

Голові разової спеціалізованої вченої ради в
Національному університеті «Одеська
політехніка»,
доктору технічних наук, професору кафедри
комп'ютеризованих систем та програмних
технологій
Олександру ФОМІНУ

РЕЦЕНЗІЯ

офіційного рецензента, доктора технічних наук, професора
професора кафедри інформаційних систем
Національного університету «Одеська політехніка»
Щербакової Галини Юріївни
на дисертаційну роботу здобувача Швандта Максима Альбертовича
«Моделі та методи детектування, трекінгу та наскрізної ідентифікації
лабораторних тварин у відеоданих»,
подану до захисту в разову спеціалізовану вчену раду
Національного університету «Одеська політехніка»
Міністерства освіти і науки України
на здобуття наукового ступеня доктора філософії
за спеціальністю 122 – Комп'ютерні науки
галузь знань 12 – Інформаційні технології

1. Актуальність обраної теми дисертації

Дисертаційна робота Швандта М. А. присвячена актуальній задачі автоматизованого аналізу відеоданих із лабораторними тваринами для отримання кількісних характеристик їхньої поведінки. Тривале відеоспостереження дає змогу визначати координати й індивідуальні траєкторії тварин, тривалість їх перебування у функціональних зонах та показники взаємодії між особинами. Значні обсяги відеоданих ускладнюють ручне опрацювання та зумовлюють потребу в його автоматизації.

За наявності змінного фону, відблисків світла, невисокої контрастності та оклюзій можливе накопичення помилок трекінгу. Похибки локалізації спричиняють викривлення траєкторій, втрати об'єктів – розриви треків, а помилкова асоціація – спотворення індивідуальної статистики. Тому актуальним є

узгодження процедур сегментації, детектування, супроводу та повторної ідентифікації.

Дисертація відповідає зазначеній постановці, оскільки поєднує інтерпретований метод суперпиксельної обробки, багатоканальну нейромережеву сегментацію та навчання представлень треклетів у єдиній технологічній послідовності. Запропоновані рішення забезпечують належну якість автоматизованого аналізу відеоданих за умови використання локальних обчислювальних ресурсів і реалізовані у спеціалізованому застосунку. Отже, обрана тема є актуальною для розвитку методів комп'ютерного зору та відповідає змісту спеціальності 122 – Комп'ютерні науки.

2. Оцінка обґрунтованості наукових положень, висновків і рекомендацій

Обґрунтованість наукових положень забезпечується послідовною декомпозицією загальної задачі на локалізацію, сегментацію, формування траєкторій і відновлення ідентичності. Для їх розв'язання автор поєднав цифрову обробку зображень, суперпиксельну сегментацію, згорткові нейронні мережі та асоціацію треклетів. Кожне удосконалення зменшує вплив окремих похибок лабораторної сцени та виконує визначену функцію в інтегрованому конвеєрі.

Експериментальна перевірка побудована поетапно: оцінено багат шаровий простір ознак і контекстну фільтрацію SLIC; досліджено двомагістральну Mask R-CNN з керованим SE-злиттям; проаналізовано формування 256-вимірних нормалізованих дескрипторів та асоціацію сегментованих треклетів; інтегрований конвеєр випробувано на синтетичній і двох реальних відеопослідовностях. Це дає змогу співвіднести кількісні показники з відповідними науковими рішеннями.

Програмна реалізація й апробація у лабораторних дослідженнях та навчальному процесі додатково підтверджують практичну обґрунтованість результатів. Сукупність теоретичного пояснення, експериментів і впровадження дає підстави визнати положення, висновки та рекомендації достатньо обґрунтованими.

3. Наукова новизна одержаних результатів

Наукова новизна дисертації визначається послідовним розвитком засобів локалізації та збереження ідентичності лабораторних тварин на трьох взаємопов'язаних рівнях: суперпiксельної обробки зображення, нейромережевого видiлення екземплярiв i асоцiацiї треклетiв. Запропонованi рiшення узгоджуються в межах цiлiсного конвеєра аналізу вiдеоданих. Формулювання результатiв, винесених на захист, вiдповiдають змiсту розроблених методiв i моделей та пiдтверджуються наведеними в роботi експериментами. До основних результатiв, що мають ознаки наукової новизни, належать:

