

A. P. KOVTUNENKO, O. V. YAKOVLEVA, V. A. LYUBCHENKO, O. V. YANGOLENKO

ДОСЛІДЖЕННЯ СУМІСНОГО ВИКОРИСТАННЯ МАТЕМАТИЧНОЇ МОРФОЛОГІЇ ТА ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ЦІННИКІВ

Робота присвячена вирішенню задачі розпізнавання зображень, що містять інформацію символічного типу, штрих коди, логотипи, або інші знаки. Прикладом таких зображень є цітники в торговельних центрах, флаера, запрошення, білети на різні заходи. Інформація на таких зображеннях має різний тип і її розпізнавання потребує відмінних підходів. В роботі розглядалось питання розпізнавання ціників в торговельних мережах. Для розпізнавання елементів зображення суттєву роль відіграє точність їх детектування. Було досліджено поєднання класичних методів аналізу зображень та нейромережевого підходу. Особливу увагу було приділено дослідженню у порівняльному аспекті детектування об'єктів методами морфології та шляхом обробки згортковою нейронною мережею. Дослідження показали, що морфологія дає значно нижчу якість детектування ніж нейронна мережа, але у декілька разів перевищує її у швидкодії. Оскільки швидкодія має велике значення для реалізації алгоритмів на мобільних пристроях, до морфології була додана обробка додатковими фільтрами та нормалізація геометричних спотворень, що суттєво поліпшило точність детектування та подальшого розпізнавання. За результатами досліджень питань детектування та розпізнавання штрих коду і символічної інформації, що присутня на ціниках, зроблено висновки щодо обрання підходів та технологій для вирішення цих задач, розроблено алгоритм та на його основі застосунок для розпізнавання ціників різних торговельних мереж. Також була розроблена мобільна версія застосунку. Алгоритм побудовано таким чином, що першим кроком є детектування опорного елемента, наприклад, штрих коду, далі відносно опорного елемента відбувається детектування інших елементів ціника. Штрих код детектується за допомогою математичної морфології та методів математичної статистики, яка використовується для підвищення точності алгоритму, або за допомогою згорткових нейронних мереж. Для детектування ціни та назви товару використовується згорткова нейронна мережа CRAFT, що вмє обробляти зображеннями низької якості. Знайдені назва та ціна нормалізуються для усунення геометричних спотворень та передаються для розпізнавання бібліотеці Tesseract. Ця бібліотека працює з багатьма мовами та знаходиться у відкритому доступі. Застосунок для розпізнавання ціників був створений мовою C++ з використанням бібліотек OpenCV, ZXing, Libtorch, Tesseract.

Ключові слова: розпізнавання зображень, детектування об'єктів, морфологія, нормалізація геометричних перетворень, згорткова нейронна мережа, рекурентна нейронна мережа, навчання нейронної мережі, торговельна мережа, цітники, штрих код, програмний застосунок.

A. P. KOVTUNENKO, E. V. YAKOVLEVA, V. A. LYUBCHENKO, O. V. YANGOLENKO

ИССЛЕДОВАНИЕ СОВМЕСТНОГО ИСПОЛЬЗОВАНИЯ МАТЕМАТИЧЕСКОЙ МОРФОЛОГИИ И СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ЦЕННИКОВ

Работа посвящена решению задачи распознавания изображений, содержащих информацию символического типа, штрих коды, логотипы и другие знаки. Примером таких изображений являются ценники в торговых центрах, флаера, приглашения, билеты на различные мероприятия. Информация на таких изображениях разнотипная и ее распознавание требует различных подходов. В работе рассматривался вопрос распознавания ценников в торговых сетях. Для распознавания элементов изображения существенную роль играет точность их детектирования. Было исследовано сочетание классических методов анализа изображений и нейросетевого подхода. Особое внимание было уделено исследованию в сравнительном аспекте детектирования объектов методами морфологии и путем обработки сверточной нейронной сетью. Исследования показали, что морфология дает значительно более низкое качество детектирования, чем нейронная сеть, но в несколько раз превышает ее в скорости. Поскольку быстрдействие имеет большое значение для реализации алгоритмов на мобильных устройствах, к морфологии была добавлена постобработка дополнительными фильтрами и нормализация геометрических искажений, что существенно улучшило точность детектирования и последующего распознавания. По результатам исследований детектирования и распознавания штрих кода и символической информации, присутствующей на ценниках, сделаны выводы об избрании подходов и технологий для решения этих задач, разработан алгоритм и на его основе приложение для распознавания ценников различных торговых сетей. Также была разработана мобильная версия приложения. Алгоритм построен таким образом, что первым шагом происходит детектирование опорного элемента, например, штрих кода, далее относительно опорного элемента детектируются другие элементы ценника. Штрих код детектируется с помощью математической морфологии и методов математической статистики, используемых для повышения точности алгоритма, или с помощью сверточных нейронных сетей. Для детектирования цены и названия товара используются сверточная нейронная сеть CRAFT, которая умеет обрабатывать изображения низкого качества. Найденные название и цена нормализуются для устранения геометрических искажений и передаются для распознавания библиотеке Tesseract. Данная библиотека работает со многими языками и находится в открытом доступе. Приложение для распознавания ценников было создано на языке C++ с использованием библиотек OpenCV, ZXing, Libtorch, Tesseract.

