

РЕЦЕНЗІЯ

на дисертаційну роботу
Здора Костянтина Андрійовича
на тему «Моделі та програмні засоби підвищення швидкодії визначення
відеоатрибутів за допомогою розбиття на сцени»,
представлену на здобуття ступеня доктора філософії
в галузі знань 12 – Інформаційні технології
за спеціальністю 121 – Інженерія програмного забезпечення

Актуальність теми дисертації.

Обрана для дисертаційного дослідження тема є беззаперечно актуальною. В умовах стрімкого збільшення обсягів відеоданих у глобальному інформаційному просторі, виникає нагальна потреба у створенні та вдосконаленні високопродуктивних методів їх автоматизованої обробки. Ідентифікація відеоатрибутів та сегментація відео на сцени є фундаментальними завданнями, що забезпечують можливість ефективного структурування, індексування, пошуку та узагальнення відеоінформації. Існуючі традиційні підходи до аналізу відео часто не задовольняють сучасні вимоги щодо точності та швидкодії, особливо при роботі з динамічним контентом. Використання методів машинного навчання, зокрема глибоких нейронних мереж та інноваційних архітектур типу візуальних трансформерів, пропонує суттєві переваги для підвищення якості відеоаналізу. Водночас, значна обчислювальна складність таких моделей обмежує їх широке впровадження, передусім у системах, що функціонують в режимі реального часу, або на пристроях з обмеженими апаратними можливостями. Таким чином, розробка нових моделей та програмних рішень, орієнтованих на підвищення швидкодії процесів визначення відеоатрибутів через оптимізацію методів розбиття на сцени та самих моделей, є важливим і своєчасним науково-технічним завданням.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Основним науковим результатом дисертаційної роботи є розроблений здобувачем комплекс моделей та програмних інструментів, спрямованих на підвищення точності та ефективності процесів сегментації відеоданих на сцени та подальшого визначення атрибутів цих сцен. Наукова новизна отриманих результатів полягає у такому:

Вперше розроблено архітектуру розподіленого програмного забезпечення для визначення атрибутів на відео, характерною особливістю якої є оперування відеопотоками для їхнього розбиття відео на плани та сцени, що дозволило збільшити швидкість аналізу відеоконтенту мінімум в 2.5-3 рази.

Вперше розроблено метод для виявлення переходів планів у відеоконтенті на основі поєднання математичних підходів та рекурентних нейронних мереж, який на відміну від існуючих методів швидко та ефективно виділяє просторові та часові ознаки кадрів, що дозволило збільшити точність влучання та F1-оцінку для знаходження зміни планів досягаючи інноваційних результатів.

Вперше розроблено метод виявлення зміни сцени для відеоконтенту з використанням нейронної мережі на основі архітектури візуального трансформеру для відео з застосуванням методу прунінгу перед навчанням, що на відміну від існуючих методів виділяє контекстуальні особливості сцен, що дозволило збільшити F1-оцінку на 5.1% та пришвидшити час виконання на 10%.

Набув подальшого розвитку метод прунінгу перед навчанням для моделей архітектури візуальних трансформерів для відео, який на відміну від існуючих методів враховує важливість механізму «уваги» та дозволяє пришвидшити час виконання моделі на 10%.

Достовірність отриманих у роботі наукових результатів та сформульованих висновків забезпечується застосуванням апробованих сучасних методів дослідження, коректним використанням математичного апарату та підтверджується результатами проведених експериментів. Отримані дані узгоджуються з фундаментальними теоретичними положеннями у сфері машинного навчання та комп'ютерного зору.

Наукові дослідження були виконані здобувачем на базі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» в межах науково-дослідної роботи «Методи і алгоритми оптимізації розпізнавання образів на основі методів машинного навчання» (№0121U109207) під науковим керівництвом к.т.н., доц. Шалденка Олексія Вікторовича та д.т.н., доц. Недашківського Олексія Леонідовича.

Отже, в дисертаційній роботі наукове завдання, що полягає у розробці моделей та програмних засобів для підвищення швидкодії визначення атрибутів у відео за допомогою розбиття на плани та сцени, вирішено у

повному обсязі, а здобувач продемонстрував глибоке володіння методологією наукової діяльності.

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

За своїм змістом дисертаційна робота Здора К.А. цілком відповідає Стандарту вищої освіти за спеціальністю 121 – Інженерія програмного забезпечення та узгоджується з науковими напрямками освітньої програми «Інженерія програмного забезпечення». Праця є цілісним та завершеним науковим дослідженням, яке свідчить про особистий творчий внесок здобувача у розвиток теорії та практики аналізу відеоконтенту.

