собі ознаки отримувані в результаті трансферного навчання:

- перша модель донавчалася внаслідок штучного збільшення існуючого датасету шляхом повороту зображень;
- друга модель на датасеті збільшеного шляхом кадрівання та передискретизації.

Необхідно зауважити, що все ще спостерігаються не спрацювання класифікатора, особливо протягом перших секунд від моменту початку пожежі, що обумовлено особливостями використання детектора руху.

Зокрема, за умов різкої зміни заднього фону спостерігаються деякі помилкові спрацювання, що знову ж на нашу думку, вимагає доопрацювання методу визначення рухомих об'єктів.

Окремо необхідно відзначити про підходи, використання яких не покращило ефективність роботи нейронної мережі: штучне збільшення

датасету (масштабування, перекося, випадкові зрушення), поділ датасету на день/ніч (результати вимагають доопрацювання оскільки зображення на яких міститься полум'я в нічний час становить лише 17% від загальної кількості зображень), використання іншої архітектури SqueezeNet (dense→sparse→dense [48]), додавання додаткових моделей до ансамблю.

В результаті проведених експериментальних досліджень зроблені висновки, що внаслідок використання запропонованих підходів до генерування ділянок зображень, які візуально нагадують полум'я, вдалося досягнути високого показника повноти з однієї сторони, а з іншої – досягнути високого значення точності на відеопотоках з хорошою якістю зображення – в середньому 0.85, дещо меншою у відеопотоках з можливим зашумленням/низькою якістю – 0.75. У відеопотоках, де знаходяться об'єкти, які є візуально схожі до полум'я спостерігається значна кількість хибних викликів, що і зумовлює використання згорткових нейронних мереж.

ВИСНОВКИ

У кваліфікаційній роботі магістра розв'язано актуальне науково-прикладне завдання розроблення нейромережевої системи та методів ідентифікації полум'я у відеопотоках для автоматизованого вироблення оптимальних дій у випадках виникнення пожежі. Відповідно до сформульованих в першому розділі задач можна зробити наступні висновки:

– на основі проведеного аналізу літературних джерел виявлено переваги і недоліки існуючих систем та методів виявлення полум'я. Запропонована концепція використання засобів відеоспостереження для знаходження та супроводу полум'я, яка дозволить мінімізувати необхідні часові затрати для здійснення оповіщення про пожежу;

– розроблено нейромережеву систему ідентифікації полум'я на основі генерування ділянок зображень, які візуально нагадують полум'я та подальшої їх верифікації за допомогою згорткових нейронних мереж. Особливістю розробленої системи є орієнтація на роботу з зображеннями з низькою роздільною здатністю та врахування реальних умов роботи з камерами відеоспостереження, з метою мінімізації кількості можливих хибних викликів;

– розроблено метод генерування ділянок зображення, які за візуальними ознаками можуть нагадувати процес горіння, на основі використання колірної моделі L^*a^*b , методу різниці кадрів та локальних бінарних шаблонів. Експериментальні результати свідчать, що середній показник точності запропонованого методу на зображеннях з хорошою якістю – 85 %, а у випадку з можливим зашумленням/низькою якістю вхідного зображення – 75 %;

– розроблено метод верифікації отриманих ділянок зображень з використанням глибоких згорткових нейронних мереж, що завдяки структурній модифікації нейронної мережі дав змогу покращити процес

ідентифікації полум'я в складних умовах. Середній показник ефективності нейронної мережі без додаткових налаштувань склав 83.5%;

