

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра Інформатики
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

ПРОГРАМНІ ЗАСОБИ АНАЛІЗУ ЗОБРАЖЕНЬ ОБ'ЄКТІВ
ДЛЯ УПРАВЛІННЯ ФОНОМ

(тема)

Виконав:
студент 4 курсу, групи ІТІНФ-20-3

Шокун А.С.
(прізвище, ініціали)

Спеціальності 122 Комп'ютерні науки
(код і повна назва спеціальності)

Тип програми освітньо-професійна

Освітня програма Інформатика
(повна назва освітньої програми)

Керівник проф. Гороховатський В.О.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Кобилін О.А.
(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)Кафедра Інформатики
(повна назва)Рівень вищої освіти перший (бакалаврський)Спеціальність 122 Комп'ютерні науки
(код і повна назва)Тип програми освітньо-професійнаОсвітня програма Інформатика
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«____» _____ 2024 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУстудентові Шокун Анні Сергіївні
(прізвище, ім'я, по батькові)1. Тема роботи Програмні засоби аналізу зображень об'єктів для управління фоном

затверджена наказом університету від 20 травня 2024 року № 464 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 01 червня 2024 р.

3. Вихідні дані до роботи науково-методична та науково-технічна література, матеріали конференцій, дані інтернет-мережі, теорія оброблення і аналізу зображень, методи аналізу фону на зображенні, нейромережеві засоби управління фоном.

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Актуальність та сфера застосування засобів аналізу зображень об'єктів для управління фоном.2. Аналіз принципу роботи мережі.3. Нейромережеві засоби управління фоном.4. Розробка програмного забезпечення.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Актуальність проблеми розробки програмного забезпечення для аналізу зображень об'єктів для управління фоном, постановка задачі, тестові зображення, аналіз результатів моделювання.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	08.04.2024	
2	Аналіз завдання, підбір літератури	08.04.24-15.04.24	
3	Аналіз літератури з досліджуваної проблеми	16.04.24-19.04.24	
4	Аналіз принципу роботи мережі	20.04.24-29.04.24	
5	Розробка програмного забезпечення	30.04.24-19.05.24	
6	Програмна реалізація	20.05.24-23.05.24	
7	Оформлення пояснювальної записки	24.05.24-31.05.24	
8	Перевірка на плагіат	01.06.24	
9	Рецензування	02.06.24	
10	Підготовка презентації та доповіді	03.06.24-09.06.24	
11	Занесення роботи в електронний архів	10.06.24	
12	Попередній захист кваліфікаційної роботи	10.06.24	

Дата видачі завдання 8 квітня 2024 р.

Студент _____
(підпис)

Керівник роботи _____ проф. Гороховатський В.О.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ/ABSTRACT

Пояснювальна записка до кваліфікаційної роботи: 52 с., 5 рис., 1 дод., 54 джерела.

УПРАВЛІННЯ ФОНОМ, НЕЙРОННІ МЕРЕЖІ, СЕГМЕНТАЦІЯ, ГАУСОВЕ РОЗМИТТЯ, ВІДЕОПОТІК, ВІДОБРАЖЕННЯ ВІДЕОСИГНАЛУ.

Об'єктом роботи є методи розмиття фону в онлайн-відеопотоках за допомогою розробки спеціалізованої нейронної мережі.

Метою роботи є підвищення конфіденційності та естетичності в реальному часі для відеозв'язку шляхом вибіркового розмиття фону, зберігаючи при цьому об'єкт на передньому плані в чіткому фокусі.

Застосовано передові методів обробки зображень, включаючи згорткові нейронні мережі, а також методи гаусового розмиття. Продуктивність системи була оцінена за результатом програмного моделювання.

Результати роботи роблять внесок у сферу комп'ютерного зору та обробки цифрових медіа.

BACKGROUND CONTROL, NEURAL NETWORKS, SEGMENTATION, GAUSSIAN BLUR, VIDEO FLOW, VIDEO SIGNAL MAPPING.

The object of the work is methods of blurring the background in online video streams using the development of a specialized neural network.

The goal of the work is to improve real-time privacy and aesthetics for video communication by selectively blurring the background while keeping the foreground object in sharp focus.

Advanced image processing techniques are applied, including convolutional neural networks, as well as Gaussian blurring methods. The performance of the system was evaluated based on the results of the software simulation.

The results of the work contribute to the field of computer vision and digital media processing.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	6
Вступ.....	7
1 Актуальність та сфера застосування.....	8
1.1 Передумови та еволюція обробки відео в реальному часі	8
1.2 Важливість розмиття фону в онлайн-відеоспілкуванні	9
1.3 Сфери застосування.....	11
1.4 Сучасні технології та їхні обмеження	13
1.5 Постановка задачі	15
2 Аналіз принципу роботи мережі	16
2.1 Розуміння нейронних мереж у комп'ютерному зорі.....	16
2.2 Нейронні мережі у розмитті фону.....	18
2.3 Механізм виявлення силуету людини	19
2.4 Принципи розмиття фону: технічні деталі та методи обробки зображень	22
2.5 Проблеми обробки в реальному часі	24
3 Розробка та демонстрація програмного забезпечення.....	27
3.1 Основні компоненти програми.....	27
3.2 Функція розпізнавання людини.....	29
3.3 Реалізація розмиття фону: технічні деталі	31
3.4 Виконання в реальному часі та аналіз ефективності	33
Висновки	35
Перелік джерел посилання	37
Додаток А Лістинг програми	43