Ключевые слова: распознавание изображений, детектирование объектов, морфология, нормализация геометрических преобразований, сверточная нейронная сеть, рекуррентная нейронная сеть, обучение нейронной сети, торговая сеть, ценники, штрих код, программное приложение.

A. R. KOVTUNENKO, O. V. YAKOVLEVA, V. A. LYUBCHENKO, O. V. YANHOLENKO

RESEARCH OF THE JOINT USE OF MATHEMATICAL MORPHOLOGY AND CONVOLUTIONAL NEURAL NETWORKS FOR THE SOLUTION OF THE PRICE TAG RECOGNITION PROBLEM

The work is devoted to solving the problem of recognizing images containing symbolic type information, barcodes, logos and other signs. Example of such images are price tags in shopping centers, flyers, invitations, tickets to various events. The information on such images is of a different type and its recognition requires various approaches. The work addressed the recognition of price tags in retail chains. The accuracy of object detection has a significant role for their recognition. A significant role for recognition of image elements has the accuracy of their detection. The combination of classical methods of image analysis and the neural network approach were investigated. Particular attention was paid to the study in the comparative aspect of the object detection by morphological method and by processing a convolutional neural network. Studies have shown that morphology yields a significantly lower detection quality than a neural network, but is several times faster than it. Since speed has a great importance for the implementation of algorithms on mobile devices, post-processing with additional filters and normalization of geometric distortions were added to the morphology, that significantly improved the accuracy of detection and subsequent recognition. Based on the research results of detection and recognition of barcodes and symbolic information presented on price tags, conclusions are drawn about choosing approaches and technologies for solving these problems, an algorithm has been developed and, on its basis, an application for recognizing price tags of various retail chains. A mobile version of the application has been developed

© A. P. Kовтуненко, O. V. Яковлева, V. A. Любченко, O. V. Янголенко, 2020

also. The algorithm is constructed in such a way that the first step is the detection of the supporting element, for example, a barcode, then other price tag elements are detected relative to this supporting element. The barcode is detected with the methods based on mathematical morphology and mathematical statistics which were used to improve the accuracy of the algorithm, or convolutional neural networks. To detect prices and product names, the convolutional neural network CRAFT is exploited, which can process low-quality images. The found name and price are normalized to eliminate geometric distortions and transferred to the Tesseract library for recognition. This library works with many languages and is in the public access. The price tag recognition application was created in C++ using the OpenCV, ZXing, Libtorch, Tesseract libraries.

Keywords: image recognition, object detection, morphology, normalization of geometric transformations, convolutional neural network, recurrent neural networks, neural network training, retail chains, price tags, barcode, software application

Вступ. В останнє десятиріччя в вирішенні багатьох задач комп'ютерного зору досягнуті суттєві успіхи, особливо, в області класифікації, детектування, сегментації, розпізнавання об'єктів. Особливі успіхи пов'язані із стрімким розвитком нейронних мереж, але, як і раніше, велика кількість задач потребують використання класичних методів. Незважаючи на успіхи, існує потреба у подальших дослідженнях питань детектування та розпізнавання об'єктів на зображеннях, наприклад, для вирішення конкретних задач, де на одному зображенні присутні об'єкти принципово відмінного типу (текст, штрих код, логотип та ін.) та вимагається робота в режимі реального часу. Такі задачі потребують поєднання багатьох підходів обробки та аналізу зорової інформації. Робота присвячена вирішенню задачі розпізнавання цінників у торговельних мережах, а саме, детектуванню на них штрих коду, ціна та назви товару.

Аналіз проблеми та обґрунтування теми дослідження полягає у наступному:

- відсутня єдина база цін товарів різних торговельних мереж;
- складність автоматичного контролю коректності цін на цінниках товарів у торговельному залі, або порівняння ціни однакових товарів у різних торговельних мережах;
- відсутність єдиного стандарту цінника (див. рис. 1);
- складність детектування та розпізнавання об'єктів на зображенні за наявності значних перешкод.



Рис. 1. Приклади цінників різних форматів

Постановка задачі досліджень. Метою даної роботи є дослідження сумісного використання математичної морфології та згорткових нейронних мереж для вирішення задачі розпізнавання цінників у торговельних мережах.

Для досягнення мети необхідно було розглянути такі питання:

- розробка підходу для розпізнавання цінників на зображенні;

- детектування штрих коду (мат. морфологія, нейронна мережа);
- підготовка даних для навчання мережі для пошуку штрих кодів;
- навчання нейронної мережі для пошуку штрих кодів;
- порівняння результатів пошуку штрих кодів методами математичної морфології та нейронною мережею;
- розпізнавання штрих коду;
- пошук та розпізнавання текстової інформації на ціннику.