– здійснено програмну реалізацію детектора полум'я у відеопотоці, яка, завдяки реалізації ансамблевого підходу (генерування ділянок зображень та їх подальша класифікація), дозволяє працювати в online режимі, забезпечити ефективне виявлення та супровід об'єкта спостереження. Результати експериментального дослідження свідчать, що середнє значення точності ідентифікації на зображеннях попередньо навченою моделлю складає 86%. Середнє значення точності ідентифікації у відеопотоці (враховуються особливості руху) 92%.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Agrawal D., Mishra P. Smoke Detection using Local Binary Pattern. *IEEE International Symposium on Industrial Electronics (ISIE) International Journal of Current Engineering and Technology*. No 4(6). 2014. Pp. 4052- 4056.
2. Bo Jiang, Yongyi Lu, Xiyang Li, Liang Lin Towards a Solid Solution of Real-time Fire and Flame Detection. *Multimedia Tools and Applications*, No 74 (3). 2015. Pp. 689–705.
3. Boureau Y-L., Bach F., LeCun Y., Ponce J. Learning Mid-Level Features For Recognition *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. 2010*. Pp. 2559-2566.
4. Boureau Y-L., Ponce J., Lecun Y A Theoretical Analysis of Feature Pooling in Visual Recognition. *27th International Conference On Machine Learning. 2010*. Pp. 111-118.
5. Celik T. Fast and Efficient Method for Fire Detection Using Image Processing. *ETRI Journal F*. No 32(6). 2010. Pp. 881-890.
6. Celik T., Demirel H., Ozkaramanli H., Uyguroglu M. Fire detection using statistical color model in video sequences. *Journal of Visual Communication and Image Representation*. No 18(2). 2007. Pp. 176-185.
7. CIFAR-10 and CIFAR-100 datasets. URL: <https://www.cs.toronto.edu/~kriz/cifar.html> (дата звернення 14.04.2022).
8. Deng J., Dong W., Socher R., Li Li-Jia, Li Kai, Li Fei-Fei ImageNet: A Large- Scale Hierarchical Image Database. 2009 IEEE conference on computer vision and pattern recognition. 2009. Pp. 248-255.
9. Emmy Prema C., Vinsley S. S., Suresh S. Multi Feature Analysis of Smoke in YUV Color Space for Early Forest Fire Detection. *Fire Technology*. 2016. T. 52, No 5. Pp. 1319–1342.
10. Everingham M., Luc Van Gool, Williams C., Winn J., Zisserman A. The PASCAL Visual Object Classes (VOC) Challenge. *Int J Comput Vis*. 2009. Pp. 303–338.

11. Fire protection association report. URL: <https://www.nfpa.org/-/media/Files/News-and-Research/Fire-statistics-and-reports/Detection-and-signaling/ossmokealarms.pdf> (дата звернення 14.04.2022).

12. Foggia P., Saggese A. and Vento M., "Real-Time Fire Detection for Video- Surveillance Applications Using a Combination of Experts Based on Color, Shape, and Motion," *IEEE Transactions on Circuits and Systems for Video Technology*, No 25(9). 2015. Pp. 1545-1556.

13. Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, Kurt Keutzer, «SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,» ArXiv, 2016.

14. Fraser N. Neural Network Follies. 1998. URL: <https://neil.fraser.name/writing/tank/> (дата звернення 11.04.2022).

15. Gomes P., Santana P., Barata J. A Vision-Based Approach to Fire Detection. *International Journal of Advanced Robotic Systems*. т. 11. No 9. 2014. Pp. 1–12.

16. Gunn R. Support Vector Machines for Classification and Regression. 1998. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.7171&rep=rep1&type=pdf> (дата звернення 24.04.2022).

17. Guo Y., Liu Y., Oerlemans A., Lao S., Wu S., S. Lew M. Deep learning for visual understanding: A review. *Neurocomputing*, No 187. 2016. Pp. 27-48.

18. Han S., Liu X., Mao H., Pu J., Pedram A., Horowitz M. Dally W. Eie: Efficient inference engine on compressed deep neural network. URL: <https://arxiv.org/pdf/1602.01528.pdf> (дата звернення 04.04.2022).

19. Haralick R., Sternberg S., Xinhua Zhuang Image Analysis Using Mathematical Morphology. *IEEE Transactions On Pattern Analysis And Machine Intelligence*. No 9(4). 1987. Pp. 532-550.

20. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. 2015. URL: <https://arxiv.org/pdf/1512.03385.pdf> (дата звернення 24.04.2022).