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CNN – Convolutional Neural Network (згорткові нейронні мережі)

FPS – Frames Per Second (кадри в секунду)

GPU – Graphics Processing Unit (графічний процесор)

TPU – Tensor Processing Unit (процесор тензорних обчислень)

CV – Computer Vision (комп'ютерний зір)

AI – Artificial Intelligence (штучний інтелект)

ML – Machine Learning (машинне навчання)

DL – Deep Learning (глибоке навчання)

ROI – Region Of Interest (область інтересу)

OPENCV – бібліотека комп'ютерного зору

PYTORCH – бібліотека машинного навчання на основі бібліотеки Torch

TENSORFLOW – фреймворк машинного навчання

ВСТУП

Стрімкий розвиток цифрових комунікаційних технологій висунув численні виклики та можливості у сфері обробки відео, зокрема, у покращенні користувацького досвіду завдяки підвищенню якості зображення та конфіденційності [1-4]. Ця кваліфікаційна робота зосереджений на ключовому аспекті цієї технологічної еволюції: принципі розмиття фону в онлайн-відеоспілкуванні. Метою роботи є всебічний аналіз цього принципу та розробка спеціальної нейронної мережі, здатної ефективно виконувати це завдання.

Методологія проєктування охоплює поєднання теоретичного аналізу та практичних експериментів. Це включає огляд існуючої літератури та технологій, а також розробку та тестування нейромережевої моделі. Основні методи включають комп'ютерне моделювання, алгоритмічний аналіз та емпіричне тестування за різних умов експлуатації.

Актуальність цього дослідження зумовлена зростаючим попитом на сучасні інструменти обробки відео в різних галузях. Зі збільшенням поширеності віддаленої роботи та віртуальних зустрічей потреба в технологіях, які можуть забезпечити конфіденційність і професійну презентацію в різних умовах, є більш вираженою, ніж будь-коли. Крім того, ці технології відіграють важливу роль у створенні контенту, де естетична якість має першорядне значення [5-9]. Таким чином, розробка спеціалізованої нейронної мережі для розмиття фону є своєчасною і практичною потребою в цифровому комунікаційному ландшафті.

1 АКТУАЛЬНІСТЬ ТА СФЕРА ЗАСТОСУВАННЯ

1.1 Передумови та еволюція обробки відео в реальному часі

Сфера обробки відео в реальному часі зазнала значних змін, пройшовши значний шлях від свого зародження до нинішнього стану. Ця еволюція є не тільки свідченням технологічного прогресу, але також підкреслює зростаючий попит на складну обробку відео в різних додатках. Мета цього розділу – підкреслити актуальність технологічного прогресу обробки відео в реальному часі і методології, що застосовуються для його вдосконалення [1-4].

Принцип розмиття фону при обробці відео в реальному часі є відносно недавньою розробкою, яка набуває популярності з поширенням відеоконференцій і зростаючою потребою в конфіденційності та естетичному контролі в цифровому спілкуванні. Ця технологія ґрунтується на складних алгоритмах, здатних ідентифікувати та відокремлювати об'єкти на передньому плані (як правило, людей) від їхнього фону в режимі реального часу. Процес передбачає застосування методів комп'ютерного зору за допомогою моделей машинного навчання для точного виявлення та ізоляції об'єкта. Згодом використовуються алгоритми цифрової обробки сигналів, щоб застосувати ефект розмиття до фону, зберігаючи при цьому чіткість об'єкта на передньому плані.

Ці зміни в обробці відео в реальному часі були зумовлені поєднанням кількох факторів. Зростаючий попит на сучасні засоби відеозв'язку в професійній та особистій сферах був значним рушієм. Крім того, розвиток нейронних мереж і алгоритмів глибокого навчання надав необхідні інструменти для вирішення складних завдань обробки зображень, які раніше були недосяжними в режимі реального часу [5, 6].

Крім того, еволюція апаратного забезпечення, зокрема графічних процесорів (GPU), мала вирішальне значення. Ці спеціалізовані процесори забезпечили можливості паралельної обробки, необхідні для інтенсивних

обчислень при обробці відео в реальному часі. Інтеграція такого обладнання зі складними програмними алгоритмами дозволила безперешкодно та ефективно виконувати такі завдання, як розмиття фону, що вимагає безперервної та швидкої обробки великих обсягів відеоданих.

Обробка відео в реальному часі має важливе застосування в таких сферах, як спостереження, автономні транспортні засоби та медична візуалізація. У кожній з цих сфер здатність обробляти та інтерпретувати відеодані в режимі реального часу може мати вирішальне значення для операційної ефективності та безпеки.

