

УДК 004.932.2

канд. техн. наук, доцент Грінков В. О. ORCID: 0000-0002-9574-3792 (ВІТІ ім. Героїв Крут)
Грінкова Г. В. ORCID: 0009-0003-4896-364X (НДІ ВР)
Грінков С. В. ORCID: 0009-0009-7350-1360 (ДУ «Відкриті публічні фінанси»)

МЕТОДИКА ОЦІНКИ ТОЧНОСТІ РОЗПІЗНАВАННЯ СИМВОЛІВ І ТЕКСТУ З ЗОБРАЖЕННЯ ДЛЯ АНАЛІЗУ ЯКОСТІ СУЧАСНИХ ІНСТРУМЕНТІВ OCR

Оптичне розпізнавання символів – це потужна технологія, яка перетворює зображення з текстом у редагований і пошуковий формат. Це забезпечує ефективність у роботі з документами, підвищує доступність інформації і сприяє автоматизації багатьох процесів.

Вперше цю технологію почали використовувати на початку 90-х, при оцифруванні історичних газет для створення електронного архіву. За останні роки систему оптичного розпізнавання символів вдалося доопрацювати до «ідеалу»: нинішні системи оптичного розпізнавання символів показують майже ідеальну точність розпізнавання тексту. Але для цього потрібно виконання наступних вимог:

*рівність і контрастність символів;
однотипність фону тексту;
контраст між фоном та літерами.*

Набагато складніше процес розпізнавання символів і тексту з зображень, коли не виконуються вищеперелічені вимоги, а саме вирішення таких задач є вимогами сьогодення для вирішення практичних завдань у військовій сфері. Розпізнавання тексту з зображень має багато важливих застосувань, що робить його актуальним і необхідним.

В статті, з використанням конкретних типів зображень, проаналізовано декілька найбільш відомих та популярних моделей штучного інтелекту для розпізнавання тексту з зображення, такі як Tesseract OCR, PyTorch, EasyOCR, Keras OCR5, OpenCV, отримані результати розпізнавання тексту і символів з зображень різної складності.

Для оцінки результатів точності розпізнавання символів і тексту, розроблена методика оцінки точності розпізнавання на базі спеціальних метрик оцінки, яка основана на порівнянні розпізнаного тексту із еталонним (правильним) текстом. Найбільш поширені метрики включають точність розпізнавання символів (Character Accuracy Rate, CAR) та точність розпізнавання слів (Word Accuracy Rate, WAR).

За допомогою розробленої методики оцінки точності розпізнавання проведено аналіз точності розпізнавання найбільш популярних інструментів технології оптичного розпізнавання тексту і символів з зображень різної складності. Проведений аналіз показав, що найбільшу ефективність і точність розпізнавання демонструє модель EasyOcr, яка навіть в умовах сильної «зашумленості» і неякісної контрастності зображення демонструвала стабільний результат і при умовах подальшого налаштування для потреб користувача, може бути застосована для рішення конкретного завдання.

Ключові слова: методика оцінки, метрики оцінки, розпізнавання тексту, технології OCR, моделі OCR, точність розпізнавання, аналіз зображень, машинне навчання.

V. Hrinkov, G. Hrinkova, S. Hrinkov. Analysis of modern optical character recognition tools for character recognition and text from the image

Optical character recognition is a powerful technology that converts images with text into an editable and searchable format. This ensures efficiency in working with documents, increases the availability of information and contributes to the automation of many processes.

This technology was first used in the early 1990s, when digitizing historical newspapers to create an electronic archive. In recent years, the optical character recognition system has been refined to the "ideal": current optical character recognition systems show almost perfect text recognition accuracy. But for this, the following requirements must be met:

*equality and contrast of symbols;
homogeneity of the background of the text;
contrast between background and letters.*

The process of recognizing symbols and text from images is much more complicated when the above requirements are not met, and the solution of such problems is the current requirement for solving practical tasks in the military sphere. Text recognition from images has many important applications, which makes it relevant and necessary.

The article, using specific types of images, analyzed several of the most famous and popular models of artificial intelligence for recognizing text from an image, such as Tesseract OCR, PyTorch, EasyOCR, Keras OCR5, OpenCV, and obtained the results of recognizing text and symbols from images of various complexity.

