

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

РЕЦЕНЗІЯ

кандидата технічних наук, доцента кафедри інформаційних систем
інституту комп'ютерних систем
Національного університету «Одеська політехніка»,
Годовиченка Миколи Анатолійовича,
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 Максима Альбертовича присвячена розв'язанню актуальної науково-практичної задачі підвищення якості та стійкості детектування, трекінгу та наскрізної ідентифікації лабораторних тварин у відеоданих за умов нестаціонарного фону, оклюзій, злиття та розділення об'єктів.

Актуальність обраного напрямку дослідження визначається зростаючими потребами у застосуванні засобів комп'ютерного зору для автоматизації

біологічних і медико-біологічних експериментів. Автоматизоване відеоспостереження дозволяє реєструвати активність та переміщення лабораторних тварин, їх взаємодію, реакції на зовнішні стимули, формувати траєкторії руху та розраховувати кількісні поведінкові характеристики.

Незважаючи на контрольований характер лабораторного середовища, відеодані можуть містити відблиски, тіні, шуми, локальні зміни освітлення та інші артефакти. Додаткові труднощі виникають при одночасному спостереженні за декількома тваринами внаслідок їх часткового або повного перекриття, злиття та подальшого розділення. У таких випадках помилки детектування та локалізації призводять до фрагментації траєкторій, а помилки асоціації – до перемикання ідентифікаторів об'єктів.

Сучасні нейромережеві методи здатні забезпечити високу якість детектування та сегментації, однак їх практичне застосування для довготривалого відеомоніторингу потребує врахування вимог до обчислювальних ресурсів, стабільності навчання на обмежених наборах даних та оперативності обробки.

У зв'язку з цим розроблення моделей і методів, які поєднують детектування, сегментацію, трекінг та повторну ідентифікацію лабораторних тварин у єдиному конвеєрі обробки відеоданих, є актуальним науковим завданням, що відповідає сучасним напрямкам розвитку комп'ютерного зору та спеціальності 122 «Комп'ютерні науки».

2. Ступінь обґрунтованості наукових положень, висновків і рекомендацій

Обґрунтованість наукових положень, висновків та рекомендацій дисертаційної роботи забезпечується комплексним використанням сучасних методів цифрової обробки зображень і відео, сегментації, машинного та глибокого навчання, методів побудови й злиття ознак, трекінгу та повторної ідентифікації об'єктів.

Автором проведено ґрунтовний аналіз класичних і нейромережових підходів до детектування, трекінгу та збереження ідентичності об'єктів, що дозволило визначити основні проблеми їх використання у лабораторних відеосценах та обґрунтувати напрями подальших досліджень.

Для підвищення якості локалізації лабораторних тварин запропоновано удосконалення методу суперпiксельної сегментації SLIC із використанням багат шарового простору ознак. Для детектування та сегментації екземплярів удосконалено нейромережову модель Mask R-CNN шляхом введення подвійної магістралі екстракції ознак із керованим SE-злиттям. Для наскрізного збереження ідентичності об'єктів запропоновано концептуальну модель Mask-TAUDL із використанням сегментованих треклетів та масково-зважених ROI-ознак.

Достовірність отриманих результатів підтверджується проведеними експериментальними дослідженнями, кількісним порівнянням розроблених рішень із базовими методами та моделями, використанням загальноприйнятих метрик оцінювання якості детектування, сегментації та повторної ідентифікації, а також практичною апробацією інтегрованого конвеєра.

Таким чином, наукові положення, висновки й рекомендації дисертаційної роботи є достатньо обґрунтованими та підтверджуються результатами експериментальних досліджень.

3. Наукова новизна одержаних результатів

Дисертаційна робота містить нові науково обґрунтовані результати, спрямовані на підвищення якості та стійкості автоматизованої обробки відеоданих лабораторних тварин.

До основних результатів наукової новизни дисертаційної роботи належать:

– удосконалено метод суперпiксельної сегментації SLIC за рахунок використання багат шарового простору ознак, який базується на LAB-моделі зображення, враховує різницеве зображення та результати перевірки однорідності пікселів у локальному оточенні, що дозволило підвищити якість локалізації

лабораторних тварин у відеоданих за умов нестаціонарного фону та наявності артефактів;

– удосконалено метод детектування лабораторних тварин у відеоданих шляхом використання методу суперпiксельної сегментації SLIC і контекстної фільтрації, що дозволило підвищити якість детектування лабораторних тварин за умов нестаціонарного фону;

– удосконалено нейромережеву модель Mask R-CNN за рахунок введення подвійної магістралі екстракції ознак із керованим SE-злиттям ($B \rightarrow A$), що дозволило підвищити якість детектування та сегментації лабораторних тварин у відеоданих за умов часткових перекриттів, злиття та подальшого розділення об'єктів зі збереженням прийнятної оперативності;

– вперше запропоновано концептуальну модель наскрізної ідентифікації лабораторних тварин у відеоданих без використання ручного маркування Mask-TAUDL, яка інтегрує механізм отримання ознак на основі удосконаленої нейромережевої моделі Mask R-CNN із подвійною магістраллю екстракції ознак, масково-зважені ROI-ознаки та механізми асоціації треклетів без використання ручного маркування TAUDL, що забезпечує формування нормалізованих векторів вбудовувань і стабільне збереження ідентичності лабораторних тварин протягом відеопослідовності.

