

## **ВІДГУК**

офіційного опонента 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.**

Дисертація «Методи та програмне забезпечення розмітки відеоданих для задач комп'ютерного зору» присвячена автоматизації підготовки розмічених відеоданих для задач машинного навчання, що пов'язані з використанням комп'ютерного зору таких у різних сферах, як робототехніка, медична діагностика, автономне відеоспостереження, автономне управління. Швидке розповсюдження методів штучного інтелекту, зокрема, нейронних мереж на заміну традиційним методам обробки зображень і розпізнавання образів, створює попит на значно більші обсяги і різноманіття спеціалізованих навчальних вибірок розмічених відеоданих, що у свою чергу стимулює інтерес до методів та інструментів, здатних знизити зусилля на розмітку великих масивів даних та вартість розмітки за рахунок кращої автоматизації і більшої селективності їх роботи.

Насамперед йдеться про теоретичний і практичний інтерес з боку наукової спільноти та розробників відповідного програмного забезпечення до нових підходів і методів забезпечення ефективності машинного навчання в сенсі досягнення кращого компромісу між швидкістю роботи моделей та їхньою точністю і стійкістю до змін у вхідних даних через зовнішні фактори такі як процедури відбору зразків (кадрів) для навчання, використання допоміжних заздалегідь навчених моделей штучного інтелекту для розмітки, інтеграції додаткового навчання моделей у процедури автоматизованої розмітки для зниження кількості операцій, що виконує людина.

### **Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.**

Тези дисертаційного дослідження представлені у контексті сучасного стану відповідних напрямків технології розмітки відеоряду і машинного навчання, логічно обґрунтовані та підтверджені результатами експериментальних досліджень з хорошим охопленням відомих альтернатив.

Достовірність наукових результатів підтверджена 4 публікаціями у рецензованих закордонних журналах, а також наведеними даними випробувань

їх програмної реалізації та порівняння з даними відомими з публікацій. Коди програм частково наведені у тексті дисертації.

Наукова новизна результатів дисертаційного дослідження ґрунтовно сформульована у дисертації. Я б виділив наступні елементи:

1. Запропоновано метод поєднання готових моделей автоматичної розмітки з моделями активного навчання з управлінням на основі моніторингу якості розмітки. Автор назвав таку архітектуру дуальною, а відповідний алгоритм - адаптивно-агрегованим навчанням нейромережі. Роль готової моделі може бути більшою, чи меншою у випадках більшої чи меншої типовості розмітки відеоряду, що потенційно дозволяє уникати апріорного вибору інструменту розмітки. Потенційно дуальна архітектура дозволяє знизити участь людини при розмітці відео змішаного характеру за рахунок більшої обчислювальної складності та потреби у більш потужному комп'ютері для досягнення аналогічної швидкодії автоматичних операцій.
2. Запропоновано оригінальний алгоритм пріоритезації складних зразків, що реалізує відомий підхід надання переваги «викидам» у формі попередньої кластеризації та заміни певної кількості зразків з центрів кластерів на зразки, що розташовані на великій відстані від центрів кластерів.
3. Запропоновано високопродуктивний метод узагальнення змісту відеоряду, що автор назвав ітеративним методом вибору ключових кадрів: спочатку у прорідженому відеоряді, а потім у більш щільному, але тільки у найбільш інформативних сегментах. Його особливість на мій погляд – орієнтація на попередню розмітку, що дає можливість формально оцінювати релевантність кадрів/сегментів.
4. Запропоновано використовувати текстової інформації від великих мовних моделей, мовної складової зі звукової доріжки, та/чи субтитрів відео як джерела додаткової інформації при розпізнаванні та описуванні сцен на засадах мультимодального підходу до машинного навчання. В деяких випадках така додаткова інформація може позитивно вплинути на автоматичне визначення цільових об'єктів при розмітці.