На основі результатів досліджень розробити застосунок для автоматичного зчитування інформації про ціну, назву, штрих код товару з фото, яке завантажує користувач, за рахунок знаходження певних об'єктів, їх вилучення та розпізнавання.

Вимоги до завантажених зображень:

- система повинна працювати з зображеннями форматів: jpg, png, bmp, роздільна здатність яких не нижча 300 dpi;
- зображення цінників повинні бути у фокусі та займати не менше 15% зображення;
- на одному зображенні повинні бути лише однакові типи цінників (див. рис. 2).



Рис. 2. Приклад завантаженого зображення

Дослідження питання детектування штрих коду методом математичної морфології. Математична морфологія представляє собою набір алгоритмів обробки зображень, в основі яких лежить операція згортки із попередньо визначеними структурними елементами – шаблонами [1]. Структурний елемент – це матриця, яка складається з елементів $\{0; 1\}$, і визначає піксель, який обробляється, та як на нього будуть впливати пікселі з зображення по сусідству. Для виконання морфологічних операцій структурний елемент проходить ковзним вікном по усьому зображенню. В основі математичної морфології знаходяться операції ерозія (erosion) та розширення (dilation).

Ерозія – операція математичної морфології, при якій піксель у вихідному зображенні, що визначається структурним елементом, приймає значення 1 лише у

тому випадку, якщо кожен одиничний піксель структурного елемента співпадає з одиничним пікселем бінарного зображення, інакше приймає значення 0. В результаті застосування операції ерозії всі об'єкти, менші ніж структурний елемент, стираються, об'єкти, з'єднані тонкими лініями стають роз'єднаними і розміри всіх об'єктів зменшуються. Операція ерозії корисна для видалення малих об'єктів і різних шумів, але у цій операції є недолік – усі об'єкти, які залишилися, зменшуються в розмірі. Цей ефект можна прибрати, якщо після операції ерозії застосувати операцію розширення з тим же структурним елементом. Така операція часто повторювалася, тому її винесли у окрему операцію – відкриття. Відкриття фільтрує всі об'єкти, менші ніж структурний елемент, але при цьому допомагає уникнути сильного зменшення їх розміру. Ця операція ідеально підходить для видалення ліній, товщина яких менше, ніж діаметр структурного елемента. Також важливо пам'ятати, що після цієї операції контури об'єктів стають більш гладкими.

Розширення – операція математичної морфології, при якій піксель у вихідному зображенні, що визначається структурним елементом, приймає значення 1 лише у тому випадку, якщо хоча б один одиничний піксель структурного елемента співпадає з одиничним пікселем бінарного зображення, інакше приймає значення 0. Якщо до зображення застосувати спочатку операцію розширення, то вона заповнить області з малими дірками та щілинами, але це призведе до збільшення контурів об'єкта. Уникнути цього збільшення дозволяє операція ерозії, виконана відразу після операції розширення з тим же структурним елементом. Така комбінація базових операцій виділена у окрему операцію – закриття.

Одним з типових застосувань математичної морфології є виділення на бінарному зображенні компонент, у яких форма і розміри задовольняють заданим обмеженням. У таких завданнях можлива побудова структурного елемента, який після застосування до бінарного зображення, збільшить необхідну область. Оскільки штрих код складається з вертикальних ліній з проміжками, для його знаходження можна застосувати морфологічну операцію закриття. Після її застосування вертикальні лінії у штрих коді об'єднуються і утворюють прямокутник. Далі, для усунення можливого злиття з іншими елементами поряд, необхідно застосувати операції ерозії та виконати аналіз отриманих контурів по відношенню сторін для відкидання зайвих елементів. Значення розмірів структурних елементів було отримано експериментальним шляхом, вони залежать від діапазону розбігу розмірів штрих кодів на завантажених зображеннях. Тому при практичній реалізації запропонованого підходу треба для завантажених зображень встановити обмеження на частину, що займає цінник, у розмірі 15%, зображення цінників повинні бути у фокусі, на одному зображенні можуть бути лише однакові типи цінників.

Для додаткової фільтрації зайвих елементів, які за відношенням сторін мають задовільний результат,

необхідно порохувати метрики математичного очікування та дисперсії частоти переходів у кожному рядку кожного виділеного регіону з бінарного зображення, але до морфологічних операцій. Експериментальним шляхом були виявлені такі норми: математичне очікування у діапазоні від 22 до 31, дисперсія у діапазоні від 5 до 9. Цими константами можна регулювати пороги точності детектування, бо на знімках, які зроблені камерою з малою роздільною здатністю, після бінаризації втрачається багато ліній самого штрих коду, але таким чином зростає частка помилково знайдених елементів.

Оскільки на зображенні штрих коду можуть знаходитися під невеликим нахилом, це впливає на результат подальшого розпізнавання. Тому в роботі пропонується ще крок нормалізації геометричних перетворень [2].