To evaluate the results of character and text recognition accuracy, a recognition accuracy evaluation method has been developed based on special evaluation metrics, which is based on comparing the recognized text with the reference (correct) text. The most common metrics include Character Accuracy Rate (CAR) and Word Accuracy Rate (WAR).

With the help of the developed method of recognition accuracy assessment, an analysis of the recognition accuracy of the most popular tools of the technology of optical recognition of text and symbols from images of different complexity was carried out. The conducted analysis showed that the EasyOcr model demonstrates the highest efficiency and accuracy of recognition, which even in conditions of strong "noise" and low-quality contrast of the image showed a stable result, and under the conditions of further adjustment for the needs of the user, it can be applied to solve a specific task.

Keywords: evaluation method, evaluation metrics, text recognition, OCR technologies, OCR models, recognition accuracy, image analysis, machine learning.

Актуальність та постановка завдання в загальному вигляді. Зі збільшенням обсягу цифрових та друкованих даних все більш актуальним стає використання технологій автоматичного розпізнавання тексту. Інструменти *OCR (Optical Character Recognition)* дозволяють перетворювати відскановані документи, фотографії та інші зображення, що містять текст, у редагованому цифровому форматі. Це сприяє підвищенню ефективності обробки інформації в різних сферах. У військовій сфері інформація є ключовим ресурсом для забезпечення оперативного управління та ухвалення тактичних і стратегічних рішень. В умовах бойових дій та військових операцій необхідно швидко обробляти великі обсяги даних із різних джерел. Технології *OCR*, які дозволяють автоматично розпізнавати текст із зображення, залишаються надзвичайно актуальними і необхідними в різних напрямках діяльності Збройних сил України:

1. Розвідка та аналіз даних. *OCR*-технології можуть бути застосовані для автоматичного розпізнавання та аналізу текстової інформації з розвідувальних зображень, таких як фотографії з дронів, супутників та інших розвідувальних платформ. Це допомагає витягувати корисні дані із захоплених у противника документів, написів на транспортних засобах та об'єктах.

2. Підтримка в бойових операціях. Розпізнавання тексту із зображень може використовуватися для оновлення карт і навігаційних систем у реальному часі, що є критично важливим для планування операцій і переміщення військ, виявлення та аналіз написів на зображеннях може допомогти в ідентифікації прихованих загроз, таких як міни або вибухові пристрої, позначені попереджувальними знаками.

3. Автоматизація документації та логістики. *OCR*-технології можуть застосовуватись для автоматизації обробки різних документів та логістичних даних у військовій сфері. Це включає розпізнавання та обробку рахунків, накладних, контрактів та іншої документації, що прискорює процеси та знижує ризик помилок.

4. Кібербезпека. *OCR*-технології можуть застосовуватись для аналізу текстів та кодів у рамках забезпечення кібербезпеки. Це включає розпізнавання шкідливих текстів, сканування документів на наявність прихованих загроз та аналіз логів систем.

5. Забезпечення безпеки та контролю доступу. *OCR*-технології можуть використовуватися для автоматичного розпізнавання текстів на перепустках, посвідченнях особи та інших документах, що сприяє посиленню контролю доступу та забезпечення безпеки на військових об'єктах.

6. Технічна документація та навчання. *OCR*-технології можуть використовуватися для оцифрування та обробки технічної документації, інструкцій та навчальних матеріалів, що сприяє більш ефективному навчанню та підвищенню кваліфікації військовослужбовців.

7. Розпізнавання текстів на транспортних засобах та устаткуванні. *OCR*-технології можуть застосовуватись для розпізнавання текстів, серійних номерів та міток на військовій техніці, обладнанні та транспортних засобах, що допомагає в ідентифікації та відстеженні активів.

8. Переклад та локалізація. *OCR*-технології у поєднанні із системами машинного перекладу може допомогти у перекладі текстів різними мовами, що важливо для взаємодії з міжнародними партнерами та коаліційними силами.

9. Зв'язок та координація. Автоматичне вилучення інформації з польових звітів дозволить командирам отримувати актуальну інформацію та приймати рішення в режимі реального часу, розпізнавання тексту може бути інтегроване з іншими системами управління і контролю, що покращує загальну координацію дій.