21. He K., Zhang X., Ren S., Sun J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. 2014. URL: <https://arxiv.org/pdf/1301.3557.pdf> (дата звернення 08.04.2022).

22. Hemangi Tawade, Patane R.D. Optimized fire detection using image processing based techniques. *International Journal of Science, Engineering and Technology Research*. 2015. No 4(6), pp. 2288-2292.

23. Krizhevsky A., Sutskever I., Hinton G. ImageNet classification with deep convolutional neural networks. *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing System*. No. 1. 2012. Pp. 1097- 1105.

24. Kumarguru Poobalan, Siau-Chuin Liew. Fire detection algorithm using image processing techniques. *3rd International Conference on Artificial Intelligence and Computer Science (AICS2015)*. 2015. Pp. 160-168.

25. Kurup A. R. Vision based fire flame detection system using optical flow features and artificial neural network. *International Journal of Science and Research (IJSR)*, т. 3. No 10. Pp. 2161-2168.

26. LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D. Handwritten digit recognition with a back-propagation network. *Advances in Neural Information Processing Systems*. 1990. Pp. 396-404.

27. Li P., Zhao W. Image fire detection algorithms based on convolutional neural networks. *Case Studies in Thermal Engineering*. No 19. 2020. 100625.

28. Maas A., Hannun A., Ng A. Rectifier Nonlinearities Improve Neural Network Acoustic Models. 2013. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.693.1422&rep=rep1&type=pdf> (дата звернення 19.04.2022).

29. Mary L., Comer and Edward J. Delp III Morphological operations for color image processing. *Journal of Electronic Imaging*. No 3(8). 1999. Pp. 279-289.

30. Nguyen T., Zong Z., Ogunbona P. Li W. Object detection using Non-Redundant Local Binary Patterns. *Image Processing (ICIP) 2010 17th IEEE International Conference on. 2010*. Pp. 4609-4612.

31. Ojala T., Pietikainen M., Maenpaa T. Multiresolution gray-scale and rotation invariant texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. No 24(7). 2002. Pp. 971-987.

32. Z., Loy Chen-C., Tang X. DeepID-Net: Deformable Deep Convolutional Neural Networks for Object Detection. 2015. URL: <https://arxiv.org/pdf/1412.5661.pdf> (дата звернення 16.04.2022).

33. Patel P., Tiwari S. Flame Detection using Image Processing Techniques. *International Journal of Computer Applications (0975 – 8887)*. 2012. No58(18). Pp. 13-16.

34. Phillips W., Shah M., Da Vitoria Lobo N. Flame recognition in video. *Applications of Computer Vision. 2020 Fifth IEEE Workshop on. 2000*. Pp. 224-229.

35. Qi X., Ebert J. A Computer Vision-Based Method for Fire Detection in Color Videos. *Spring*. т. 2. No 9. 2009. Pp. 22-34.

36. Rui Chen, Yuanyuan Luo, Mohanmad Reza Alsharif Forest Fire Detection Algorithm Based on Digital *Journal of software*, No 8 (8). 2013. Pp. 1897- 1905.

37. Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A., Fei-Fei L. ImageNet Large Scale Visual Recognition Challenge. 2015. URL: <https://arxiv.org/pdf/1409.0575.pdf> (дата звернення 24.04.2022).

38. Scholkopf B., Smola A. Learning with Kernels. 2001. URL: <http://www.leg.ufpr.br/~eferreira/CE064/learning-with-kernels.pdf> (дата звернення 17.04.2022).

39. Smoke Alarms in U.S. Home Fires. 2015. URL: [http://www.nfpa.org/News-and-Research/Fire-statistics-](http://www.nfpa.org/News-and-Research/Fire-statistics-and-reports/Fire-) and-reports/Fire-

statistics/Fire-safety-equipment/Smoke-Alarms-in-US-Home-Fires (дата звернення 26.04.2022).

40. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going Deeper with Convolutions. 2014. URL: <https://arxiv.org/pdf/1409.4842.pdf> (дата звернення 14.04.2022).

