УДК 004.048, 004.855.5, 004.896, 004.931

О. Максимів[1], Т. Рак[1], Д. Пелешко[2]
[1]Львівський державний університет безпеки життєдіяльності,
кафедра управління інформаційною безпекою,
[2]Національний університет "Львівська політехніка",
кафедра інформаційних технологій видавничої справи

ЗМЕНШЕННЯ КІЛЬКОСТІ ХИБНИХ ВИКЛИКІВ ПІД ЧАС РОЗВ'ЯЗАННЯ ЗАДАЧІ ДЕТЕКТУВАННЯ ПОЛУМЯ У ВІДЕОПОТОЦІ З ВИКОРИСТАННЯМ ГЛИБОКИХ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

© Максимів О., Рак Т., Пелешко Д., 2017

Розроблено новий підхід до детектування полум'я на зображеннях, який ґрунтується на використанні згорткових нейронних мереж. Запропоновано структуру реалізації каскадного підходу до детектування вогню, яка забезпечує покращену ефективність розпізнавання на зображеннях з низькою роздільною здатністю, та об'єктів, які можуть візуально нагадувати полум'я. Проведено експерименти з дослідження запропонованого методу порівняно з сучасним методом детектування об'єктів Faster R-CNN. У результаті проведених експериментів було виявлено покращення показника ефективності в середньому на 20 %.

Ключові слова: комп'ютерний зір, нейронні мережі, детектування полум'я.

In this paper, we develop a new approach for detecting fire in images based on convolutional neural networks. Cascade structure, which provides improved efficiency of recognition in images with low resolution and objects that can visually resemble flames, was proposed. We have performed an experimental comparison with the modern method of objects detecting Faster R-CNN. As a result of the experiments, it was found that performance of fire recognition improved on average by 20 %.

Key words: computer vision, neural networks, flame detection.

Вступ

Надзвичайні ситуації можуть призвести до значних матеріальних втрат, травмування чи навіть загибелі людей. Одними із найнебезпечніших видів надзвичайних ситуацій є пожежі та вибухи, що і зумовило значний інтерес дослідників до розроблення систем автоматичного детектування полум'я та диму. Сьогодні пожежні сповіщувачі за видом контрольованої ознаки пожежі поділяють на теплові (реагують на певне значення температури і (або) швидкість її зростання), димові (реагують на аерозольні продукти горіння), полум'я (реагують на електромагнітне випромінювання), комбіновані (реагують на декілька ознак пожежі). Проте з метою оперативного реагування на надзвичайну ситуацію ці сповіщувачі повинні бути розміщені неподалік від центру виникнення вибуху або пожежі, що супроводжується складнощами під час установки та значними матеріальними затратами під час як установлення, так і подальшої експлуатації, особливо на великих об'єктах.

На відміну від давачів, камери можуть візуально охопити значно більший район та надати додаткову інформацію про те, що відбувається на об'єкті. Зокрема протягом останніх років спостерігається тенденція до запровадження систем пасивного спостереження на підприємствах, установах і в житлових приміщеннях, основою яких є камери відеоспостереження. Проте систем такого типу, які б здійснювали ефективне реагування на загрози надзвичайних ситуацій (передусім

полум’я), сьогодні в Україні не існує. Саме тому вважають за доцільне використовувати сучасні методи машинного зору і можливості відеокамер у контексті розв’язання задачі детектування полум’я.

Постановка задачі

Розглянемо рис. 1, на якому зображено полум’я на зображеннях з низькою роздільною здатністю та об’єкти, які візуально можуть нагадувати полум’я. Людина легко розпізнає на цьому наборі зображень наявність полум’я або відсутність будь-яких загроз для виникнення та розвитку пожежі. Проте сучасні системи детектування проявів надзвичайних ситуацій, які ґрунтуються на використанні алгоритмів комп’ютерного зору, у цьому випадку не розпізнають полум’я з низькою роздільною здатністю чи розпізнають об’єкти, які за текстурами можуть мати певну візуальну схожість на пожежу. Такі некоректні спрацювання, особливо за умов інтеграції розроблюваного детектора в систему безпеки життєдіяльності людини, є не допустимими.