Наукові результати логічно пов'язані між собою та утворюють основу інтегрованого конвеєра детектування, трекінгу та наскрізної ідентифікації лабораторних тварин.

4. Структура та зміст дисертаційної роботи

Дисертаційна робота має логічну структуру та складається зі вступу, п'яти розділів, висновків, списку використаних джерел і додатків.

У вступі обґрунтовано актуальність теми дисертації, сформульовано мету та задачі дослідження, визначено об'єкт, предмет і методи дослідження, наведено

наукову новизну та практичне значення одержаних результатів, представлено інформацію щодо особистого внеску здобувача та апробації результатів.

У першому розділі проведено аналіз методів та підходів до детектування, трекінгу і збереження ідентичності лабораторних тварин у відеоданих. Розглянуто класичні методи комп'ютерного зору, сучасні нейромережеві детектори, сегментаційні моделі та методи трекінгу. Визначено переваги та обмеження існуючих підходів, а також сформовано вимоги до побудови інтегрованого конвеєра обробки відеоданих.

У другому розділі розроблено та досліджено методи детектування й локалізації лабораторних тварин на основі суперпиксельної сегментації SLIC. Запропоновано удосконалення методу SLIC шляхом використання багатопланового простору ознак та перевірки локальної однорідності. Експериментальна перевірка показала підвищення якості детектування приблизно на 6,8 % порівняно з базовим методом SLIC.

У третьому розділі запропоновано удосконалену нейромережеву модель Mask R-CNN із подвійною магістраллю ResNet18 та керованим SE-злиттям. Проведено порівняння з базовими конфігураціями Mask R-CNN із різними екстракторами ознак. Для запропонованої моделі отримано $mAP50 = 0,87$ для обмежувальних рамок і $mAP50 = 0,90$ для масок сегментації за середнього часу виведення близько 105,5 мс.

У четвертому розділі розроблено концептуальну модель Mask-TAUDL для наскрізної ідентифікації лабораторних тварин без використання ручного ID-маркування. Модель використовує сегментовані треклети, масково-зважені ROI-ознаки та 256-вимірні нормалізовані вектори вбудовувань. Експериментальні результати підтверджують працездатність запропонованого підходу до повторної ідентифікації.

У п'ятому розділі розроблено та експериментально перевірено інтегрований конвеєр, який поєднує попередню обробку відеоданих, детектування, сегментацію, арбітраж результатів, трекінг, формування треклетів та їх подальшу

асоціацію. Тестування виконано на синтетичній контрольованій сцені та реальних лабораторних відеозаписах із мишами та бичками.

Висновки дисертації відповідають поставленим задачам та системно узагальнюють основні результати дослідження.

5. Практичне значення одержаних результатів

Практичне значення дисертаційної роботи полягає у доведенні розроблених моделей і методів до рівня інтегрованого програмного конвеєра автоматизованої обробки лабораторних відеоданих.

Запропоновані рішення дозволяють здійснювати детектування та сегментацію лабораторних тварин, формувати їх траєкторії, зберігати ідентичність об'єктів після перекриттів і фрагментації треків та створюють основу для подальшого обчислення поведінкових характеристик.

На основі запропонованих моделей та методів розроблено спеціалізований програмний застосунок для автоматизованої обробки відеоспостережень за лабораторними тваринами.

Результати роботи впроваджено у лабораторні дослідження, що проводяться на базі біологічного факультету Одеського національного університету імені І. І. Мечникова, а також у навчальний процес Національного університету «Одеська політехніка» та Одеського національного університету імені І. І. Мечникова.

Наведене підтверджує практичну значущість отриманих здобувачем результатів.

6. Повнота викладення результатів дисертації в опублікованих працях та дотримання принципів академічної доброчесності

Основні результати дисертаційного дослідження достатньо повно представлені в наукових публікаціях здобувача.

За матеріалами дисертації опубліковано 12 наукових праць, серед яких 6 статей у наукових фахових виданнях, три з яких опубліковано у виданнях, представлених у міжнародній наукометричній базі Scopus, а також 6 публікацій апробаційного характеру у матеріалах наукових конференцій.