#### **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 співпадіння, можна зробити висновок, що дисертаційна робота Зарічкового Олександра Анатолійовича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Збіги в тексті обмежуються співпадінням з власними публікаціями здобувача, що були створені з метою попереднього висвітлення наукових 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.

Дисертація написана з використанням сучасного наукового стилю, що склався в технічних науках, логічно і послідовно. Наукова термінологія в галузі машинного навчання використана вміло і не протирічить вимогам термінологічних стандартів України.

Тези розумно структуровані, складаються з вступу, 6 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації – дуже великий: 261 сторінок (198 без додатків).

У вступі викладено актуальність теми, мету, завдання, об'єкт і предмет дослідження. Розкрито наукову новизну і практичне значення отриманих результатів.

Перший розділ зосереджено на аналізі методів розмітки відеоданих, детекції об'єктів і відповідного програмного забезпечення. Наведено огляд сучасних методів, таких як YOLO і R-CNN, та поставлено завдання дослідження.

Другий розділ присвячено методу пріоритезації складних зразків, що забезпечує підвищення якості набору даних та точності детекції без попередньої розмітки.

У третьому розділі запропоновано ітеративний метод вибору ключових кадрів для довгих відео, що зменшує обсяг даних для обробки і покращує точність сегментації.

Четвертий розділ описує метод агрегації знань у мультимодальних моделях (VLM), який забезпечує точніше розпізнавання складних сцен на відео завдяки генерації додаткових атрибутів.

П'ятий розділ присвячено опису програмного забезпечення, зокрема дуальної архітектури, та огляду інструментів для автоматизації розмітки.

Шостий розділ містить експериментальні результати, оцінку ефективності різних підходів до автоматизації розмітки та висновки щодо досягнення мети роботи.

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 демонструють високий науковий рівень, належну якість та відповідність принципам академічної доброчесності.

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

### **Недоліки та зауваження до дисертаційної роботи.**

1. Експериментальне дослідження ефективності застосування розроблених інструментів, результати якого наведені у розділі 6.4, не витримує критики як з точки зору методології (учасники знали різницю між продуктами та мету порівняння), так і з точки зору кількісних параметрів (лише 10 осіб, лише 10 відео). Я б його зовсім не включав до тез на місті здобувача.
2. У роботі правильно, ґрунтовно і детально висвітлено переваги запропонованих методів, проте нічого не написано про їх недоліки і обмеження. Я бачу якнайменше проблему надлишкового використання обчислювальних ресурсів у дуальній архітектурі, обмеження на тип відеоряду при застосуванні великих мовних моделей (що не здатні, наприклад, озвучувати зображення з медичних зондів). Є певні запитання і до стійкості алгоритму пріоритезації складних зразків. Наявність обмежень і недоліків не знижує наукову цінність результатів, але їх обговорення самим автором відповідає кращим зразкам детального дослідження.
3. Тези – великі, і їх можна було би без шкоди скоротити, майже повністю прибравши оглядовий розділ 1. Більшість розглянутих у ньому питань стосуються тематики роботи радше опосередковано. Натомість якісний огляд наукових робіт, щільно пов'язаних з відповідним дослідженням, наведено в кожному розділі.
4. Європа і світ переходять до відкритої науки, одним з ключових елементів якої є оприлюднення даних і кодів програм у формі, що дозволяє самостійно перевірити отримані наукові результати. Замість додатків до дисертації, відкриті коди розроблених програми разом з інструкцією щодо завантаження тестових даних і потрібними параметрами запуску варто було оприлюднити у GitHub, GitLab, Bitbucket, чи іншому репозитарії.



Вважаю, що висловлені зауваження не є визначальними, не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на мою цілком 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».

### **Офіційний опонент:**

завідувач відділу  
автоматизації програмування (№ 145)  
Інституту кібернетики  
ім. В.М. Глушкова НАН України,  
д. ф.-м. н., с.н.с.



/ Вадим ТУЛЬЧИНСЬКИЙ

« \_\_\_\_ » \_\_\_\_\_ 2025 року