Таким чином, шлях обробки відео в реальному часі від аналогових методів до сучасних цифрових технологій відображає поєднання досягнень в галузі інформатики, штучного інтелекту і розвитку апаратного забезпечення. Конкретне застосування розмиття фону в онлайн-відеозв'язку є прямим результатом цих досягнень, задовольняючи сучасні потреби в приватності, естетиці та професіоналізмі в цифровій взаємодії. Цей історичний контекст закладає основу для розуміння поточного стану технології та її потенційної майбутньої траєкторії, особливо в тому, що стосується розробки та застосування спеціальних нейронних мереж для обробки відео [6-9].

1.2 Важливість розмиття фону в онлайн-відеоспілкуванні

Поява платформ для відеозв'язку в Інтернеті докорінно змінила спосіб, у який ми спілкуємося, працюємо та навчаємося. Однак ця повсюдна присутність також створила проблеми, пов'язані з конфіденційністю, професіоналізмом і зосередженістю. Розмиття фону стає потужним інструментом для вирішення цих проблем і покращення загального досвіду онлайн відеозв'язку.

Однією з основних переваг розмиття фону є його здатність захищати особисту інформацію, ненавмисно розкрити під час відеодзвінків. У домашніх

офісах або спільних житлових приміщеннях на захарашеному фоні можуть відобразитися конфіденційні документи, сімейні фотографії або навіть пробіски сусідів по кімнаті чи домашніх улюбленців. Розмиття приховує ці деталі, надаючи людям відчуття контролю над своїм особистим простором і запобігаючи ненавмисному порушенню конфіденційності. Це особливо важливо для віддалених працівників, які працюють з конфіденційною інформацією, або для людей, які занепокоєні потенційним стеженням [10-13].

У професійному середовищі відеодзвінки є звичним явищем для зустрічей, презентацій та співбесід. Підтримання професійного іміджу має вирішальне значення, і розмиття фону може суттєво сприяти цьому. Усуваючи відволікаючі або захарашені фони, воно дозволяє повністю зосередитися на мовцеві та його повідомленні. Це робить аудиторію більш уважною та зацікавленою, сприяє кращій комунікації та передає відчуття компетентності та професіоналізму.

Розмиття фону також може створити більш інклюзивне середовище для різних користувачів. Деякі люди можуть мати захарашене або безладне тло через особисті обставини або обмежений життєвий простір. Це може призвести до сором'язливості або небажання брати участь у відеодзвінках. Розмиття пом'якшує ці проблеми, дозволяючи людям відчувати себе комфортно і впевнено, не турбуючись про своє фонове оточення.

Окрім практичних переваг, розмиття фону очевидно покращує загальний досвід користувача під час відеозв'язку в Інтернеті. Дослідження показали, що розмитий фон створює більш естетичне і заспокійливе візуальне середовище, зменшуючи стрес і втому очей під час тривалих відеодзвінків. Це може сприяти підвищенню задоволеності користувачів та їхньої взаємодії з платформою.

Хоча розмиття фону пропонує різні переваги, важливо визнати потенційні обмеження і майбутні міркування. Технологічний прогрес необхідний для досягнення більшої точності та надійності в обробці складних фонів, динамічних умов освітлення та об'єктів, що швидко рухаються. Крім

того, для забезпечення відповідального застосування технології розмиття фону необхідно враховувати етичні міркування щодо конфіденційності та потенційних зловживань [14, 15].

Отже, розмиття фону стало цінним інструментом в онлайн-відеоспілкуванні, пропонуючи практичні та психологічні переваги для користувачів. Захищаючи конфіденційність, підвищуючи професіоналізм, покращуючи фокус і сприяючи інклюзивності, воно значно покращує загальний досвід спілкування. Однак для усунення обмежень і забезпечення етичного впровадження цієї технології необхідні постійні дослідження і розробки, оскільки вона продовжує відігравати вирішальну роль у майбутньому онлайн-спілкуванні.

1.3 Сфери застосування

У цифровій комунікації конфіденційність є першочерговим питанням. Поява технології розмиття фону при обробці відео в реальному часі безпосередньо вирішує цю проблему. Завдяки вибіркового затушовуванню фону, конфіденційна інформація, чи то особисті речі, індикатори місцезнаходження або інші особи, що знаходяться поблизу, залишаються поза увагою. Цей аспект особливо важливий у таких сценаріях, як віддалена робота, телемедицина та онлайн-освіта, де недоторканність приватного простору і конфіденційність інформації мають вирішальне значення.

Крім того, в системах відеоспостереження і безпеки ця технологія полегшує захист приватного життя в громадських місцях. Хоча камери спостереження мають важливе значення для безпеки, вони часто фіксують людей, які не є об'єктами інтересу. Розмиття фону в реальному часі може бути використане для анонімізації осіб у цих стрічках, таким чином балансує між потребами безпеки та правами на приватність.

Естетичний аспект відеозв'язку набув значного значення, особливо з поширенням створення контенту і трансляції на цифрових платформах. Технологія розмиття заднього фону дозволяє підвищити візуальну привабливість відео, створюючи ефект глибини різкості, який часто зустрічається у професійному кінематографі. Цей ефект гарантує, що об'єкт відео знаходиться в різкому фокусі, тоді як фон художньо розмитий, тим самим привертаючи увагу глядача до об'єкта.