а б

Рис. 1 Проблемні ситуації під час детектування полум’я: а – низька роздільна здатність (червоним прямокутником зазначено місця розташування полум’я); б – текстура полум’я як частина певного об’єкта

Основна задача полягає в тому, щоб з використанням сучасних методів комп’ютерного зору розробити систему детектування полум’я, яка б, з одного боку, ефективно розпізнавала вогонь на зображеннях навіть невисокої роздільної здатності, а з іншого – мінімізувала кількість хибних викликів на об’єктах, які візуально можуть нагадувати його.

Вимоги до розроблення такого методу:

– необхідно використовувати згорткові нейронні мережі, оскільки протягом останніх років вони мають набагато кращі показники ефективності детектування порівняно з іншими методами;

– розроблюваний детектор полум’я повинен характеризуватися значною швидкодією та пониженою ймовірністю віднесення об’єктів, які візуально нагадують полум’я, до категорії “полум’я”.

Аналіз існуючих методів

Узагальнюючи науковий досвід попередніх дослідників, необхідно зазначити, що незважаючи на відсутність напрацювань в Україні щодо проблематики детектування надзвичайних ситуацій у відеопотоці, закордонні науковці присвятили цьому питанню низку наукових праць. Так, існуючі роботи можна поділити за двома основними методами детектування полум’я: використання різноманітних колірних моделей та інформації щодо руху об’єктів.

Під час використання колірних моделей передбачено інтерактивну сегментацію зображення в ділянці, де спостерігається полум’я, та відповідно до отриманих статистичних розподілів пікселів формуються граничні умови, в межах яких визначається належність пікселя до шуканої області. Для такої сегментації використовують моделі: RGB [1], HSV [2], YCbCr [3], CIE Lab [4] з певними граничними межами або (та) набором правил. Зокрема відомі підходи до комбінації наборів двох і більше колірних моделей для покращення результату детектування (табл. 1).

Таблиця 1

Оцінювання ефективності використання колірних моделей у задачі детектування полум'я [5]

Умови	К-сть зображень	H	Y	R	HYR	HY
Місто	124	0.78	0.51	0.19	0.81	0.83
Сільська місцевість	73	0.73	0.57	0.17	0.77	0.79
Всередині приміщень	20	0.36	0.47	0.12	0.43	0.46
Ніч	21	0.60	0.61	0.06	0.64	0.64
Загалом	217	0.74	0.54	0.15	0.78	0.79

Результати експериментів, які наведено у табл. 1, дають можливість констатувати, що використання колірних моделей є недостатнім для побудови ефективної системи детектування полум'я на зображені. На нашу думку, передусім це зумовлено неоднорідністю кольору горіння (колір полум'я чи диму залежить від його температури та речовини, яка горить), що зумовлює використання додаткових методів комп'ютерного зору для покращення ефективності роботи детектора.

Для виявлення руху у відеопотоці широко використовують методи обчислення оптичного потоку [6] та різниці кадрів [7]. Незважаючи на те, що комбінація таких методів разом з колірними моделями дають змогу значною мірою підвищити ефективність роботи детектора полум'я все ще залишається проблема помилкових спрацювань на рухомих об'єктах, які за зовнішніми ознаками можуть нагадувати полум'я.

З моменту перемоги CNN у конкурсі класифікації зображень ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012 їх популярність значно зросла. Мережі цього типу почали широко використовувати для розв'язання задач детектування об'єктів [8], класифікації та навіть сегментації. За умов наявності значної репрезентативної навчальної вибірки CNN демонструє набагато кращі показники ефективності порівняно з традиційними методами детектування об'єктів. Саме тому ми пропонуємо використовувати глибокі згорткові нейронні мережі для вирішення завдання детектування полум'я на зображеннях.

Питанню використання НМ для детектування полум'я було присвячено роботи [9–11]. Зауважимо, що жодна з цих робіт не враховує можливості використання глибоких згорткових нейронних мереж, які сьогодні показують якнайкращі результати для детектування об'єктів на зображеннях. Проте на відміну від існуючих напрацювань постає задача не лише розроблення детектора на основі CNN, а передусім мінімізації можливих хибних викликів, що покращить загальну ефективність детектування.