Таким чином, розпізнавання тексту із зображень є важливим інструментом для підвищення ефективності, точності та швидкості виконання військових завдань. Технологічні досягнення в цій області можуть суттєво підвищити бойову здатність та інформаційну перевагу Збройних сил України.

Зважаючи на вищезазначене, методика оцінки точності розпізнавання, на базі якої здійснюється пошук найбільш ефективного інструменту розпізнавання тексту і символів із зображень є актуальним завданням.

Аналіз попередніх досліджень. Науковим дослідженням щодо використання інструментів *OCR* для розпізнавання символів і тексту присвячено значну кількість робіт [1–12], аналіз деяких з них викладено нижче.

Дослідження [1] присвячено аналізу та дослідженню сучасних підходів до обробки зображення та розпізнавання текстової інформації у технології *OCR*. У ході дослідження виявлено та проаналізовано найбільш популярні методи класифікації (шаблонний, структурний, ознаковий, статистичний, із застосуванням штучних нейронних мереж) та зроблено висновки стосовно ефективних напрямів їх застосування.

В роботі [2] описуються способи використання інструментів *OCR* для перетворення будь-яких надрукованих або відсканованих текстових зображень в цифровий формат.

В роботі [3] наводиться історичний екскурс розвитку технологій *OCR*, поточний стан використання технологій *OCR* для розпізнавання символів і перспективи його розвитку і застосування.

Робота [4] розглядає сучасні підходи до *OCR*, включаючи методи глибинного навчання. Вона також містить огляд різних інструментів і програм *OCR*, ілюстрованих прикладами для кращого розуміння технології.

Робота [5] досліджує новітні досягнення у розпізнаванні тексту на основі глибокого навчання. Автори описують моделі для розпізнавання та розпізнавання тексту в складних сценах, таких як вуличні фотографії та відео.

Робота [6] досліджує використання згорткових нейронних мереж (CNN) для розпізнавання тексту в зображеннях із природних сцен (вивіски, вуличні таблички тощо). В роботі представляється наскрізна система для виявлення тексту, локалізації та розпізнавання тексту в природному вигляді і пошук текстових зображень. Система побудована на визнанні і ранжуванні пропозиції, тренуванні великої згортки нейронної мережі для розпізнавання слів у цілому фрагменті зображення одночасно, завдяки систем, заснованих на класифікаторі символів минулого.

З розвитком нейронних мереж і машинного навчання сучасні *OCR*-системи досягли значних успіхів у точності розпізнавання тексту, однак проблеми зчитування тексту на низькоякісних зображеннях, різноманітність мов, шрифтів та складні макети документів все ще залишаються актуальними. Аналіз наявних інструментів *OCR* дозволяє розуміти їхні можливості, обмеження та перспективи подальшого вдосконалення.

Метою статті є розробка методики оцінки точності розпізнавання символів і тексту моделями та проведення аналізу і порівняння сучасних інструментів OCR розпізнавання символів і тексту із зображенням різної складності.

Для досягнення цієї мети необхідно провести:

1. Огляд існуючих алгоритмів та технологій, що використовуються в сучасних OCR-системах.
2. Вибір найпопулярніших моделей OCR для аналізу.
3. Вибір опцій зображень.
4. Розробка методики оцінки точності розпізнавання символів та тексту.
5. Оцінка точності інструментів при розпізнаванні текстів на зображеннях різної якості.
6. Виявлення переваг та обмеження OCR-системи для роботи з різними типами зображень.
7. Розробка рекомендацій щодо вибору та використання OCR-інструментів залежно від вирішення конкретної задачі.

Результати аналізу користувальницьких програм застосовуються з вибором найбільш відповідного інструменту OCR для конкретних потреб та завдань, а також сприяють розвитку нових підходів до вдосконалення системного розпізнавання.

Виклад основного матеріалу. Існує багато моделей і бібліотек, котрі можна використовувати для розпізнавання тексту і символів з зображення. У цій статті буде використовуватись мова програмування Python, як більш популярна мова для штучного інтелекту та машинного навчання та більш лаконічна. Приведемо декілька найбільш відомих і популярних моделей.