У створенні контенту, особливо на таких платформах, як відеоблоги, вебінари та прямі трансляції, розмиття фону допомагає створити більш відшліфований і професійний вигляд. Це дозволяє творцям контенту підтримувати постійну візуальну якість, незалежно від фізичного середовища. Ця можливість є безцінною в ситуаціях, коли фон захаращений, відволікає або не є естетично привабливим [16].

У професійному середовищі, наприклад, під час віртуальних зустрічей та інтерв'ю, оточення, зображене на відео, може суттєво вплинути на сприйняття професіоналізму. Технологія розмиття фону дозволяє людям презентувати себе в більш контрольованій та професійній манері. Вона усуває візуальні відволікаючі фактори і гарантує, що увага буде зосереджена на людині та її змісті, а не на навколишньому середовищі. Цей аспект особливо корисний для людей, які працюють вдома або у спільних приміщеннях, де контроль над фоном оточенням обмежений.

Крім того, у сценаріях обслуговування клієнтів і взаємодії з ними, таких як віртуальні консультації або віддалена підтримка клієнтів, розмиття фону допомагає підтримувати єдиний і професійний візуальний стандарт. Це допомагає створити середовище, сприятливе для сфокусованої та ефективної комунікації, вільне від візуального шуму та відволікаючих чинників [17-19].

Окрім цих основних сфер, сфера застосування розмиття фону в обробці відео в реальному часі поширюється на різні інші сектори. У телемедицині воно підвищує конфіденційність пацієнта і фокусує увагу на медичному працівнику або конкретній області, що його цікавить. У навчальних відео та

онлайн-навчанні вона допомагає звести до мінімуму відволікаючі фактори, тим самим покращуючи навчальний процес. Технологія також знаходить застосування в середовищах віртуальної та доповненої реальності, де її можна використовувати для плавного поєднання реальних і віртуальних елементів [20].

Сфера застосування технології розмиття фону дуже широка, її інтеграція відображає не лише технологічний прогрес, але й є відповіддю на зростаючі потреби цифрової комунікації та медіа-виробництва. З подальшим розвитком цієї технології її застосування розширюватиметься.

1.4 Сучасні технології та їхні обмеження

Розмиття фону в режимі реального часу в онлайн-відеоспілкуванні залежить від поєднання потужних сучасних технологій, кожна з яких має свої переваги та недоліки. Розуміння цих проблем має вирішальне значення для розробки та вдосконалення методів розмиття фону.

Нині в основі розмиття фону лежать згорткові нейронні мережі (CNN). Навчені на величезних масивах даних з міченими зображеннями, ці мережі ефективно ідентифікують і сегментують людський передній план від фону. Однак залишаються обмеження. ШНМ можуть не впоратися зі складним фоном (наприклад, захаращеність кімнати, листя), динамічними умовами освітлення та об'єктами, що швидко рухаються. Крім того, обчислювальні вимоги до висновків у реальному часі на пристроях з обмеженими ресурсами можуть бути високими [21-25].

Окрім CNN, спеціальні алгоритми виявлення об'єктів можуть підвищити точність розмиття фону. Такі алгоритми, як YOLO і Faster R-CNN, ще більше уточнюють силует людини, особливо для великих планів або об'єктів, що перекривають один одного. Однак ці алгоритми призводять до

додаткових обчислювальних витрат і можуть бути чутливими до оклюзії або відеопотоків низької роздільної здатності.

Різні методи обробки зображень доповнюють основні нейронні мережі та алгоритми. Гаусове розмиття, виявлення країв та імітація глибини різкості відіграють важливу роль у згладжуванні артефактів, покращенні країв об'єктів та створенні природних ефектів розмиття. Однак вибір оптимальної комбінації методів залежить від конкретних уподобань користувача та обчислювальних обмежень [26-28].

Основними перешкодами на шляху до ідеального виконання розмиття фону в реальному часі є:

- неправильна ідентифікація об'єктів на передньому плані, особливо волосся або складних деталей, може призвести до неприродних артефактів розмиття, а оклюзія або динамічний фон можуть ще більше знизити точність;

- складність CNN і алгоритмів виявлення об'єктів може призвести до перевантаження обчислювальної потужності, особливо на пристроях низького класу, що впливає на продуктивність у реальному часі та час роботи від батареї;

- варіації освітлення (тіні, віддзеркалення) і складне оточення (зовнішні сцени, текстуровані стіни) можуть заплутати алгоритми розмиття, що призводить до неточностей;

- зловживання технологією розмиття фону для стеження або порушення приватності, що вимагає відповідальної розробки та впровадження [29-31].

Постійні дослідження і розробки спрямовані на усунення описаних вище обмежень і досягнення більш надійного та ефективного розмиття фону:

- дослідження нових архітектур ШНМ, спеціально розроблених для розмиття фону, з акцентом на зменшенні обчислювальної складності та підвищенні точності для різних сценаріїв;

- розробка більш ефективних алгоритмів виявлення об'єктів з мінімальним впливом на обчислювальну потужність для пристроїв з обмеженими ресурсами;

- динамічне налаштування параметрів розмиття на основі аналізу сцени в реальному часі для підвищення точності та природності;

- встановлення чітких етичних принципів і найкращих практик для розробки та впровадження технологій розмиття фону, щоб захистити конфіденційність користувачів і запобігти зловживанням [32].