В результаті досліджень був розроблено алгоритм на основі операцій математичної морфології для детектування штрих коду, який складається із наступних кроків:

1. Бінаризувати зображення (див. рис.3) та провести фільтрацію медіанним фільтром.
2. Виконати морфологічні операції закриття та ерозії (див. рис. 4).
3. Отримати контури, провести фільтрацію (по відношенню сторін), апроксимувати контури у прямокутники (див. рис. 4).
4. Виконати додаткову фільтрацію по математичному очікуванню (див. рис. 5).
5. Провести нормалізацію геометричних перетворень отриманих прямокутників.



Рис. 3. Вхідне зображення (зліва) та бінаризоване (справа)



Рис. 4. Результати застосування морфології (зліва) та заливки контурів після неї (справа)

Дослідження питання детектування штрих коду за допомогою згорткових нейронних мереж. Нейромережовий підхід потребує велику кількість даних для навчання. Для досліджень було створено датасет цінників з розміткою областей, де знаходилися штрих коду (див. рис. 6) [3]. Для розмітки

використовувався застосунок LabelMe [4], розмічені зображення було переведено у формат COCO [5], анотації зберігалися у одному JSON файлі.

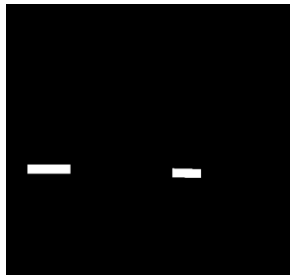


Рис. 5. Результат фільтрації за допомогою математичного очікування



Рис. 6. Приклад зображень з датасету та їх розмітка

Ідея архітектури мережі належить Zharkov A., Zagaunov I. [6]. Вона проводить пошук за допомогою семантичної сегментації. Так як штрих коди не накладаються один на одного, то задачу екземплярної сегментації вирішувати було не потрібно. Виходом мережі є зображення, у якому пікселі приймають значення від 0 до 1 і показують розподіл ймовірності належності цього пікселя до штрих коду.

В оригінальній статті не було шарів для пакетної нормалізації [7], але виникла проблема перенавчання, тому в роботі було запропоновано додати шари пакетної нормалізації для підвищення стабільності нейронної мережі. Пакетна нормалізація нормалізує вихід попереднього шару активації шляхом віднімання середнього значення пакету та ділення на стандартне відхилення:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i,$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2,$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma^2 + \epsilon}},$$

$$y_i = \gamma \hat{x}_i + \beta,$$

де B – пакет з зображеннями;
 m – кількість зображень у пакеті;
 x_i – зображення з пакету B ;
 $\epsilon > 0$ – скільки завгодно мале число, яке запобігає діленню на нуль;

γ, β – параметри нормалізації, які система буде вивчати.

Кінцева структура нейронної мережі наведена на рис. 7.

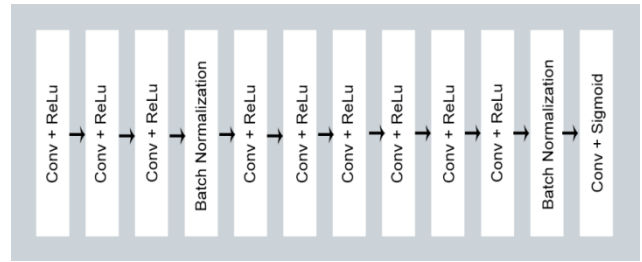


Рис. 7. Структура згорткової нейронної мережі для детектування штрих коду

Для покращення якості навчання нейронної мережі була використана аугментація (див. рис. 8). Ця техніка полягає у застосуванні заданих перетворень для зображень з датасету на кожній ітерації навчання з метою збільшення кількості різних зображень у вибірці під час навчання.

Під час навчання нейронної мережі помилка на валідаційній вибірці мала вигляд, як на рис. 9, та модель видавала результат, який наведено на рис. 10.



Рис. 8. Приклади аугментації

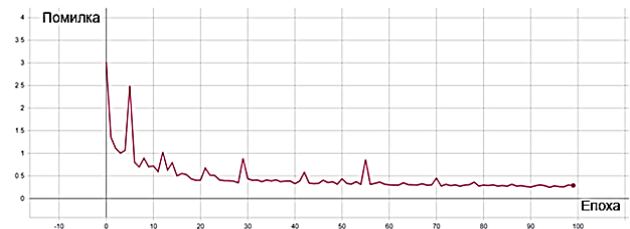


Рис. 9. Помилка на валідаційному наборі даних



Рис. 9. Вхідне зображення та вихід нейронної мережі

Порівняння результатів детектування штрих коду на основі математичної морфології та згорткової нейронної мережі. Для порівняння якості розглянутих підходів було застосовано метрики точності (precision) та повноти (recall):

$$\text{precision} = \frac{TP}{TP + FP},$$

$$\text{recall} = \frac{TP}{TP + FN},$$

де TP – кількість штрих кодів, які вірно були знайдені;

FP – кількість штрих кодів, які були знайдені помилково;

FN – кількість не знайдених штрих кодів.

Порівнювались результати детектування штрих кодів методом морфології без використання фільтрації та нормалізації наприкінці, тобто без пост обробки, та за допомогою нейронної мережі. Для порівняння використовувалися зображення, які не були присутні у тренувальному наборі даних під час навчання мережі. Після проведення замірів на центральному процесорі швидкодії було отримано результати, що наведені у табл. 1.