1. *TesseractOCR1*. Інструмент оптичного розпізнавання символів, розроблений Google, який підтримує більше 100 мов. Він може працювати з різноманітними форматами зображень та обробляти складні тексти. Для використання *Tesseract OCR* на *Python*, необхідно встановити бібліотеку *pytesseract2*, котра надає простий інтерфейс для роботи з *Tesseract OCR*. Офіційний ресурс моделі *Tesseract OCR*: [GitHub - tesseract-ocr/tesseract: Tesseract Open Source OCR Engine \(main repository\)](#).

2. *EasyOCR3*. Бібліотека, котра використовує нейронні мережі для розпізнавання тексту з зображення. Вона підтримує більше 80 мов, включаючи кирилицю, може розпізнавати текст у різних орієнтаціях та умовах освітлення. *EasyOCR* також дозволяє налаштовувати параметри розпізнавання, такі як розмір тексту, колір фону та контраст. Офіційний ресурс моделі *EasyOCR*: [\[GitHub - JaidedAI/EasyOCR: Ready-to-use OCR with 80+ supported languages and all popular writing scripts including Latin, Chinese, Arabic, Devanagari, Cyrillic and etc\]\[7\]](#).

3. *PyTorch OCR4*. Бібліотека, яка реалізує кілька сучасних моделей для розпізнавання тексту із зображення, таких як *CRNN*, *STAR-Net*, *RobustScanner* і *NRTR*. Вона заснована на фреймворці *PyTorch* та підтримує *GPU* – прискорення. *PyTorch OCR* також надає вже обучені моделі для різних мов та сценаріїв, таких як документи, таблиці та рукописний текст. Офіційний ресурс моделі *PyTorch OCR*: [\[GitHub - DYJNG/PyTorchOCR: OCR Toolkits based on PyTorch\] \[8\]](#).

4. *Keras OCR5*. Бібліотека, яка реалізує двоетапний підхід для розпізнавання тексту із зображення. Спочатку вона використовує модель *CRAFT* для виявлення областей тексту на зображенні, а потім використовує модель *CRNN* для розпізнавання символів у кожній області. Вона заснована на фреймворці *Keras* та підтримує *GPU*-прискорення. *Keras OCR* також надає вже навчені моделі для англійської та французької мов. Офіційний ресурс моделі *Keras OCR*: [\[GitHub - faustomorales/keras-ocr: A packaged and flexible version of the CRAFT text detector and Keras CRNN recognition model\] \[9\]](#).

5. *OpenCV* – це бібліотека, яка надає велику кількість функцій для роботи із зображеннями та відео, включаючи розпізнавання тексту. Вона підтримує різні алгоритми для

виявлення та вилучення тексту із зображення, такі як MSER, SWT та EAST. Вона також дозволяє використовувати різні двигуни для розпізнавання символів, такі як *Tesseract OCR*, *LSTM* та *CNN*. Офіційний ресурс моделі *OpenCV*: [*OpenCV - Open Computer Vision Library*] [10].

EasyOCR. Встановлюємо модель через *pip install pip install easyocr*. Модель дуже зручна, оскільки має підтримку багатьох мов та автоматично завантажує їх при додаванні аббревіатури цих мов, також є можливість одночасно використовувати кілька мов для розпізнавання. Модель постійно оновлюється та покращується.

Tesseract OCR. Встановлюємо її через *pip install pip install pytesseract*. Для цієї моделі в стартовому наборі існує лише модель для розпізнавання англійської мови, якщо використовується інша мова, необхідно завантажити мовні дані додатково з репозиторію, що є доволі незручним.

PyTorch. Для розпізнавання тексту на зображенні з використанням *PyTorch* використовується *Tesseract* (для розпізнавання тексту) та *PyTorch* для розпізнавання областей зображення з текстом. Встановлюємо ці бібліотеки за допомогою *pip install: pip install pytesseract Pillow torchvision*.

Keras OCR5. Встановлюємо цю бібліотеку за допомогою *pip install: pip install keras-ocr*
OpenCV. Встановлюємо цю бібліотеку за допомогою *pip install: pip install pytesseract pip install opencv-python*

Протестуємо вказані моделі на трьох зображеннях різного рівня складності. Для першого зображення з розміром 73 кБ (рис. 1), результати розпізнавання тексту (символів) приведені в таблиці 1.