Обмеження, пов'язані з точністю, обчислювальними витратами і факторами навколишнього середовища, продовжують створювати проблеми. Подолання цих обмежень за допомогою майбутніх досліджень і розробок прокладе шлях до більш плавних, ефективних і етично відповідальних рішень для розмиття фону, що ще більше збагатить досвід онлайн-спілкування.

1.5 Постановка задачі

Дослідження присвячено аналізу та розробці нейронної мережі для розмиття фону у відео-спілкуванні в реальному часі.

Об'єктом роботи є методи розмиття фону в онлайн-відеопотоках за допомогою розробки спеціалізованої нейронної мережі.

Метою роботи є підвищення конфіденційності та естетичності в реальному часі для відеозв'язку шляхом вибіркового розмиття фону, зберігаючи при цьому об'єкт на передньому плані в чіткому фокусі.

Для досягнення мети необхідно вирішити такі задачі:

- застосувати передові методи обробки зображень, включаючи згорткові нейронні мережі (CNN), оцінку глибини, а також методи гауссового розмиття для розробки нейромережевої моделі;

- здійснити програмну реалізацію;

- оцінити продуктивність системи за різних умов.

2 АНАЛІЗ ПРИНЦИПУ РОБОТИ МЕРЕЖІ

2.1 Розуміння нейронних мереж у комп'ютерному зорі

Нейронні мережі, особливо у сфері комп'ютерного зору, зробили революцію в тому, як машини інтерпретують і розуміють візуальні дані. У цьому розділі розглядаються принципи роботи нейронних мереж у комп'ютерному зорі, з'ясовується їхня структура, функціонування та застосування у таких задачах, як розпізнавання зображень, виявлення об'єктів і, зокрема, розмиття фону при обробці відео [33, 34].

Нейронні мережі, натхненні біологічними нейронними мережами, які складають мозок тварин, є серією алгоритмів, які розпізнають основні взаємозв'язки в наборі даних за допомогою процесу, що імітує роботу людського мозку. У контексті комп'ютерного зору ці мережі призначені для інтерпретації візуальних даних шляхом розпізнавання патернів і характеристик на зображеннях і відео.

Типова нейронна мережа складається з шарів взаємопов'язаних вузлів або нейронів, кожен з яких призначений для виконання певних перетворень над вхідними даними. Ці шари включають вхідний шар, один або кілька прихованих шарів і вихідний шар. Вузли в кожному шарі з'єднані за допомогою ваг, і ці зв'язки передають сигнал від одного вузла до іншого. Сила цих зв'язків регулюється під час навчання, що є вирішальним процесом у розвитку здатності нейронної мережі робити точні прогнози або класифікації [33-38].

Спеціалізованою формою нейронних мереж, яка широко використовується в комп'ютерному зорі, є згорткові нейронні мережі (CNN) (рис. 2.1).

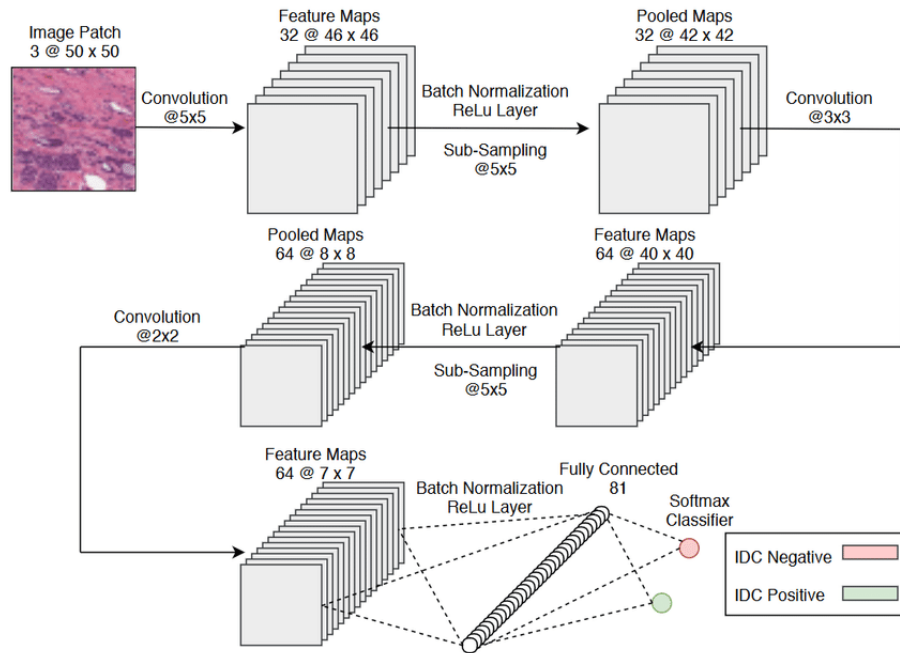


Рисунок 2.1 – Базова архітектура CNN

ШНМ особливо вправні в обробці піксельних даних і використовуються в різних завданнях, починаючи від класифікації зображень і закінчуючи складними завданнями, такими як виявлення та сегментація об'єктів. ШНМ використовують математичну операцію, відому як згортка, яка включає фільтр або ядро, що пропускає вхідні дані (наприклад, зображення), створюючи карту особливостей, яка виділяє специфічні особливості, такі як краї, текстури або форми (рис. 2.2) [35-39].

