

РЕЦЕНЗІЯ

на дисертаційну роботу
Зарічкового Олександра Анатолійовича
на тему «Методи та програмне забезпечення розмітки відеоданих для задач
комп'ютерного зору»,
представлену на здобуття ступеня доктора філософії

в галузі знань 12 Інформаційні технології
за спеціальністю 121 Інженерія програмного забезпечення

Актуальність теми дисертації.

Актуальність дисертаційного дослідження на тему «Методи та програмне забезпечення розмітки відеоданих для задач комп'ютерного зору» визначається зростанням ролі штучного інтелекту та комп'ютерного зору у сучасному цифровому світі. Завдання комп'ютерного зору, такі як детекція об'єктів, класифікація відеоданих і розпізнавання складних сцен, значною мірою залежать від наявності великих і якісно розмічених наборів даних. Однак існуючі інструменти розмітки часто не відповідають сучасним вимогам індустрії, через що процес стає трудомістким, дорогим і не завжди точним.

У роботі пропонується вирішення цієї проблеми шляхом розробки методів та програмного забезпечення, які забезпечують автоматизацію та оптимізацію процесу розмітки відеоданих. Результати дослідження мають практичну значимість у різних сферах, включаючи автономне управління, відеоспостереження, медицину та інші галузі, де якість і швидкість обробки відеоданих є критичними. Запропоновані методи у сукупності дозволяють підвищити ефективність процесу.

Робота сприяє розвитку інструментів, які відповідають сучасним викликам, що стоять перед системами штучного інтелекту, і забезпечує нові можливості для подальшого вдосконалення методів обробки даних при вирішенні задач комп'ютерного зору.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

1. Вперше запропоновано дуальну архітектуру програмного забезпечення для автоматизованої розмітки даних, яка, за рахунок методу адаптивно-агрегованого навчання нейромережі, забезпечує пришвидшення процесу розмітки та, на відміну від існуючих аналогів, дає змогу ефективного застосування нульового та активного навчання нейромережі для розмітки

даних та більш гнучкого використання програмного забезпечення для різноманітних задач комп'ютерного зору;

2. Вперше запропоновано метод пріоритезації складних зразків для навчання нейронної мережі, який, за рахунок відбору найскладніших зразків для навчання, підвищує якість набору даних без проведення попередньої розмітки відео, внаслідок чого збільшується точність детекції об'єктів на відео, та, на відміну від існуючих підходів, базується виключно на автоматично згенерованій репрезентації даних;
3. Вперше запропоновано ітеративний метод вибору ключових кадрів на довгих відео, що дає змогу визначати ключові кадри та сегменти відео з поступовим підвищенням точності, та, на відміну від існуючих методів, враховувати динамічно зміни контенту відео для вибору ключових кадрів, підвищуючи точність сегментації та зменшуючи обсяг відеоданих для обробки;
4. Вперше запропоновано метод агрегації знань між текстовою та візуальною частинами у візуально-мовній моделі (VLM) для обробки складних мультимодальних взаємодій, що забезпечує більш високу точність розпізнавання складних сцен на відео та їх опису у порівнянні з існуючими аналогами.

Наукові дослідження були виконані здобувачем на кафедрі інформатики та програмної інженерії КПІ ім. Ігоря Сікорського в рамках НДР «Методи та технології високопродуктивних обчислень та обробки надвеликих масивів даних» (державний реєстраційний номер 0117U000924) під керівництвом проф., д.т.н., проф. Стеценко Інни Вячеславівни.

Поставлене у дисертаційній роботі наукове завдання підвищення швидкості розмітки відеоданих у контексті завдань детекції об'єктів виконано повністю, достовірність отриманих результатів підтверджується серією експериментів та обґрунтованих висновків.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.

За своїм змістом дисертаційна робота здобувача Зарічкового О. А. повністю відповідає Стандарту вищої освіти зі спеціальності 121 Інженерія програмного забезпечення та напрямкам досліджень відповідно до освітньої програми Інженерія програмного забезпечення.

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у науковий напрям інженерії програмного забезпечення.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадиння, можна зробити висновок, що дисертаційна робота Зарічкового Олександра Анатолійовича є результатом самостійних

досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Виявлені збіги пов'язані з власними науковими працями здобувача, які були раніше опубліковані для представлення основних результатів дисертації. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

Мова та стиль викладення результатів.