Матеріали і методи

Загальна методологія детектування полум'я на зображеннях

Як уже було зазначено, CNN можуть продемонструвати кращі результати лише за умов наявності значної репрезентативної навчальної вибірки. Саме тому на початковому етапі важливою задачею є створення навчального датасету $D = \{(x, y)_i\}_{i=1}^{|D|}$, де x – вхідне зображення з параметрами $x = \Omega \to \mathbb{R}^3$, де Ω вказує на сітку зображення з розмірами $w \times h \times d$,, де d = 3. Класову належність визначають за рівнянням $y \in Y = \{0,1\}^C$, що визначає наявність або відсутність полум'я на i -му зображенні (C=2). Для кожного вхідного зображення x_i детектор використовують для відображення можливих ділянок із шуканим об'єктом із зазначенням оцінок його "впевненості" в цьому. Це може бути формалізовано так $G = \{(x(B), I)_i\}_{i=1}^{|G|}$. Своєю чергою, $B = \{b_k\}_{i=1}^{|B|}$ являє собою набір координат обмежувальних вікон (bounding box) з $b_k = [x_k, y_k, w_k, h_k] \in \mathbb{R}^4$; $I = \{s_k\}_{i=1}^{|I|}$; – оцінки, які призначені для пропозицій $x(B)$.

Здатність до узагальнення CNN можна покращити завдяки використанню уже навчених моделей замість випадкової ініціалізації вагових значень. З цією метою було використано нейронну мережу VGG-16 [12], яка натренована на більш ніж 14 млн. зображень, які належать до 1000 класів. Формально, було використано власноруч розроблену базу зображень для донавчання CNN: $\widetilde{D} = \{(\tilde{x}, \tilde{y})_i\}_{i=1}^{|\widetilde{D}|}$,

Згорткові нейронні мережі за специфікою своєї роботи містять декілька стадій обробки. Кожна стадія характеризується наявністю двох типів шарів: згортковий, що містить нелінійну функцію активації, та нелінійний підвибірковий шар. У першому шарі згортковий фільтр подається на вхід, а у другому відбувається зменшення розміру вхідних даних. Формально CNN можна описати так:

$$f\left(v, \theta^{(1)}\right) = v^* = f_{out} \circ f_L \circ \ldots \circ f_2 \circ f_1(v^{(0)}),$$

де $v^{(0)} = v$ – вхідні дані, $\circ$ – оператор суперпозиції, θ – параметри CNN. Вихід згорткової нейронної мережі можна розглядати як наближення вхідних даних (представлених v^*). Кожен шар мережі містить набір фільтрів, де кожен фільтр формально визначається як:

$$v^{(k)} = f_k\left(v^{(k-1)}\right) = \sigma(W_k(i,j)^T v^{(k-1)} + \beta_k,$$

де $\sigma(.)$ є нелінійністю і де згорткові фільтри представлені ваговою матрицею W_k і вектором зміщення β_k. Нелінійні шари підвибірки визначаються $v^{(k)} = \downarrow v^{(k-1)}$, де $\downarrow$ – функція субдискретизації, яка об'єднує значення з області вхідних даних. Шар поліноміальної логістичної регресії обчислює ймовірність класу i за допомогою ознак v^L з L-го шару з багатозмінною логістичною втратою (softmax) $y(i) = \frac{\exp(v^L(i))}{\sum_j \exp(v^L(j))}$. У нашому випадку (1) можна описати як:

$$f\left(x(B), \theta^{(1)}\right) = x(B)^* = f_{out} \circ f_L \circ \ldots \circ f_2 \circ f_1(x(B^{(0)})),$$

де входами є пропозиції (полум'я на зображенні із виділенням його границь), які позначені як $x(B^{(0)})$. Основна ідея полягає в тому, що беруть пропозиції $x(B)$, які будуть оброблятися CNN, і наводять ймовірність класифікації, що задана пропозиція містить полум'я. Пропозиції, які класифікуються як не полум'я, відкидаються, що дає змогу уникнути помилкових спрацювань детектора. Ті пропозиції, які класифікуються як не полум'я, зберігаються, включаючи надану детектором оцінку. Вихідне передбачення CNN можна подати у вигляді:

$$f(x(B), \theta) = y^*.$$

Кінцевою метою є модифікація самої структури CNN моделі для зменшення кількості хибних викликів на основі існуючої бази зображень D.