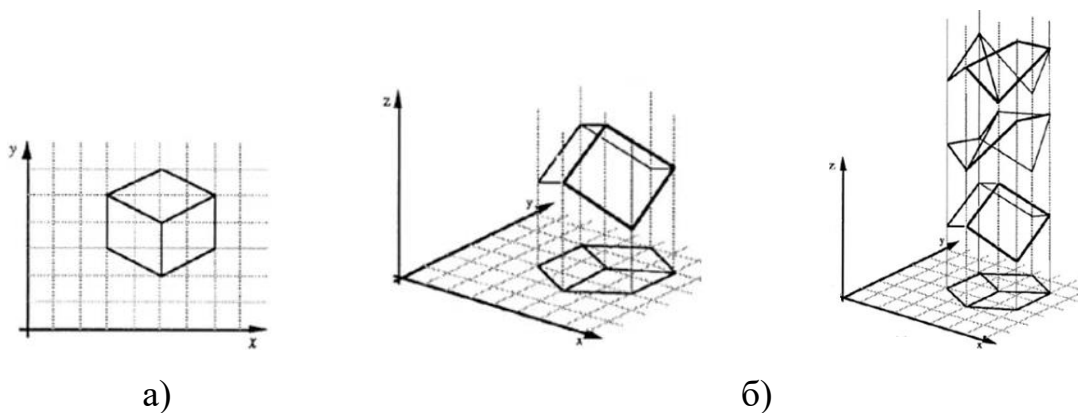


Рисунок 2.2 – Типи згортки:

а) звичайна згортка з наступною висхідною дискретизацією; б) атрозисна згортка, що призводить до більш щільної карти ознак

Архітектура CNN зазвичай включає кілька шарів, зокрема згорткові шари, об'єднані шари та повністю з'єднані шари. Згорткові шари застосовують різні фільтри до вхідних даних, об'єднувальні шари зменшують просторовий розмір (а отже, і обчислювальне навантаження) представлення, а повністю з'єднані шари обчислюють оцінки класів, що призводить до кінцевого результату [39, 40].

2.2 Нейронні мережі у розмитті фону

У контексті розмиття фону при обробці відео в реальному часі нейронні мережі відіграють ключову роль. Завдання полягає насамперед у сегментації переднього плану (зазвичай це людина) від фону, а потім у застосуванні ефекту розмиття до фону. Ця сегментація є формою класифікації зображення на рівні пікселів, де кожен піксель класифікується як такий, що належить до переднього або заднього плану.

Для цього нейронну мережу навчають на великих наборах даних, що складаються із зображень або відео з позначеними переднім і заднім планами. Навчання передбачає подачу мережі вхідних даних (зображень або відеокадрів) і налаштування вагових коефіцієнтів зв'язків у мережі, щоб мінімізувати різницю між виходом мережі та фактичними мітками. Процес навчання використовує алгоритми зворотного поширення та градієнтного спуску для ітеративного налаштування ваг.

Після навчання нейронна мережа може обробляти нові відеокадри, ідентифікувати передній план і відокремлювати його від фону в режимі реального часу. Сегментований передній план потім поєднується з розмитим фоном, в результаті чого створюється відеокадр, де об'єкт знаходиться у фокусі, а фон розмитий [41-46].

Застосування нейронних мереж для розмиття фону стикається з кількома проблемами. Вони включають забезпечення швидкості обробки в реальному

часі, обробку різноманітних і динамічних фонів, роботу з різними умовами освітлення і управління розмиттям при русі. Щоб вирішити ці проблеми, останні досягнення були зосереджені на оптимізації архітектури нейронних мереж, розробці більш ефективних алгоритмів навчання та використанні апаратного прискорення, такого як графічні процесори [40, 42].

Крім того, досліджуються такі інновації, як використання датчиків глибини та інтеграція з технологіями доповненої реальності, щоб підвищити продуктивність і можливості нейронних мереж у додатках для розмиття фону.

Нейронні мережі, зокрема CNN, стали фундаментальними в галузі комп'ютерного зору, пропонуючи безпрецедентні можливості в аналізі зображень і відео. Їх застосування для розмиття фону при обробці відео в реальному часі свідчить про їх універсальність і ефективність. Оскільки дослідження в цій галузі продовжують розвиватися, очікується, що нейронні мережі відіграватимуть все більш важливу роль у різних сферах застосування, розширюючи межі можливого в комп'ютерному зорі та цифровій обробці зображень.

2.3 Механізм виявлення силуету людини

Виявлення людського силуету, важливий крок у розмитті фону в реальному часі, передбачає точне відокремлення людини на передньому плані від фонового оточення. Ця складна задача покладається на різні алгоритми та архітектури моделей, кожен з яких має свої сильні та слабкі сторони.