Таблиця 1 – Оцінювання якості та швидкодії детектування на основі математичної морфології та згорткової нейронної мережі

Метод	precision	recall	Швидкість, мс
Математична морфологія	0.48	0.7058	100
Згорткова нейронна мережа	0.9714	1.0	400

Як видно з табл. 1, морфологічна морфологія поступається у якості, але вона набагато швидша. Головною проблемою методу, який побудований на математичній морфології є те, що багато помилково виділений об'єктів. Якщо додатково проводити перевірку на декодування знайденого регіону, то можна покращити результати, але впаде швидкість, приблизно до 1 секунди на зображення, і перевага у ній буде втрачена. Недоліком згорткової нейронної мережі є те, що вона не досить точно виділяє границі об'єкта і дає малі зайві елементи, але можна додатково провести пост обробку для довіділення країв та відфільтрувати елементи по значенню їх площі. На рис. 11 представлений приклад порівняння згорткової нейронної мережі та математичної морфології.

Детектування тексту за допомогою згорткової нейронної мережі. Після аналізу існуючих згорткових мереж для детектування тексту була обрана модель CRAFT [8]. Головна мета її архітектури – точно локалізувати кожний окремий символ на зображеннях. Для вилучення характерних ознак у даній архітектурі використовується мережа VGG16 [9] з пакетною нормалізацією. Модель робить об'єднання згорткових карт низького рівня подібно до U-net [10]. Кінцевий вихід має два канали: вірогідність центру символу та оцінку зв'язності символів між собою. Оцінка центру використовується для локалізації окремих символів на зображенні, а оцінка спорідненості використовується для групування кожного символу в слова. На відміну від бінарної карти сегментації, яка позначає кожний піксель дискретно, в архітектурі CRAFT використовується кодування ймовірності центру символу за допомогою Гаусової теплової карти (див. рис. 12). CRAFT дозволяє виявити великі екземпляри тексту, незважаючи на використання невеликих рецептивних полів у архітектурі. Згорткові фільтри

виявляють внутрішньо символні та міжсимвольні ознаки зв'язності, а не ознаки усього текстового рядка. Попередні підходи, такі як обмежувальні рамки, вимагають великого рецептивного поля в таких випадках.



а



б

в

Рис. 11. Результати детектування штрих коду на основі математичної морфології та згорткової нейронної мережі: а – завантажене зображення; б – регіони, які виділила математична морфологія (без пост обробки); в – регіони, які виділила нейронна мережа

Також модель CRAFT не потребує подальшої пост обробки результату, такої як non-maximum suppression, на відміну від моделей, які відносяться до two-stage методів. Вона одразу видає на вихідному зображенні, яке має два канали: перший – вірогідність центру літери, закодовану Гаусіаном; другий – вірогідність міжсимвольної ознаки, також закодовану Гаусіаном. На рис. 12 наведені приклади детектування тексту. За допомогою CRAFT на ціннику знаходяться регіони з текстом, які далі аналізуються за зв'язністю, розміром, координатами відносно зображення. Вилучені символи об'єднуються у полігони, нормалізуються та передаються на розпізнавання.



Рис. 12. Приклад знаходження тексту моделлю CRAFT

Далі необхідно провести дослідження з метою зменшення параметрів нейронної мережі, що дало би змогу запускати її на малопотужних системах, у тому

числі і на мобільних. Для підвищення швидкої під час проектування мобільного застосунку був розглянутий варіант заміни нейронної мережі на математичну морфологію.

Далі необхідно провести дослідження з метою зменшення параметрів нейронної мережі, що дало би змогу запускати її на малопотужних системах, у томі числі і на мобільних. Для підвищення швидкої під час проектування мобільного застосунку був розглянутий варіант заміни нейронної мережі на математичну морфологію.

Розпізнавання текстової інформації та штрих коду. Розпізнавання штрих коду відбувалось відкритою бібліотекою ZXing [11]. ZXing має вбудовані функції покращення зображення, які допомагають декодувати штрих код навіть на поганих фото, аналізує зображення попіксельно і видає тип коду, та його числове значення. Але має суттєвий недолік – шукає на зображенні тільки один штрих код, тому й була потреба розробляти додатково метод для пошуку кодів, а саму бібліотеку використовувати тільки для розпізнавання вже виділених об'єктів.

Для розпізнавання текстової інформації запропонована відкрита бібліотека Tesseract, яка використовує рекурентну нейронну мережу Long short-term memory (LSTM) [12, 13] (див. рис. 13). Рекурентна нейронна мережа (Recurrent neural network, RNN) – це мережа, яка має внутрішню пам'ять. Вихід для поточного значення залежить від попередніх. Кожен елемент послідовності для обробки по черзі передається одним і тим же нейронам, вихід яких повертається і разом з наступним елементом послідовності знов передається на вхід, до тих пір поки послідовність не закінчиться. Цей клас нейронних мереж добре підходить для обробки часових послідовностей або розпізнавання тексту. У цих задачах важливий контекст – тобто знання про те, що було раніше. LSTM – це модифікація звичних рекурентних мереж з метою покращення механізму запам'ятовування за забування інформації з попередніх даних та вирішення проблеми зникаючого градієнту, для цього були додані три елементи (вентилі): forget gate, input gate, output gate [13]. На рис. 13 продемонстровані результати розпізнавання текстової інформації з цінників бібліотекою Tesseract.