Огляд запропонованого методу

Запропонована структура мережі містить генератор гіпотез для знаходження можливих ділянок, на яких може з'явитися полум'я, та каскадний ансамбль з двох нейронних мереж для детектування полум'я великих та малих розмірів відповідно. Результати детектування проходять через порогову функцію з метою їх верифікації. За цим підходом можна врахувати можливі характеристики полум'я незалежно від якості зображення, з одного боку, та зменшити кількість хибних спрацювань детектора завдяки масштабуванню з метою урахування повної картини зображення з іншої.

Генератор гіпотез

Ідеальний генератор гіпотез повинен генерувати якомога більше пропозицій, щоб охопити всі можливі екземпляри об'єктів. На основі високої абстрактної здатності CNN, RPN здатна вловлювати схожості навіть серед неоднакових об'єктів. Внаслідок цього виникає проблема, коли під час класифікації регіонів насамперед вивчаються візуальні схожості об'єктів, що, з одного боку, дає змогу відокремити шуканий клас об'єкта від інших об'єктів (такими особливостями можуть бути колірні ділянки, гострі та замкнені краї), проте, з іншого, може призвести до ряду хибних викликів під час класифікації регіонів, які можуть візуально нагадувати об'єкт. Враховуючи різноманітність форм

полум'я, використання одного лише RPN, який запропоновано у роботі [13], під час розв'язання задачі детектування вогню, безперечно, призведе до виникнення проблеми такого роду.

Для вирішення цієї проблеми автори запропонували використовувати метод "CRAFT" [14], в основу якого покладено стратегію "поділяй та володарюй" (генерація гіпотез та класифікація), що передбачає поділ кожної задачі на дві підзадачі. За цією стратегією можна покращити результати одразу в двох основних напрямках детектування полум'я: під час генерації ділянок – отримати меншу кількість ROI для аналізу; під час класифікації – зменшити кількість хибних спрацювань.

Для навчання генератора гіпотез було перенавчено мережу VGG, яка натренована на базі Imagenet на основі власної бази зображень. Замість навчання бінарного детектора наявності/відсутності полум'я появилась змога врахувати особливості об'єктів, які можуть за зовнішніми ознаками його нагадувати (лампи освітлення, гірлянди тощо). Загальний коефіцієнт відклику всіх категорій об'єктів становить 94,87 %.

Проте, як зазначено у роботі [14], різноманітність співвідношення сторін та масштабу в багатьох випадках не дає змоги RPN робити правильні припущення. Саме тому передбачено використання додаткової двокласової мережі Fast R-CNN, в якій використакно вихід RPN як навчальні даніх. За умов використання такої каскадної структури, RPN приймає універсальні ознаки зображення як вхідні дані і відповідає за виділення загальних патернів – таких, як текстури об'єкта. Своєю чергою, FRCN на вході отримує регіони, які мають візуальну схожість з об'єктами та вивчає їх детальніше, щоб відділити їх від заднього фону.

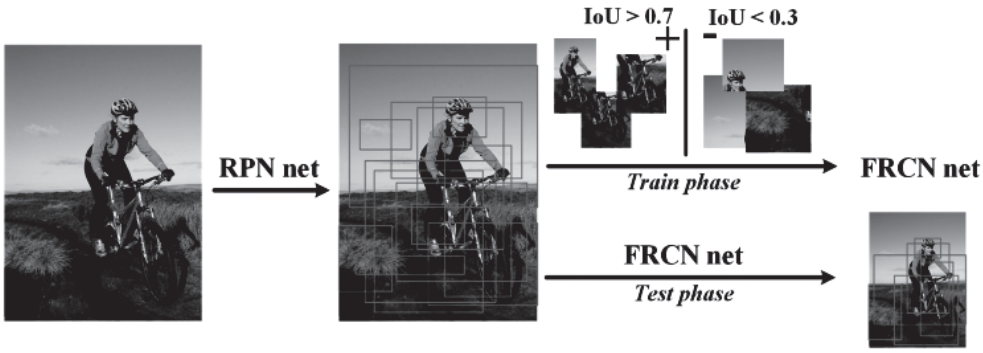

Рис. 1 Каскадний підхід на етапі генерації можливих гіпотез [14]