Класичними алгоритмами для виявлення людського силуету є блюрення фону. Цей класичний підхід порівнює послідовні відеокадри, визначаючи пікселі, які значно змінюються, як такі, що належать рухомому об'єкту на передньому плані (людині). Хоча цей метод простий і ефективний, він погано справляється зі статичним фоном або коли кілька об'єктів рухаються одночасно.

Алгоритм визначення порогових значень аналізує інтенсивність пікселів, відокремлюючи пікселі переднього плану (яскравіші за встановлений поріг) від фону. Однак цей метод чутливий до змін освітлення та складного фону. Приклад ядра простого фільтра нижніх частот:

$$L = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \quad (2.1)$$

Алгоритми на кшталт виявлення країв Кенні зосереджуються на виявленні різких змін контрасту на зображенні, які часто відповідають межам об'єктів. Хоча вони корисні для базового виділення силуетів, їх можна легко ввести в оману через шум або складні текстури [43-47].

Сегментація на основі ознак виокремлює із зображення специфічні ознаки, такі як колір, текстура та рух, щоб відрізнити передній план від фону. Він пропонує більшу гнучкість, ніж простіші методи, але вимагає ретельного вибору ознак і налаштування параметрів (рис. 2.3).

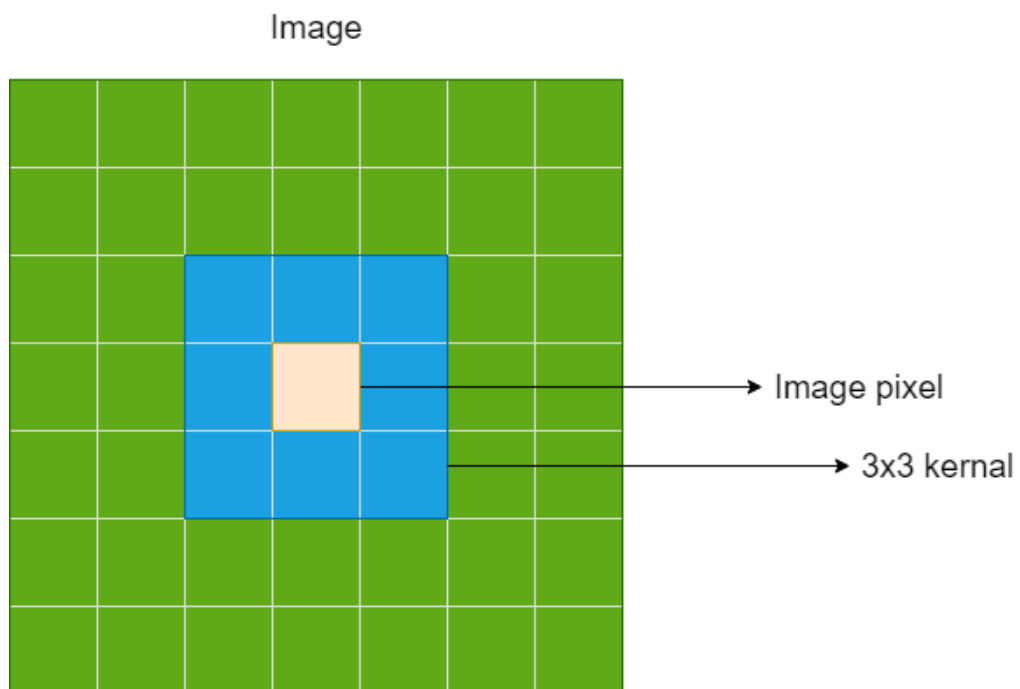


Рисунок 2.3 – Матриця зображення 7×7 та ядро 3×3, що застосовується до пікселів зображення

Поява глибинного навчання зробила революцію в розпізнаванні людських силуетів, запропонувавши більш надійні та точні рішення. Згорткові нейронні мережі (CNN) стали домінуючою силою, ефективно вивчаючи складні закономірності та взаємозв'язки в даних зображення, щоб відрізнити людей від їхнього оточення.

Одною із популярних архітектур ШНМ для розпізнавання людських силуетів є повністю згорткові мережі (FCN). Ці мережі обробляють все зображення одночасно, генеруючи попіксельні прогнози ймовірності переднього/заднього плану. Популярні варіанти FCN, такі як U-Net і SegNet, використовують пропускні з'єднання для збереження просторової інформації та досягнення точної сегментації.

Архітектури кодер-декодер використовують частину кодера для вилучення ознак із зображення, а частину декодера – для підвищення частоти дискретизації та реконструкції отриманої маски, виділяючи людський силует (рис. 2.4). ResNets і DenseNets є поширеними прикладами цієї архітектури.