Алгоритм розпізнавання цінників. Таким чином, загальний алгоритм складається з трьох головних кроків:

Крок 1. Знаходження та розпізнавання опорного елемента (наприклад, штрих коду):



Рис. 13. Результат розпізнавання текстової інформації за допомогою бібліотеки Tesseract

1. Обрати спосіб знаходження опорного елемента:

Варіант 1. Математична морфологія (алгоритм наведено вище)

Варіант 2. Згорткова нейронна мережа (перевести зображення у градацію сірого; нормувати значення яскравості; обробити зображення нейронною мережею; отримати контури; провести фільтрацію; апроксимувати прямокутники та нормалізувати).

2. Віддати для розпізнавання бібліотеці ZXing. Крок 2. Розмітка цінника за опорним елементом.

1. Розмітити по шаблону елементи цінника відповідно опорного елемента;

2. Провести нормалізацію розмічених елементів. Крок 3. Детектування та розпізнавання інших елементів (ціни та назви)

1. У розмічених регіонах знайти інформацію для розпізнавання за допомогою згорткової мережі або математичної морфології;

2. Нормалізувати знайдені елементи;

3. Віддати елементи на розпізнавання Tesseract.

Використанні технології та розробка застосунку. Основна бібліотека написана мовою C++. Це дає змогу перенести вже існуючий код на мобільні платформи, такі як Android та iOS, без переписування усього функціоналу. Під час розробки бібліотеки для розпізнавання цінників використовувались відкриті бібліотеки OpenCV, Tesseract, ZXing, Libtorch.

Для перенесення розробленого функціоналу на систему Android використана технологія Java Native Interface (JNI). Вона зв'яже байт код Java Virtual Machine (JVM) та нативний, що був розроблений на C++. За допомогою JNI бібліотека для розпізнавання цінників була імпортована під Android та розроблений мобільний застосунок. Робота з застосунком починається з обрання шаблону цінника, що потрібно розпізнати. Шаблон цінника залежить від торгівельної мережі. Далі стає доступною камера і потрібно зробити фото, після чого запускається процес розпізнавання, результат якого виводиться на наступному екрані, разом із зробленим фото (див. рис. 14). Так як телефон, це малопотужний пристрій, то нейронна мережа для знаходження тексту була замінена на методи математичної морфології, які шукають текст у заданій шаблонній області.

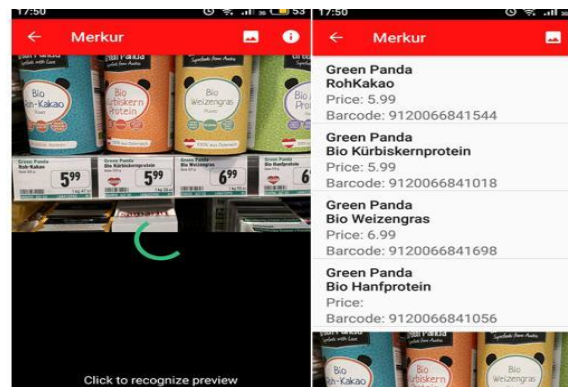


Рис. 14. Приклад роботи мобільного застосунку

Розроблений застосунок може використовуватися для автоматизації деяких операцій обліку товарів та маркетингу в торговельних секторі, наприклад, створення єдиної бази товарів та їх цін, або швидкого порівняння товарів у різних торговельних мережах.

Висновки. В роботі досліджено питання розпізнавання зображень, що містять інформацію різного типу, шляхом спільного використання класичних підходів, таких, як математична морфологія, бінаризація, фільтрація, нормалізація геометричних перетворень, та нейронних мереж для аналізу зображень та тексту. У рамках досліджень було розглянуто такі питання: розробка підходу для розпізнавання цінників на зображенні; детектування штрих коду (математична морфологія з пост обробкою, нейронна мережа); підготовка даних для навчання мережі для детектування штрих кодів (створення датасету); навчання згорткової нейронної мережі для детектування штрих кодів; порівняння результатів пошуку штрих кодів методами математичної морфології та нейронною мережею; розпізнавання штрих коду (бібліотека ZXing); детектування текстової інформації на ціннику (модель CRAFT); розпізнавання текстової інформації (бібліотека Tesseract); проектування та реалізація застосунку (мова C++, бібліотеки OpenCV, Libtorch, технологія JNI).