Детектування об'єкта

Аналізуючи набір зображень, на яких попередньо розроблені детектори полум'я [15] здійснювали помилкові спрацювання або не спрацювали за умов наявності вогню, було виділено дві ключові проблеми, за умов вирішення яких можливо забезпечити надійну ефективну роботу цього детектора. Перша проблема полягає у низькій якості як усього зображення, так і самого осередку полум'я, яке може находитися на значній відстані від місця розташування камери відеоспостереження і відповідно характеризуватиметься абсолютно інакшими ознаками (рис. 1, а). Друга проблема – неможливість врахування повної картини зображення традиційними методами детектування (рис. 1, б).

Для вирішення цих проблем було запропоновано використовувати додаткову CNN з метою перевірки правильності інформації наданої генератором гіпотез, про можливість наявності об'єкта. Структура даної мережі (рис. 2), з метою розв'язання проблеми низького розширення полум'я на зображенні поділяється на дві підмережі.

Для оцінювання розмірів полум'я з метою вибору НМ, яка повинна використовуватися, аналізується гіпотеза, яку було надано в попередньому кроці. Якщо її розмір не перевищує 1/15 висоти усього зображення, то використовується підмережа два (n_l). В інших випадках – підмережа один (n_h) :

$$\begin{matrix} n_l if\ b_b \leq \frac{1}{15} image_height \\ n_h if\ other \end{matrix} .$$

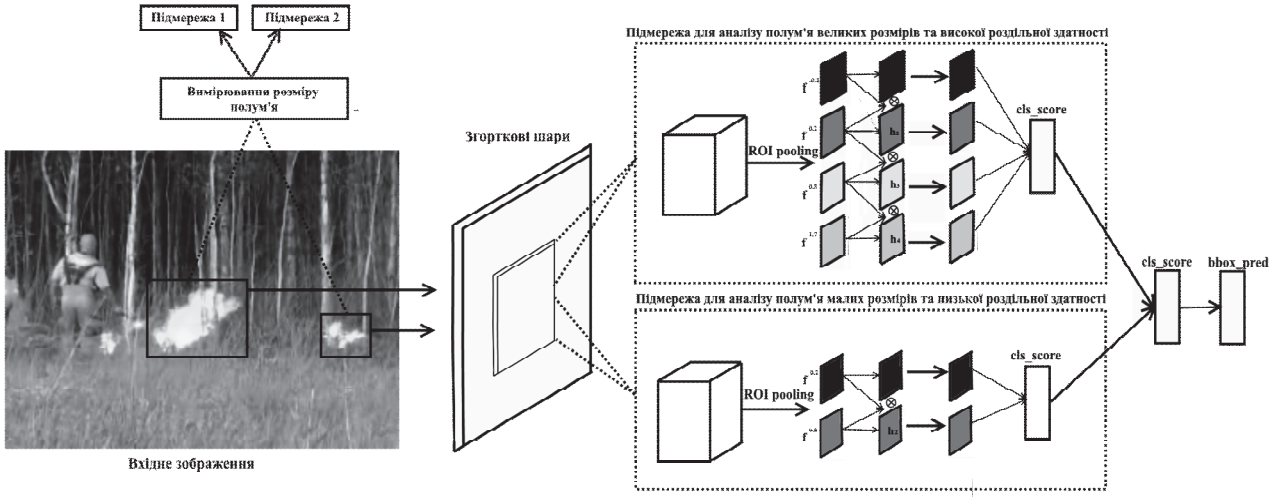

Рис. 2. Запропонована структура каскадного методу детектування полум'я на основі згорткових нейронних мереж з використанням порогових функцій