Рис. 1. Зображення 73 кБ

Таблиця 1

Результати розпізнавання тексту і символів із зображення на рис. 1

<i>EasyOCR</i>	<i>Tesseract OCR</i>	<i>PyTorch</i>	<i>Keras OCR5</i>	<i>OpenCV</i>
[ENLI', [Ч@ПИ', [EISTRГ', 'Б1ош', '89', 'Шрота', 'Допота', 'Скавость', 'Цепь Нет даннымо GPS', 'NSmsLId', 'БПА', '[Нет даШНЬМЕТ GP S', '178', 'Крс']	ВОЛр ор Розраои е ВАг Бтом] @.° ЦИ Ц рота олгота @ Скорость Ё. МА еп Негданных т оРУ он _В щ аЁЁЗ,ЁЁЁ.ЁЁН"ЁБЦ оате оАИЫ Болефк сооаещесщес оО в Рар о СОя оь о соой *- ощ сиее аНЕО кО,, оо овосорееоноНонН Н УН —° СОИ ТР о сснынеьннь чesyЕ Н @т ОБЕ о КОЧЕ стр ЯОМ ИО р палек оаЫ ооа нааВЬссесв > —чача оо Б ЬАМ о, с ПИЗВаЕ На ы й ' : , Ь:, _ , \ \ < ^ ста РОИ оо аЕН, МЕОар оР Воаео ол ОАар НЕЙ ЛВ ВЕИ ОАА ; * @ ЙОЕ МОНЕа КОЙ ЗЕВ ОН ов с ГА^ оааа н г""г,»"щ Ь""?@Ё"" В РОИ #Гi.ь;?... \; *_ ,; ,т VM ,ч#, >ё... ж ОкБ` ААНЫ " ___ - !.: ' _ ,; @, : ; ; ; ; э ' Пр асаа —	Для першого зображення модель не розпізнала нічого	['alt', 'es', 'baadi', 'ins', 'start', 'stons', 'for', 'bt', '9', 'gind', '89', 'donteta', 'Impeta', 'ckopocт', 'керс', 'het', 'grs', 'nsnotreh', 'lenb', 'sahtae', 'et', 'si', 'het', 'gps', 'ot', 'aahnae', 't8']	Для першого зображення модель не розпізнала нічого

Для другого зображення з розміром 217 кБ (рис. 2) моделі видали наступні результати (таблиця 2):



Рис. 2. Зображення 217 кБ

Таблиця 2

Результати розпізнавання тексту (символів) з зображення рис. 2

EasyOCR	Tesseract OCR	PyTorch	Keras OCR5	OpenCV
['ДВТошКола #ро9', 'ДВИЖИМОНИ', 'ХАТЯЖНЬД', 'Дан" ТУРАГЕНТСТВО', ''', '8 (915) 871-21- 53', '[потолки', 'ШШШШ [ВЕРП', 'Альфадепт 722- 227', 'Стоашатообог', 'добрая', '~ЛБНЕяиг', 'ТУИГЕНТЕГНО', '1', '8-915-871-21-53', 'Луа Уашон', '~агашокш', '[']	ста + ' во, „Заснашаторасоы в о К СО '@'Єi@ Й —° И е шо \ В, - Pi,,: ""'м;—\ — Е == Ы ВВ, % ЕА ВТОШКОЛА «Город» - % -5ваа 2° ва р 60 вяниности_ , й Г каа е оое 8 Ы НЕ М_i ' роой) ПБ ТО АН т2227 стообла _a_ 'r' Н ЕЕЕ Ё i с ЗртНе в) — аМСа О == — уа Уапатом манатоулы	ime ТУРАТЕАТСТВО (1h 8 aie nee 21-53 =3 уОЕ SX KIS XS ages_tlya Varlamov varlamov.ru-	['c', 'abtolwkola', 'ropog', 'i', ", 'iyer', 'typarehtctbo', 'hatakhble', 'abiammoth', '8', '9151', '8712153', 'notojika', 'okha', 'abeph', 'goopaa', 'bawa', 'ajibpaicht', '722227', 'ctomatoiorh', 'tooupat', 'typarehtctbo', 'p', 'e', 'a', '89158112153', 'lya', 'varlamoy', 'varamoviu']	'wa ТУРАТЕНТСТВО . 8 (915) 8671-24-53 i UGK (IRIE 2-227 cxbsiensteen OO o —— nn + Mya Variamov variamov.ne:

Тут вже більше моделей видали прийнятний результат, хоча текст був не накладений на зображення, тому його виділення було ускладнено.