Порівняння методів знаходження штрих коду за допомогою математичної морфології та згорткової нейронної мережі показало, що морфологія дає значно нижчу якість детектування ніж нейронна мережа, але у декілька разів перевищує її у швидкодії. Оскільки швидкодія має велике значення для реалізації алгоритмів на мобільних пристроях, до морфології була додана обробка додатковими фільтрами та нормалізація геометричних спотворень, що суттєво поліпшило точність детектування та подальшого розпізнавання.

Також практичні дослідження щодо знаходження тексту дозволяють зробити висновок, що використана згорткова нейронна мережа CRAFT показує високу точність, але на малопотужних системах вона стає слабким місцем усього алгоритму. На малопотужних системах, як мобільні пристрої, поступаючись точністю детектування, ця мережа замінена на методи математичної морфології з фільтрацією. Тому у якості подальшого напрямку досліджень доцільно змінити частину структури моделі CRAFT, яка вилучає характерні ознаки, на більш легку, наприклад EfficientNet [14].

Також у подальшому необхідно провести тестування точності та швидкодії роботи застосунку на основі запропонованого підходу до розпізнавання цінників у цілому, поки проведено такий аналіз для окремих етапів. Розгляд даного питання дав би змогу зробити висновки щодо результатів розпізнавання більш обґрунтовано.

У цілому практичні дослідження підтвердили доцільність запропонованого підходу до розпізнавання цінників, а розроблений застосунок може бути застосований у торговельних мережах.

Список літератури

1. Serra J. *Image Analysis, Mathematical Morphology*. Academic Press, 1982. 621 p.
2. Пуятин Е. П., Яковлева Е. В., Любченко В. А. Разложение матрицы центроаффинного преобразования для нормализации изображений. *Радиоэлектроника и информатика*. 1999. № 4 (05). С. 91–94.
3. *Artificial Intelligence Development Services*. URL: <https://www.sytoss.com/data-science-and-neural-network> (дата звернення: 10.06.2020).
4. *Image Polygonal Annotation with Python (polygon, rectangle, circle, line, point and image-level flag annotation)*. URL: <https://github.com/wkentaro/labelme> (дата звернення: 15.04.2020).
5. *Common Objects in Context*. URL: <http://cocodataset.org/#home> (дата звернення: 01.04.2020).
6. Zharkov A., Zagaynov I. *Universal Barcode Detector via Semantic Segmentation*. URL: <https://arxiv.org/abs/1906.06281> (дата звернення: 10.02.2020).
7. Ioffe S., Szegedy C. *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. URL: <https://arxiv.org/abs/1502.03167> (дата звернення: 01.03.2020).
8. Zhou X. et al. EAST: an efficient and accurate scene text detector. *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2017. P. 5551–5560.
9. Simonyan K., Zisserman A. *Very deep convolutional networks for large-scale image recognition*. URL: <https://arxiv.org/abs/1409.1556> (дата звернення: 15.05.2020).
10. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham. 2015. P. 234–241.
11. ZXing (“Zebra Crossing”) barcode scanning library for Java, Android. URL: <https://github.com/zxing/zxing> (дата звернення: 10.02.2020).
12. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural computation*. 1997. Т. 9, № 8. P. 1735–1780.
13. *Understanding LSTM Networks*. URL: <https://colah.github.io/posts/2015-08-Understanding-LSTMs> (дата звернення: 10.03.2020).
14. Tan M., Le Q. V. *Efficientnet: Rethinking model scaling for convolutional neural networks*. URL: <https://arxiv.org/abs/1905.11946> (дата звернення: 20.04.2020).

References (transliterated)

1. Serra J. *Image Analysis, Mathematical Morphology*. Academic Press, 1982. 621 p.
2. Putjatin E. P., Jakovleva E. V., Ljubchenko V. A. Razlozhenie matricy centroaffinnogo preobrazovanija dlja normalizacii izobrazhenij [Centroaffine transformation matrix decomposition for image normalization]. *Radioelektronika i informatika* [Radioelectronics and informatics]. 1996, no. 4 (05), pp. 91–94.
3. *Artificial Intelligence Development Services*. Available at: <https://www.sytoss.com/data-science-and-neural-network> (accessed 10.06.2020).
4. *Image Polygonal Annotation with Python (polygon, rectangle, circle, line, point and image-level flag annotation)*. Available at: <https://github.com/wkentaro/labelme> (accessed 15.04.2020).
5. *Common Objects in Context*. Available at: <http://cocodataset.org/#home> (accessed 01.04.2020).
6. Zharkov A., Zagaynov I. *Universal Barcode Detector via Semantic Segmentation*. Available at: <https://arxiv.org/abs/1906.06281> (accessed 10.02.2020).
7. Ioffe S., Szegedy C. *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. Available at: <https://arxiv.org/abs/1502.03167> (accessed 01.03.2020).
8. Zhou X. et al. EAST: an efficient and accurate scene text detector. *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2017. pp. 5551–5560.
9. Simonyan K., Zisserman A. *Very deep convolutional networks for large-scale image recognition*. Available at: <https://arxiv.org/abs/1409.1556> (accessed 15.05.2020).
10. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham. 2015. pp. 234–241.