Перша мережа визначає наявність полум'я, враховуючи його особливості на зображеннях з високою роздільною здатністю. Своєю чергою, друга мережа враховує особливості полум'я на зображеннях з низькою роздільною здатністю. Кожна з мереж приймає як вхідні дані карту ознак, отриманих з попередніх шарів мережі Fast R-CNN на етапі генерації гіпотез, та закінчується вихідним шаром, який оцінює можливість наявності полум'я на зображенні та за умов наявності якого виробляються два вихідні вектори для кожного пропонованого об'єкта. Перший вектор виводить оцінку класифікації відповідно до класу об'єкта та заднього фону. Другий використовують для виведення обмежувальної рамки щодо розміщення об'єкта. Шар підвибірки ROI (ROI pooling layer) використовують для поєднання карти ознак у вектор фіксованої довжини.

З метою вирішення проблеми неможливості врахування усієї картини зображення пропонується введення так званих порогових функцій на основі пропонованих ділянок з різними розмірами та опорними регіонами ($f^{-0.2}, f^{0.2}, f^{0.8}, f^{1.7}$). Приклад отримання таких ознак CNN проілюстровано рис. 3.

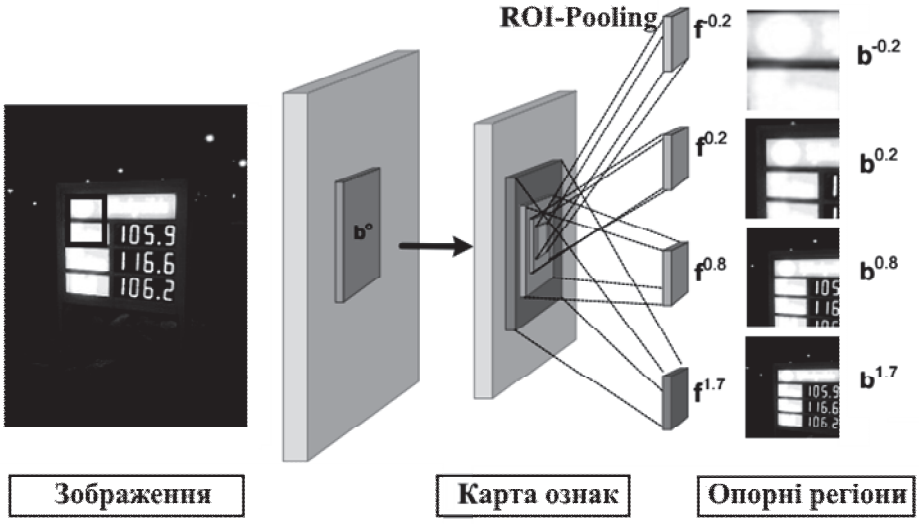

Рис. 3. Приклад використання ROI-pooling для отримання ознак CNN з різними розмірами і опорними регіонами. Червоним прямокутником позначено можливе розміщення полум'я, надане генератором гіпотез. ROI-pooling надає чотири опорні ділянки для $\{b^p\}$

Для передавання повідомлень між функціями було використано двосторонню структуру з пороговими функціями [16]. Після цього всі ознаки h_i^3 для $i = 1,2,3,4$ проходять через низку шарів CNN із загальними параметрами для отримання результату, що дає змогу здійснити прогнозування об'єкта до класу y , де $y \in Y = \{0,1\}^C$.(рис. 4).

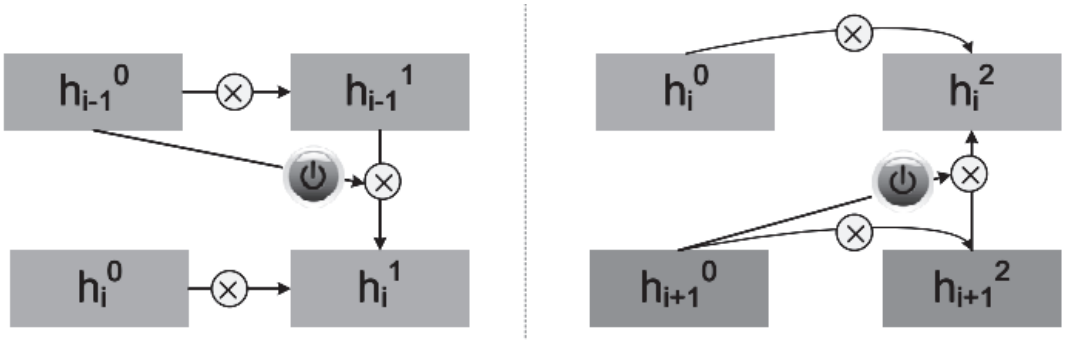

Рис. 4. Двостороння структура з пороговими функціями. ⊗ *– передбачає процес згортання, кнопка перемикання відповідає за порогову функцію [16]*