Тематика публікацій охоплює основні складові дисертаційного дослідження – аналіз методів детектування і трекінгу, попередню обробку відеоданих, суперпіксельну сегментацію, модифікацію Mask R-CNN та повторну ідентифікацію на основі сегментованих треклетів.

Використані в дисертації ідеї, положення та результати інших авторів супроводжуються відповідними посиланнями на джерела. За результатами вивчення представлених матеріалів підстав для сумнівів щодо дотримання здобувачем принципів академічної доброчесності не встановлено.

7. Дискусійні положення та зауваження щодо змісту дисертації

Наведені нижче зауваження мають переважно дискусійний та рекомендаційний характер і не знижують загальної позитивної оцінки дисертаційної роботи.

1. В першому розділі наведено ґрунтовний аналіз класичних і нейромережевих методів детектування, сегментації та трекінгу. Водночас результати проведеного аналізу викладено переважно у текстовій формі та не узагальнено у вигляді єдиної порівняльної таблиці або схеми, яка б наочно відображала зв'язок між виявленими обмеженнями існуючих підходів та вимогами до компонентів запропонованого інтегрованого конвеєра.

2. В другому розділі удосконалений метод SLIC передбачає використання низки параметрів, зокрема бажаної кількості суперпікселів, просторового кроку, радіуса локальної околиці, порогу локальної однорідності та параметрів контекстного порогування. При цьому в роботі не наведено розгорнутого аналізу чутливості отриманих результатів до зміни зазначених параметрів та їх залежності від масштабу об'єктів і характеристик відеосцени.

3. В третьому розділі базові варіанти Mask R-CNN навчалися протягом 30 епох, тоді як запропонована модель із подвійною магістраллю та SE-злиттям – протягом 50 епох. Така відмінність у тривалості навчання дещо ускладнює відокремлення ефекту запропонованих архітектурних змін від впливу збільшення кількості епох. Крім того, у роботі окремо не оцінено внесок додаткових інформаційних каналів, другої магістралі та механізму SE-злиття у загальне підвищення показників якості.

4. Експериментальна перевірка інтегрованого конвеєра у п'ятому розділі проведена на трьох тестових відеопослідовностях. Такий обсяг експериментальної вибірки підтверджує працездатність запропонованих рішень у розглянутих сценаріях, однак обмежує можливість оцінювання їх узагальнюваності для ширшого набору умов освітлення, конфігурацій лабораторних арен та інших видів лабораторних тварин.

5. В четвертому розділі обґрунтовано адаптацію механізмів TAUDL до однокамерної постановки шляхом інтерпретації окремих сеансів або фрагментів довготривалого відеозапису як умовних доменів. Разом із тим, у роботі недостатньо конкретизовано практичні правила розбиття відеозапису на такі сеанси, вибору порогу подібності або кількості найближчих сусідів під час міжтреклетної асоціації, а також критерії оцінювання та зважування надійності коротких треклетів.

6. В п'ятому розділі оцінювання інтегрованого конвеєра переважно виконується за системою проксі-метрик, що обумовлено відсутністю повної покадрової еталонної розмітки ідентифікаторів. Водночас у роботі не наведено результатів оцінювання хоча б на окремій повністю розміченій контрольній підвибірці із застосуванням стандартних метрик багатооб'єктного трекінгу та збереження ідентичності, що дещо ускладнює безпосереднє зіставлення отриманих результатів з іншими підходами.

7. В тексті дисертаційної роботи трапляються окремі редакційні та технічні неузгодженості, які не впливають на зміст отриманих результатів. Зокрема,

перелік умовних позначень не охоплює значну кількість скорочень, що використовуються у роботі.

Зазначені зауваження мають переважно дискусійний та рекомендаційний характер, не ставлять під сумнів достовірність отриманих результатів, не впливають на сформульовані положення наукової новизни та не знижують наукової і практичної значущості дисертаційної роботи.

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 цілісною, самостійною та завершеною кваліфікаційною науковою працею, в якій одержано нові науково обґрунтовані результати, що мають наукове та практичне значення.

У дисертації вирішено актуальну науково-практичну задачу підвищення якості та стійкості детектування, трекінгу та наскрізної ідентифікації лабораторних тварин у відеоданих за умов нестаціонарного фону, оклюзій, злиття та розділення об'єктів.

Поставлені в дисертації задачі вирішено, сформульована мета дослідження досягнута. Наукові положення та висновки достатньо обґрунтовані, отримані результати мають наукову новизну та практичну цінність, а основні результати дослідження відображено в наукових публікаціях здобувача.

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 пунктів 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 січня 2022 року № 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.

Доцент кафедри інформаційних систем
інституту комп’ютерних систем
Національного університету
«Одеська політехніка»,
кандидат технічних наук

Микола ГОДОВИЧЕНКО