41. Szegedy C., Zaremba W., Sutskever I., Bruna J., Erhan D., Goodfellow I., Fergus R. Intriguing properties of neural networks. 2014. URL: <https://arxiv.org/pdf/1312.6199.pdf> (дата звернення 10.04.2022).

42. Tenenbaum J. B., Freeman W. T. Separating style and content with bilinear models. *Neural computation*. 2000. Pp. 1247–1283.

43. Thou-Ho (Chao-Ho) C., Ping-Hsueh W., Yung-Chuen C. An Early Fire- Detection Method Based on Image Processing. *International Conference on Image Processing*. 2004. Pp. 1707-1710.

44. Thou-Ho Chen, Cheng-Liang Kao, Sju-Mo Chang. An intelligent real-time fire-detection method based on video processing. *Security Technology Proceedings. IEEE 37th Annual 2003 International Carnahan Conference*. Pp. 104-111.

45. Toreyin B. U. Fire Detection Algorithms Using Multimodal Signal and Image Analysis. *Bilkent University*. 2009. p. 138.

46. Toreyin B. U., Dedeoglu Y., Cetin A. E. Flame detection in video using hidden Markov models. *Image Processing, 2005. ICIP 2005. IEEE International Conference on*. 2005. Pp. 1230–1233.

47. Toreyin B.U., Dedeoglu Y., Gudukbay U., Cetin A.E. Computer vision based method for real-time fire and flame detection. *Pattern Recognition Letters*. 27. No 1. 2006. Pp. 49-58.

48. Toulouse T., Rossi L., Celik T., Akhloufi M. Automatic fire pixel detection using image processing: a comparative analysis of rule-based and machine learning-based methods. *Signal, Image and Video Processing*. т. 10. No 4. 2016. Pp. 647–654.

49. Tsung-Yu L., Aruni R., Subhransu M. Bilinear CNN Models for Fine-grained Visual Recognition. 2015. URL: <https://arxiv.org/pdf/1504.07889.pdf> (дата звернення 18.04.2022).

50. Turgay C., Demirel H., Ozkaramanli H., Uyguroglu M. Fire detection using statistical color model in video sequences. *Journal of Visual Communication and Image Representation*. 2007. No 18 (2). Pp. 176-185.

51. Turgay Ç., Özkaramanlı H., Demirel H. Fire and smoke detection without sensors: image processing based approach. *15th European Signal Processing Conference*. 2007. Pp. 1794-1798.

52. Ultimate Chase Video Production Services. URL: <http://www.ultimatechase.com/> (дата звернення 04.04.2022).

53. Wang J. Research on Flame Image Detection. *3rd International Conference on Social Science (ICSS 2016)*. 2016. Pp. 125-128.

54. Wang K., Zhang Y., Wang J., Zhang Q., Chen B., Liu D., "Fire Detection in Infrared Video Surveillance Based on Convolutional Neural Network and SVM," *2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP)*. 2018. Pp. 162-167.

55. Xia J., Sun J., H. Li, "Forward-Looking Infrared Image Segmentation Using Support Vector Machine Based on Feature Extraction," *2010 International Conference on Electrical and Control Engineering*. 2010. Pp. 1031-1034.

56. Xingyu Z., Wanli O., Junjie Y., Hongsheng L., Tong X., Kun W., Yu L., Yucong Z., Bin Y., Zhe W., Hui Z., Xiaogang W. Crafting GBD-Net for Object Detection. 2016. URL: <https://arxiv.org/pdf/1610.02579.pdf> (дата звернення 15.04.2022).

57. Zeiler M., Fergus R. Stochastic Pooling for Regularization of Deep Convolutional Neural Networks. 2013. URL: <https://arxiv.org/pdf/1301.3557.pdf> (дата звернення 19.04.2022).

58. Zhao G., Pietikainen M. Dynamic Texture Recognition Using Volume Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. No 29(6). 2007. Pp. 915 – 928.