11. ZXing ("Zebra Crossing") barcode scanning library for Java, Android. Available at: <https://github.com/zxing/zxing> (accessed 10.02.2020).
12. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural computation*. 1997, vol. 9, no. 8, pp. 1735–1780.
13. *Understanding LSTM Networks*. Available at: <https://colah.github.io/posts/2015-08-Understanding-LSTMs> (accessed 10.03.2020).
14. Tan M., Le Q. V. *Efficientnet: Rethinking model scaling for convolutional neural networks*. Available at: <https://arxiv.org/abs/1905.11946> (accessed 20.04.2020).

Надійшла (received) 11.05.2020

Відомості про авторів / Сведения об авторах / About the Authors

Ковтуненко Андрій Романович – Харківський національний університет радіоелектроніки, бакалавр, SYTOSS R&D Engineer; м. Харків, Україна; e-mail: andrii.kovtunenکو@nure.ua

Яковлева Олена Володимирівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інформатики, м. Харків, Україна; ORCID: <http://orcid.org/0000-0002-6129-6146>; e-mail: olena.yakovleva@nure.ua

Любченко Валентин Анатолійович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інформатики, SYTOSS R&D Senior Developer, м. Харків, Україна; ORCID: <http://orcid.org/0000-0002-9966-0249>; e-mail: valentyn.liubchenko@nure.ua

Янголенко Ольга Василівна – кандидат технічних наук, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри програмної інженерії та інформаційних технологій управління, м. Харків, Україна; ORCID: <http://orcid.org/0000-0001-7755-1255>; e-mail: Olha.Yanholenko@khpі.edu.ua

Ковтуненко Андрей Романович – Харьковский национальный университет радиоэлектроники, бакалавр, SYTOSS R&D Engineer; г. Харьков, Украина; e-mail: andrii.kovtunenکو@nure.ua

Яковлева Елена Владимировна – кандидат технических наук, доцент, Харьковский национальный университет радиоэлектроники, доцент кафедры информатики, г. Харьков, Украина; ORCID: <http://orcid.org/0000-0002-6129-6146>; e-mail: olena.yakovleva@nure.ua

Любченко Валентин Анатольевич – кандидат технических наук, доцент, Харьковский национальный университет радиоэлектроники, доцент кафедры информатики, SYTOSS R&D Senior Developer, г. Харьков, Украина; ORCID: <http://orcid.org/0000-0002-9966-0249>; e-mail: valentyn.liubchenko@nure.ua

Янголенко Ольга Васильевна – кандидат технических наук, Национальный технический университет «Харьковский политехнический институт», доцент кафедры программной инженерии и информационных технологий управления, г. Харьков, Украина; ORCID: <http://orcid.org/0000-0001-7755-1255>; e-mail: Olha.Yanholenko@khpі.edu.ua

Kovtunenکو Andrii – Kharkiv National University of Radio Electronics, bachelor, SYTOSS R&D Engineer; Kharkiv, Ukraine; e-mail: andrii.kovtunenکو@nure.ua

Yakovleva Olena – Candidate of Technical Sciences, Docent, Kharkiv National University of Radio Electronics, Associate Professor of Informatics Department, Kharkiv, Ukraine; ORCID: <http://orcid.org/0000-0002-6129-6146>; e-mail: olena.yakovleva@nure.ua

Liubchenko Valentyn – Candidate of Technical Sciences, Docent, Kharkiv National University of Radio Electronics, Associate Professor of Informatics Department, SYTOSS R&D Senior Developer, Kharkiv, Ukraine; ORCID: <http://orcid.org/0000-0002-9966-0249>; e-mail: valentyn.liubchenko@nure.ua

Yanholenko Olha – Candidate of Technical Sciences, National Technical University «Kharkiv Polytechnic Institute», Associate Professor of Software Engineering and Management Information Technology Department, Kharkiv, Ukraine; ORCID: <http://orcid.org/0000-0001-7755-1255>; e-mail: Olha.Yanholenko@khpі.edu.ua

УДК 311.3:778.14:004

DOI: 10.20998/2079-0023.2020.01.06

Н. Л. СИТНИК, О. Є. ВІНОГРАДОВА, Т. В. ТЯГУН, А. Б. МАЗНИЧКО

КЛАСИФІКАЦІЯ ДОКУМЕНТІВ СТРАХОВОГО ФОНДУ ДОКУМЕНТАЦІЇ УКРАЇНИ

Відповідно до Закону України «Про страховий фонд документації України» та Положення про Державну архівну службу України одним із основних завдань Укрдержархіву є координація і контроль за формуванням страхового фонду документації України, його веденням, утриманням і наданням користувачам копій документів страхового фонду документації, необхідних для поставлення на виробництво, експлуатацію та ремонт продукції оборонного, мобілізаційного і господарського призначення, проведення будівельних (відбудовчих), аварійно-рятувальних та аварійно-відновлювальних робіт під час ліквідації надзвичайних ситуацій та в особливий період, а також у сфері

© Н. Л. Ситник, О. Є. Виноградова, Т. В. Тягун, А. Б. Мазничко, 2020