Експерименти

Для навчання мережі та оцінювання ефективності запропонованого підходу до детектування полум'я на зображенні була використана розроблену авторами базу зображень. База містить 2647 зображень з полум'ям, які, своєю чергою, поділено на дві категорії: малих розмірів (1078 зображень) та великих розмірів (1569 зображень). Зокрема під час навчання було використано зображення, які за візуальними ознаками можуть його нагадувати (засоби освітлення, гірлянди тощо). Ділянки, де передбачено розміщення полум'я, було проанотовано за допомогою обмежувальних рамок. У результаті цього було сформовано набір даних (табл. 2), який, своєю чергою, було поділено на навчальний (train), перевірковий (val) та тестовий (test).

Таблиця 2

Розподіл зображень, які використовувалися для навчання та перевірки роботи детектора полум'я за розмірами

Розміри полум'я	Загальна к-сть	train	val	test
Великі	1969	1549	210	210
Малі	678	478	100	100

Результати

Для порівняння ефективності запропонованого методу було додатково використано метод Faster R-CNN, який на момент написання роботи є провідним та найбільш вживаним методом для розв'язання задач детектування об'єктів. Результати цього експерименту зображено на рис. 5.

Обговорення

На рис. 6 наведено приклади неправильних спрацювань розробленого детектора.

За вищенаведеними результатами можемо спостерігати, що під час роботи розробленого детектора полум'я все ще спостерігаються певні помилкові спрацювання/неспрацювання. Вирішити цю проблему, на нашу думку, можливо з використанням додаткових алгоритмів визначення рухомих об'єктів та збільшенням існуючої вибірки зображень полум'я з низькою роздільною здатністю, оскільки існуюча вибірка є надзвичайно малою для забезпечення ефективного процесу навчання нейронної мережі. Проте необхідно зазначити, що на переважній більшості таких зображень людині самій важко розпізнати наявність на них полум'я. Перевагою є відсутність

викликів детектора на об’єкти, які можуть містити текстуру полум’я, проте не містять загрози для життя людини чи матеріальним цінностям.

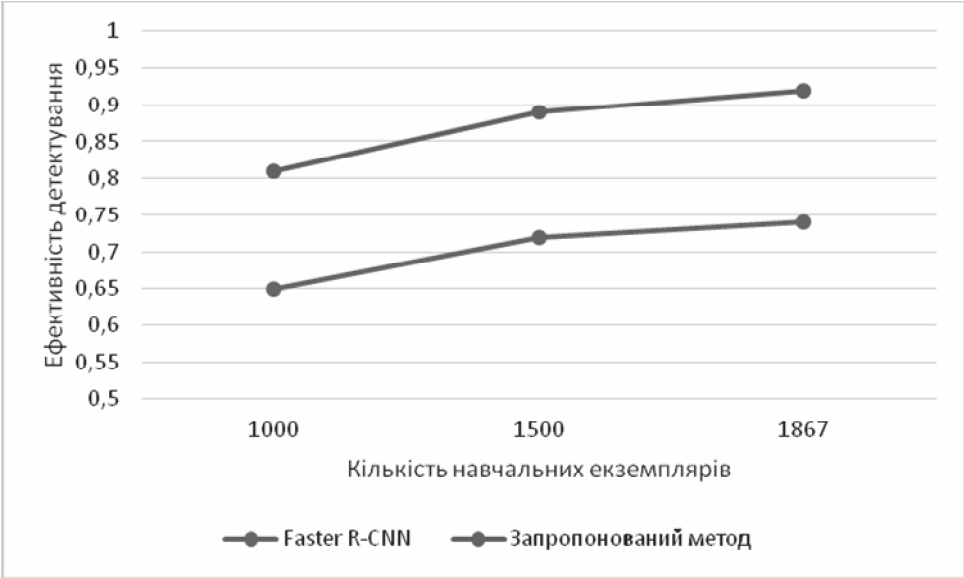

Рис. 5. Порівняння ефективності запропонованого методу з Faster R-CNN

Рис. 6. Приклади неправильної роботи детектора.
Жовтим кольором позначено ділянки, на яких полум’я знаходиться,
проте детектор не спрацював