– *удосконалено* метод суперпiксельної сегментацiї SLIC за рахунок використання багат шарового простору ознак, який базується на LAB-моделi зображення, враховує рiзницеве зображення та результати перевiрки однорiдностi пiкселiв у локальному оточеннi, що дозволило пiдвищити якiсть локалiзацiї лабораторних тварин у вiдеоданих за умов нестацiонарного фону та наявностi артефактiв;

– *удосконалено* метод детектування лабораторних тварин у вiдеоданих шляхом використання методу суперпiксельної сегментацiї SLIC i контекстної фiльтрацiї, що дозволило пiдвищити якiсть детектування лабораторних тварин за умов нестацiонарного фону;

– *удосконалено* нейромережеву модель Mask R-CNN за рахунок введення подвiйної магiстралi екстракцiї ознак iз керованим SE-злиттям ($B \rightarrow A$), що дозволило пiдвищити якiсть детектування та сегментацiї лабораторних тварин у вiдеоданих за умов часткових перекриттiв, злиття та подальшого роздiлення об'єктiв зi збереженням прийнятної оперативностi;

– *вперше* запропоновано концептуальну модель наскрiзної iдентифiкацiї лабораторних тварин у вiдеоданих без використання ручного маркування Mask-TAUDL, яка iнтегрує механiзм отримання ознак на основi удосконаленої нейромережевої моделi Mask R-CNN iз подвiйною магiстраллю екстракцiї ознак, масково-зваженi ROI-ознаки та механiзми асоцiацiї треклетiв без використання ручного маркування TAUDL, що забезпечує формування нормалiзованих векторiв

вбудовувань і стабільне збереження ідентичності лабораторних тварин протягом відеопослідовності.

4. Повнота викладення результатів дисертації в опублікованих працях

За темою дисертації опубліковано 12 праць: 6 статей у фахових виданнях України (3 у категорії «А», що індексуються у базі Scopus, і 3 у категорії «Б») та 6 матеріалів міжнародних наукових конференцій. Публікації охоплюють основні результати роботи й етапи їх апробації.

Тематика праць відображає розвиток дослідження від аналізу детектування і трекінгу та процедур попередньої обробки до суперпіксельної локалізації, модифікації Mask R-CNN і повторної ідентифікації за сегментованими треклетами. Положення наукової новизни й експериментальні результати оприлюднено належним чином; особистий внесок здобувача у спільні праці визначено в дисертації. Повнота викладення результатів є достатньою.

5. Значення роботи для науки та практики

Наукове значення дисертації полягає у розвитку узгодженого підходу до аналізу відеоданих, який пов'язує покадрову локалізацію, сегментацію екземплярів і часову асоціацію треклетів. Запропоновані рішення можуть бути використані як основа для подальших досліджень контрольованих динамічних сцен:

- удосконалений метод суперпіксельної сегментації розширює можливості інтерпретованої локалізації об'єктів за нестационарного фону та наявності артефактів;
- двомагістральна Mask R-CNN демонструє спосіб залучення додаткових каналів без переходу до надмірно ресурсомісткої магістралі;
- концептуальна модель Mask-TAUDL забезпечує формування масково-зважених дескрипторів і асоціацію треклетів без ручних глобальних ID-міток на етапі навчання відповідного модуля;

Практичне значення роботи полягає у створенні спеціалізованого застосунку для формування траєкторій, показників активності та контактів тварин. Його використано у дослідженнях біологічного факультету ОНУ імені І. І. Мечникова, а результати роботи впроваджено в освітній процес двох університетів.

Практичну ефективність підтверджують приріст якості SLIC більш ніж на 6 %, показники $mAP50 = 0,87$ для рамок і $0,90$ для масок та перевірка інтегрованого конвеєра на послідовностях тривалістю 900, 3200 і 4320 кадрів.

6. Відсутність (наявність) порушень академічної доброчесності

Порушень академічної доброчесності в дисертації та наукових публікаціях, у яких висвітлено основні наукові результати дисертації, не виявлено.

Усі результати, винесені автором на захист, отримані самостійно та відображено в опублікованих роботах. У роботах, опублікованих у співавторстві, використані тільки ті ідеї, положення та результати розрахунків, які є результатом особистих наукових пошуків.