Ознайомлення з текстом дисертації та звітом подібності дає підстави стверджувати, що робота є результатом самостійних наукових пошуків автора. Всі використані сторонні ідеї, результати та текстові матеріали супроводжуються коректними посиланнями на першоджерела. Порушень принципів академічної доброчесності, таких як плагіат, фабрикація чи фальсифікація, не виявлено.

Мова та стиль викладення результатів.

Дисертацію написано українською мовою. Матеріал роботи викладено структуровано, логічно та послідовно, науковий стиль витримано. Оформлення дисертації дозволяє чітко виділити та сприйняти ключові результати дослідження.

Структурно робота складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел (149 позицій) та трьох додатків. Загальний обсяг роботи становить 178 сторінок, з яких 154 сторінки займає основний текст. Дисертація містить 24 рисунки, 34 формули та 9 таблиць.

У вступі обґрунтовано актуальність теми, сформульовано мету і завдання, визначено об'єкт, предмет, наукову новизну та практичну цінність, а також наведено відомості про апробацію результатів.

У першому розділі здійснено аналіз існуючих методів визначення відеоатрибутів, архітектур моделей для обробки відео та способів їх оптимізації, виявлено обмеження сучасних підходів.

У другому розділі запропоновано удосконалені методи розбиття відео на сцени та плани, зокрема з використанням візуальних трансформерів та гібридних нейро-математичних підходів, та експериментально доведено їх ефективність.

У третьому розділі розроблено та досліджено метод прунінгу для моделей візуальних трансформерів, що враховує особливості механізму уваги та дозволяє оптимізувати моделі зі збереженням точності.

У четвертому розділі представлено архітектуру та реалізацію програмного забезпечення для визначення відеоатрибутів, описано використані технології та інструменти, а також результати тестування швидкодії системи. 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 наукових публікаціях здобувача, серед яких: 4 статті у фахових наукових виданнях України. Із них одна наукова стаття опублікована у періодичному видання, що входить до наукометричної бази Scopus та відноситься до першого — третього квартилів (Q1—Q3) відповідно до класифікації SCImago Journal and Country Rank або Journal Citation Reports.

Також результати дисертації були апробовані на 6 наукових фахових конференціях.

Усі публікації здобувача мають належний науковий рівень, в них достатньо повно описані головні наукові здобутки, представлені в дисертації. Особистий внесок здобувача до всіх наукових публікацій, опублікованих у співавторстві, є вагомим.

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

Недоліки та зауваження до дисертаційної роботи.

1. У першому розділі розглядаючи архітектуру моделей для роботи з відео, зокрема гібридні архітектури, було б корисним включити більш глибокий порівняльний аналіз їхньої масштабованості та вимог до обсягів навчальних даних

2. У другому розділі при описі методу розбиття відео на сцени за допомогою візуального трансформера для відео, зокрема при виборі ключових кадрів з планів, доцільно було б детальніше обґрунтувати вибір саме чотирьох планів як оптимальної кількості для аналізу та дослідити вплив зміни цієї кількості на кінцеву точність та обчислювальну складність.

3. У третьому розділі під час експериментальної оцінки удосконаленого методу прунінгу, окрім впливу на точність (F1-оцінку) та швидкість виконання, доцільно було б також дослідити його вплив на стабільність та швидкість збіжності процесу навчання моделі.

4. У четвертому розділі для кращої демонстрації аналітичних можливостей системи було б доцільно навести декілька прикладів типових

запитів або сценаріїв агрегації даних, які підтримуються розробленою структурою бази даних.

Зазначені зауваження не знижують наукової та практичної цінності роботи і можуть бути враховані здобувачем у подальших дослідженнях.

Висновок про дисертаційну роботу.

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Здора Костянтина Андрійовича на тему «Моделі та програмні засоби підвищення швидкодії визначення відеоатрибутів за допомогою розбиття на сцени» виконана на високому науковому рівні, відповідає встановленим вимогам щодо академічної доброчесності та є завершеним науковим дослідженням. Сукупність отриманих теоретичних та експериментальних результатів становить вирішення актуального наукового завдання, що має суттєве значення для галузі знань 12 – Інформаційні технології.

Дисертаційна робота за актуальністю теми, науковою новизною та практичною значимістю отриманих результатів повністю відповідає вимогам чинного законодавства України, зокрема 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 Здор Костянтин Андрійович заслуговує на присудження ступеня доктора філософії в галузі знань 12 – Інформаційні технології за спеціальністю 121 – Інженерія програмного забезпечення.

Рецензент:

Професор кафедри інженерії програмного
забезпечення в енергетиці

Національного технічного університету

України «Київський політехнічний інститут

імені Ігоря Сікорського»,

доктор технічних наук, доцент



« 16 » 06