Для третього зображення з розміром 122 кБ (рис. 3), моделі видали наступні результати (таблиця 3):

	A	B	C	D	E	F
1	Name	Type 1	Type 2	HP	Attack	Defense
2	Abra	Psychic		25	20	15
3	Kadabra	Psychic		40	35	30
4	Alakazam	Psychic		55	50	45
5	Machop	Fighting		70	80	50
6	Machoke	Fighting		80	100	70
7	Machamp	Fighting		90	130	80
8	Bellsprout	Grass	Poison	50	75	35
9	Weepinbell	Grass	Poison	65	90	50
10	Victreebel	Grass	Poison	80	105	65
11	Tentacool	Water	Poison	40	40	35
12	Tentacruel	Water	Poison	80	70	65

Рис. 3. Зображення 122 кБ

Таблиця 3

Результати розпізнавання тексту (символів) з зображення рис. 3

<i>EasyOCR</i>	<i>Tesseract OCR</i>	<i>PyTorch</i>	<i>Keras OCR5</i>	<i>OpenCV</i>
[['B', 'D', 'E', '1', 'Name', 'Type 1', 'Type 20', 'HP', 'Attack', 'Defense_', '2', 'Abra', 'Psychic', '25', '20', '15', '3', 'Kadabra', 'Psychic', '40', '35', '30', '4', 'Alakazam', 'Psychic', '55', '50', '45', '5', 'Machop', 'Fighting', '70', '80', '50', '6', 'Machoke', 'Fighting', '80', '100', '70', 'Machamp', 'Fighting', '90', '130', '80', '8', 'Bellsprout', 'Grass', 'Poison', '50', '75', '35', '9', 'Weepinbell', 'Grass', 'Poison', '65', '90', '50', '10', 'Victreebel', 'Grass', 'Poison', '80', '105', '65', '11', 'Tentacool', 'Water', 'Poison', '40', '40', '35', '12', 'Tentacruel', 'Water', 'Poison', '80', '70', '65']	A B ç C D E F G imName Ed Type1Bd Type 2R7HP (Attack BJ Defense 2 Abra Psychic 25 20 15 3 Kadabra Psychic 40 35 30 4 Alakazam Psychic 55 50 45 5 Machop Fighting 70 80 50 6 Machoke _ Fighting 80 100 70 7 Machamp _ Fighting 90 130 80 8 Bellsprout Grass Poison 50 75 35 9 Weepinbell Grass Poison 65 90 50 10 Victreebel Grass Poison 80 105 65 11 Tentacool Water Poison 40 40 35 12 Tentacruel Water Poison 80 70 65	A B ç C D E F Name PJ Type1Bd Type 2B4HP [fq Attack BJ Defense Abra Psychic 25 20 15 Kadabra Psychic 40 35 30 Alakazam _ Psychic 55 50 45 Machop Fighting 70 80 50 Machoke _ Fighting 80 100 70 Machamp _ Fighting 90 130 80 Bellsprout Grass Poison 50 75 35 Weepinbell Grass Poison 65 90 50 Victreebel Grass Poison 80 105 65 Tentacool Water Poison 40 40 35 Tentacruel Water Poison 80 70 65	['c', 'a', 'e', 'f', 'g', 'b', 'd', 'defense', 'type', 'type', 'hp', 'attack', 'name', '1', '2', '1', 'psychic', 'abra', '20', '25', '15', '2', 'kadabra', 'psychic', '40', '35', '30', '3', 'alakazam', , 'psychic', '55', '50', '45', '4', 'machop', 'fighting', '70', '80', '50', '5', 'machoke', 'fighting', '80', '100', '70', '6', 'machamp', 'fighting', '90', '130', '80', '7', 'bellsprout', 'grass', 'poison', '50', '75', '35', '8', 'weepinbell', '90', 'grass', 'po', 'ison', '65', '50', '9', 'victreebel', 'grass', '80', '105', 'po', 'ison', '65', '10', 'tentacool', 'water', 'poison', '40', '40', '35', '11', 'tentacruel', 'water', 'po', '80', '70', '65', '12', 'ison']	wo wWOnN DU BWN BPP RP WNRr OO A Name B Cc a Type 18a Type 2a HP ba Attack a Defenselig Abra Kadabra Alakazam Machop Machoke Machamp Bellsprout Weepinbell Victreebel Tentacool Tentacruel Psychic Psychic Fighting Fighting Fighting Grass Grass Grass Water Water Poison Poison Poison Poison Poison 25 40 55 70 80 90 50 65 80 40 80 20 35 50 80 100 130 75 90 105 40 70 15 30 45 50 70 80 35