7. Дискусійні положення та зауваження щодо змісту дисертації

Попри загалом високий науковий та методичний рівень роботи, окремі положення дисертації потребують додаткового уточнення або можуть бути розвинені в подальших дослідженнях:

1) у дисертації детектування, трекінг і повторна ідентифікація розглядаються як складові цілісного конвеєра. Водночас механізм покадрового трекінгу та його взаємодію з процедурою формування треклетів описано недостатньо докладно. Доцільно чіткіше визначити критерії асоціації детекцій у сусідніх кадрах і внесок автора в цей етап;

2) у концептуальній моделі Mask-TAUDL формується 256-вимірний L2-нормалізований вектор вбудовування. Було б корисно додатково обґрунтувати вибір його розмірності з урахуванням якості пошуку та обчислювальних витрат;

3) результати повторної ідентифікації показують різне співвідношення показників mAP і $Rank-1$ для магістралей ResNet18 та ResNet50. Доцільно було б докладніше пояснити, який із цих критеріїв є визначальним для цільового сценарію зшивання треклетів і чому перевагу надано конфігурації на базі ResNet18;

4) якість детектування та сегментації удосконаленої Mask R-CNN переважно оцінено за показником $mAP50$. Використання $mAP75$ або усередненого mAP за різних порогів IoU дало б змогу повніше оцінити точність геометрії рамок і масок;

5) у четвертому розділі twin-ResNet18 для формування дескрипторів Re-ID ініціалізовано вагами сегментаційної моделі. Вплив обраної стратегії перенесення параметрів висвітлено стисло. Порівняння повного донавчання із заморожуванням частини шарів дозволило б точніше відокремити внесок попереднього навчання від внеску запропонованої моделі Mask-TAUDL;

6) під час формування навчальних, валідаційних і тестових наборів використовуються кадри відеопослідовностей та перетворені варіанти вихідних зображень. Корисним було б чіткіше зазначити, як виключалася поява часово близьких або похідних кадрів у різних частинах вибірки, оскільки це може впливати на оцінку якості детектування та сегментації моделлю Mask R-CNN;

7) додаткові канали другої магістралі формуються різними процедурами попередньої обробки та мають відмінні статистичні властивості. Детальніший опис їх нормалізації та просторового узгодження з каналами першої магістралі підвищив би відтворюваність двомагістральної моделі Mask R-CNN;

8) в роботі наявні деякі незначні синтаксичні неточності. Наприклад, не зрозуміло, що саме позначено як “чи є i також найкращим партнером для знайденого $j^*(i)$ ” (с.146).

Зазначені зауваження не впливають на загальний висновок щодо наукового рівня дисертації та можуть бути враховані автором у подальшій роботі.

8. Загальний висновок про відповідність роботи встановленим вимогам

Дисертаційна робота Швандта Максима Альбертовича на тему «Моделі та методи детектування, трекінгу та наскрізної ідентифікації лабораторних тварин у відеоданих» є завершеним самостійним науковим дослідженням, яке спрямоване на забезпечення цілісного переходу від виділення тварини на окремому кадрі до формування узгоджених індивідуальних траєкторій у тривалих відеозаписах. Удосконалений SLIC, двомагістральна Mask R-CNN і концептуальна модель Mask-TAUDL утворюють взаємопов'язану систему, у якій результати кожного етапу підвищують надійність наступного. Експериментальна перевірка, програмна реалізація та впровадження підтверджують завершеність і практичну придатність роботи.

Враховуючи викладене, вважаю, що дисертаційна робота Швандта Максима Альбертовича на тему «Моделі та методи детектування, трекінгу та наскрізної ідентифікації лабораторних тварин у відеоданих» відповідає вимогам Порядку підготовки здобувачів вищої освіти ступеня доктора філософії та доктора наук у ЗВО (наукових установах), затвердженого Постановою Кабінету Міністрів України від 23 березня 2016 року № 261 (зі змінами) та вимогам п.п. 5 – 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 12.01.2022 №44 (зі змінами).

Автор дисертаційної роботи Швандт Максим Альбертович заслуговує на присудження наукового ступеня доктора філософії за спеціальністю 122 – Комп’ютерні науки, галузь знань 12 – Інформаційні технології.

Офіційний рецензент:

Доктор технічних наук, професор,
професор кафедри інформаційних систем
Національного університету «Одеська
Політехніка»

Галина ЩЕРБАКОВА