Висновок

Наведено теоретичне та експериментальне обґрунтування використання згорткових нейронних мереж для розв’язання задачі детектування полум’я на зображеннях. Запропонований метод модифікації CNN спрямований на вирішення проблеми якісної класифікації вогню з низькою роздільною здатністю та мінімізації хибних викликів на об’єктах, які можуть візуально його нагадувати. Проведені експерименти для порівняння ефективності розробленого методу з сучасним методом для детектування об’єктів Faster R-CNN дають змогу стверджувати, що запропонований метод визначення наявності полум’я на зображенні характеризується набагато кращими показниками ефективності. Так, ефективність методу Faster R-CNN становить лише 0,69 % при навчальній вибірці 1867 зображень. Ефективність розроблюваного методу 0,91 % за тієї самої кількості зображень. Під час подальших досліджень передбачено розширенні існуючої бази зображень, оскільки вона є надзвичайно малою для навчання якісної нейронної мережі.

1. Che-Bin L. Vision based fire detection / L. Che-Bin, N. Ahuja,. // Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. – 2004. – № 4. – P. 134–137. 2. Qi X. A Computer Vision-Based Method for Fire Detection in Color Videos / X. Qi, J. Ebert. // International

journal of imaging. – 2009. – №9. – P. 22–34. 3. Celik T. Fire and smoke detection without sensors: image processing based approach / T. Celik, H. Ozkaramanli, H. Demirel. // 15th European Signal processing conference. – 2007. – P. 1794–1798 4. Celik T. Fast and efficient method for fire detection using image processing / T. Celik. // ETRI journal. – 2010. – №6. – P. 881–890. 5. Gomes P. A Vision-based Approach to Fire detection / P. Gomes, P. Santana, J. Barata. // International Journal of Advanced Robotic Systems. – 2014. 6. Horn B. K. P., Schunck B. G. Determining optical flow //Artificial intelligence. – 1981. – T. 17. – №. 1–3. – P. 185-203. 7. Zhan C. et al. An improved moving object detection algorithm based on frame difference and edge detection // Image and Graphics, 2007. ICIG 2007. Fourth International Conference on. – IEEE, 2007. – P. 519–523. 8. Russakovsky O. et al. Imagenet large scale visual recognition challenge //International Journal of Computer Vision. – 2015. – T. 115. – №. 3. – P. 211–252. 9. Deep Convolutional Neural Networks for Forest Fire Detection / Z. Qingjie, X. Jiaolong, X. Liang, G. Haifeng. // International Forum on Management, Education and Information Technology Application. – 2016. – P. 568–575. 10. Caixia C. One Fire Detection Method Using Neural Networks / C. Caixia, S. Fuchun, Z. Xinquan. // Tsinghua science and technology. – 2011. – №16. – P. 31–35. 11. Banghua Y. et al. Recognition of fire detection based on neural network //Life System Modeling and Intelligent Computing. – Springer Berlin Heidelberg, 2010. – P. 250–258. 12. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015. 13. Ren S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks // Advances in neural information processing systems. – 2015. – P. 91–99. 14. Yang B. et al. Craft objects from images //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – P. 6043–6051. 15. Maksymiv O., Rak T., Menshikova O. Deep convolutional network for detecting probable emergency situations // Data Stream Mining & Processing (DSMP), IEEE First International Conference on. – IEEE, 2016. – P. 199–202. 16. Zeng X. et al. Gated bi-directional cnn for object detection //European Conference on Computer Vision. – Springer International Publishing, 2016. – P. 354–369.