Визначення точності розпізнавання тексту чи символів із зображення здійснюємо за допомогою загально визначених метрик оцінки, які порівнюють розпізнаний текст із еталонним (правильним) текстом. Найбільш поширені метрики (оцінки точності) включають точність розпізнавання символів (*Character Accuracy Rate, CAR*) та точність розпізнавання слів (*Word Accuracy Rate, WAR*) і описуються формулами (1), (2) [11]:

$$CAR = (1 - \frac{S + D + I}{N}) \times 100\%, \quad (1)$$

де S – кількість заміненних символів;

D – кількість віддалених символів;

I – кількість вставлених символів;

N – загальна кількість символів у еталонному тексті.

$$WAR = (1 - \frac{S_w + D_w + I_w}{N_w}) \times 100\%, \quad (2)$$

де S_w – кількість заміненних слів;

D_w – кількість віддалених слів;

I_w – кількість вставлених слів;

N_w – загальна кількість слів у еталонному тексті.

Сутність методики оцінки точності розпізнавання символів полягає у порівнянні еталонного зображення з результатами розпізнавання і складається у виконанні наступних кроків (для пояснення будемо використовувати таблицю 4).

Таблиця 4

Пояснення алгоритму методики

Еталонний фрагмент						Результат розпізнавання									
A	B	C	D	E	F	A	B	ç	C	D	E	F			
Name	Type 1	Type 2	HP	Attack	Defense	Nime	PJ	Type	1Bd	Type	2B4HP	[fq	Attack	BJ	Defense

Крок 1. Визначаємо загальну кількість символів у еталонному фрагменті. При розрахунку змінної пропуски не враховуються, знак дорівнює 2 символам $N = 45$.

Крок 2. Визначаємо загальну кількість символів у фрагменті розпізнавання $N^* = 47$.

Крок 3. Визначаємо значення змінних I, D . Для цього знаходимо різницю $\Delta = N^* - N$. Якщо $\Delta > 0$, то $I = \Delta, D = 0$. Якщо $\Delta < 0$, $I = 0, D = |\Delta|$. Якщо $\Delta = 0$, то $I = 0, D = 0$. Для нашого прикладу $I = 2, D = 0$.

Крок 4. Методом порівняння визначаємо значення змінної S , для нашого прикладу $S = 12$.

Крок 5. Значення змінних підставляємо в формулу 1, знаходимо точність розпізнавання символів. Для нашого прикладу $CAR = 69\%$. Якщо відношення $\frac{S + D + I}{N} > 1$, то $CAR = 0$.

Сутність методики оцінки точності розпізнавання слів також полягає у порівнянні еталонного зображення з результатами розпізнавання і складається у виконанні наступних кроків (для пояснення будемо використовувати таблицю 4). Під словом розумітимемо множину символів між пропусками.

Крок 1. Визначаємо загальну кількість слів у еталонному фрагменті $N_W = 16$.

Крок 2. Визначаємо загальну кількість слів у фрагменті розпізнавання $N_W^* = 16$.

Крок 3. Визначаємо значення змінних I_W, D_W . Для цього знаходимо різницю $\Delta_W = N_W^* - N_W$. Якщо $\Delta_W > 0$, то $I_W = \Delta_W, D_W = 0$. Якщо $\Delta_W < 0$, $I_W = 0, D_W = |\Delta_W|$. Якщо $\Delta_W = 0$, то $I_W = 0, D_W = 0$. Для нашого прикладу $I_W = 0, D_W = 0$.

Крок 4. Методом порівняння між еталонним і результуючим фрагментами визначаємо значення змінної S_W , для нашого прикладу $S_W = 8$.

Крок 5. Значення змінних підставляємо в формулу 2, знаходимо точність розпізнавання слів. Для нашого прикладу $WAR = 50\%$.

Якщо відношення $\frac{S_W + D_W + I_W}{N_W} > 1$, то $WAR = 0$.