Дисертаційна робота написана українською мовою.

У дисертації Зарічкового О.А. «Методи та програмне забезпечення розмітки відеоданих для задач комп'ютерного зору» матеріал викладено чітко, логічно та послідовно, з використанням загально вживаної професійної термінології.

Дисертація складається зі вступу, 6 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації 261 сторінок.

У вступі розкрито актуальність теми дослідження, сформульовано мету, завдання, об'єкт і предмет дослідження, а також наведено основні положення наукової новизни та практичного значення роботи.

У першому розділі розглянуто основи детекції об'єктів і програмні засоби для розмітки відеоданих. Проведено аналіз сучасних методів розмітки, включаючи R-CNN, YOLO та інші, а також сформульовано задачу дослідження.

Другий розділ присвячено методу пріоритезації складних зразків для навчання нейромереж, що підвищує якість набору даних та точність детекції об'єктів без попередньої розмітки.

Третій розділ містить ітеративний метод вибору ключових кадрів на довгих відео, що зменшує кількість нерелевантних кадрів, оптимізуючи процес сегментації.

Четвертий розділ представляє метод агрегації знань у мультимодальних моделях (VLM), що забезпечує точніше розпізнавання складних сцен на відео.

У п'ятому розділі наведено огляд програмних засобів розмітки, описано дуальну архітектуру для автоматизації розмітки даних і розроблене програмне забезпечення.

Шостий розділ включає результати експериментального дослідження, порівняння підходів до автоматизації розмітки та оцінку ефективності розробленої архітектури.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

Оприлюднення результатів дисертаційної роботи.

Наукові результати дисертації висвітлені у 6 наукових публікаціях здобувача, серед яких: 1 стаття у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України; 4 статті у

періодичних наукових виданнях, проіндексованих у базах даних Web of Science Core Collection та/або Scopus.

Також результати дисертації були апробовані на 4 наукових фахових конференціях.

Публікації здобувача повністю відображають результати дисертаційного дослідження. Поршень принципів академічної доброчесності не виявлено. Крім наукового керівника, інших співавторів у публікаціях немає, що підтверджує беззаперечний особистий внесок здобувача.

Таким чином, наукові результати описані в дисертаційній роботі повністю висвітлені у наукових публікаціях здобувача.

Недоліки та зауваження до дисертаційної роботи

1. У тексті дисертації спостерігається надмірне використання запозичених рисунків, для яких формально існує посилання на першоджерела, але для повного їх розуміння необхідне глибоке знання англомовної термінології та англомовних скорочень.
2. Бажано було б уважніше поставитись до формування заголовка розділу 2, який не виконує функцію узагальнення змісту, оскільки співпадає з назвою підрозділу 2.3 цього ж розділу.
3. Автор дисертаційної роботи у ряді розділів використовує оглядові матеріали, які варто було б віднести до розділу 1, що покращило б її структурованість.
4. У шостому розділі відсутній детальний опис параметрів, які були використані під час експериментальних досліджень, що значно ускладнює можливість їх відтворення іншими дослідниками. Зокрема, не наведено критеріїв вибору параметрів, що використовувалися в експериментах, і не обґрунтовано їхнє значення для поставлених завдань дослідження. Також бракує пояснень, чому саме ці параметри є важливими для досягнення цілей роботи та як їх вибір вплинув на отримані результати.
5. У дисертаційному дослідженні відсутній аналіз впливу якості відео, зокрема його роздільної здатності (4K, HD, SD), на результуючу ефективність запропонованих методів. Недостатня увага до цього аспекту може обмежити розуміння впливу таких факторів на точність і швидкість розмітки відеоданих, що є важливим для практичного застосування розробок.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

Висновок про дисертаційну роботу.

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Зарічкового Олександра Анатолійовича на тему «Методи та програмне забезпечення розмітки відеоданих для задач комп'ютерного зору» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі знань 12 Інформаційні технології. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 – 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Зарічковий Олександр Анатолійович заслуговує на присудження ступеня доктора філософії в галузі 12 Інформаційні технології за спеціальністю 121 Інженерія програмного забезпечення.

Рецензент:

в.о. завідуючого кафедри обчислювальної техніки
Національного технічного університету України
«Київський політехнічний інститут імені
Ігоря Сікорського»,
д.т.н., проф.



Михайло НОВОТАРСЬКИЙ

« ____ »