Результати розпізнавання символів і тексту, які отримані за допомогою вищенаведених методик наведено в таблиці 5.

Таблиця 5

Точність розпізнавання символів і тексту із зображення для вибраних моделей OCR в умовах розпізнавання різної складності

Зображення	EasyOCR, %		Tesseract OCR, %		PyTorch, %		Keras OCR5, %		OpenCV, %	
	CAR	WAR	CAR	WAR	CAR	WAR	CAR	WAR	CAR	WAR
Рисунок 1	54	49	0	0	0	0	22	11	0	0
Рисунок 2	78	72	46	0	18	15	68	47	25	18
Рисунок 3	98	93	98	92	91	71	96	93	85	68

Для розглянутих прикладів краще всього себе показала модель *EasyOcr*. Також вона є найзручнішою та однією з найпопулярніших і частіше оновлених. При використанні налаштувань для зображень, відповідно до конкретних потреб, вона покаже ще кращі результати. На другому місці модель *Keras OCR5*. У моделі є труднощі з встановленням. У моделі всього дві мови доля розпізнавання – англійська та французька, тож у випадку інших мов краще вибрати іншу модель.

Висновки. В результаті аналізу деяких моделей технології *OCR* для розпізнавання тексту і символів із зображень різної складності, які проведені за допомогою розроблених методик, найбільшу ефективність і точність розпізнавання продемонструвала модель *EasyOcr*, яка навіть в умовах сильної «зашумленості» і неякісної контрастності зображення (рис. 1) показала непоганий результат і при умовах подальшого налаштування для потреб користувача, може бути застосована для рішення конкретного завдання.

Майбутні напрями досліджень. Таким чином, отримані результати є підґрунтям для подальших наукових досліджень, які полягають в автоматизації процесу підрахунку даних для оцінки точності розпізнавання і продовженню аналізу якості розпізнавання інших моделей *OCR*, вдосконаленні метрики для оцінки якості розпізнавання, включення нових метрик, таких як семантична точність, що оцінює збереження змісту документа, і структурна точність для складних документів (таблиці, діаграми, формули). Це дозволить глибше аналізувати точність розпізнавання різних типів контенту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Переяславська С., Шевченко В., Смагіна О. Аналіз підходів до розпізнавання текстової інформації у технології *OCR*. Scientific Collection «InterConf»: SCIENTIFIC RESEARCH IN XXI CENTURY (March 6–8, 2021).
2. Hamad K, Kaya M. A Detailed Analysis of Optical Character Recognition Technology URL: <https://dergipark.org.tr/en/download/article-file/236939> (дата звернення: 25.04.2024).
3. Eikvil L. OCR Optical Character Recognition. URL: <https://www.nr.no/~eikvil/OCR.pdf> (дата звернення: 15.09.2024).
4. Optical Character Recognition: An Illustrated Guide to the Frontier. Georeg Nagy, Stephen V. Rise, Thomas A. Narker. Springer Science&Business Media. 2019 (дата звернення: 25.05.2024).
5. Scene Text Detection and Recognition: The Era of Deep Learning – Baoguan Shi, Xiang Bai, Kong Yao (2017). URL: <https://www.researchgate.net/publication/328899907> (дата звернення: 20.10.2024).
6. Reading Text in the Wild with Convolutional Neural Networks. Jaderberg M., Simonyan K., Vedaldi A., Zisserman A. URL: <https://www.robots.ox.ac.uk/~vgg/publications/2016/Jaderberg16/jaderberg16.Pdf> (дата звернення: 14.08.2024).
7. URL: <https://github.com/JaidedAI/EasyOCR> (дата звернення: 16.03.2024).
8. URL: <https://github.com/DYJNG/PyTorchOCR> (дата звернення: 09.04.2024).
9. URL: <https://github.com/faustomorales/keras-ocr> (дата звернення: 19.04.2024).
10. URL: <https://opencv.org/> (дата звернення: 18.04.2024).
11. ISO/IEC 19757-7:2009. "Information technology – Document Schema Definition Languages (DSDL) – Part 7: Character Repertoire Description Language (CREPDL)".
12. Jurafsky, D., & Martin, J. H. (2009). "Speech and Language Processing". Upper Saddle River, New Jersey 07458.