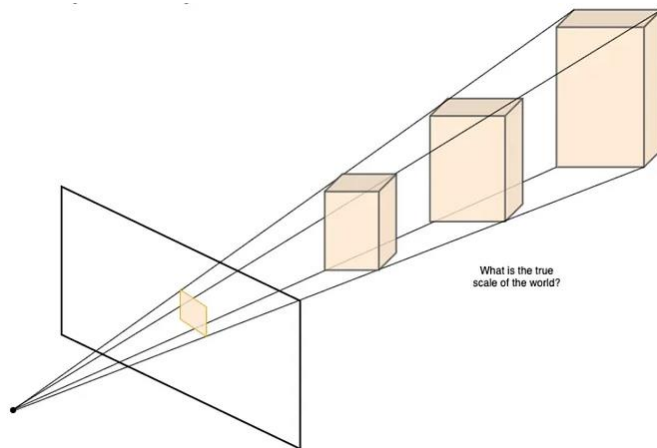


Рисунок 2.4 – Архітектура кодера-декодера з ASPP

Механізми уваги фокусують увагу мережі на відповідних ділянках зображення, підвищуючи точність для складних сценаріїв, таких як оклюзія або складний фон [48-50].

Згортання при атрофії зображено на формулі 2.2.

$$y[i] = \sum_{k=1}^K x[i + r \cdot k] \omega[k], \quad (2.2)$$

де при $r > 1$ – атрофічне згортання;

$r = 1$ – стандартне згортання.

2.4 Принципи розмиття фону: технічні деталі та методи обробки зображень

Розмиття фону у відеопотоках передбачає сегментацію зображення на дві основні складові: передній план, що зазвичай містить об'єкт зйомки, і задній план, який потрібно розмити. Ця сегментація має вирішальне значення і формує основу, на якій будується процес розмиття. Основними методами, які використовуються в цьому процесі, є сегментація зображення, оцінка глибини та застосування фільтрів розмиття.

Сегментація зображення – це процес розбиття цифрового зображення на декілька сегментів (наборів пікселів) для спрощення або зміни представлення зображення на щось більш змістовне і легше для аналізу. У контексті розмиття фону сегментація має на меті точно відрізнити об'єкт на передньому плані від фону (рис. 2.5).

Для сегментації використовуються різні алгоритми, серед яких нейронні мережі, зокрема згорткові нейронні мережі (CNN), відіграють ключову роль. Ці мережі навчаються на великих наборах даних з міченими зображеннями, щоб вивчити закономірності та особливості, які відрізняють елементи переднього плану від фону. Для досягнення високого ступеня точності використовуються передові методи, такі як семантична сегментація, де кожен піксель класифікується за категоріями [51, 52].

Оцінка глибини – ще один важливий компонент процесу розмиття фону. Вона передбачає визначення відстані кожного пікселя або об'єкта на зображенні від камери. Точна оцінка глибини дозволяє зробити розмиття більш реалістичним і ефективним, оскільки допомагає визначити, до якої міри слід розмити кожен частину фону.

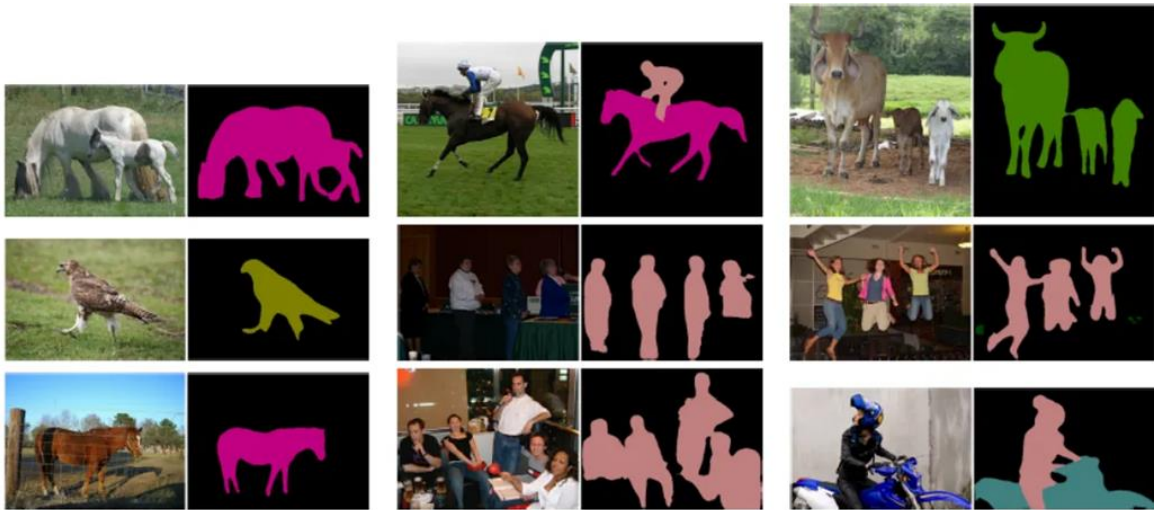


Рисунок 2.5 – Семантична сегментація за допомогою Deeplabv3

Після сегментації та оцінки глибини наступним кроком є застосування фільтрів розмиття до тла. Гаусове розмиття є однією з найпоширеніших технік для цієї мети. Він передбачає використання функції Гаусса для створення ефекту розмиття, що дає природний ефект розфокусування, який нагадує ефект об'єктива камери [53].

Ступінь та інтенсивність розмиття залежать від інформації про глибину. Пікселі або області, віддалені від площини фокусування (передній план), розмиваються інтенсивніше порівняно з тими, що розташовані ближче до площини. Ця градація створює ефект глибини різкості, додаючи візуальної привабливості та реалістичності відео.

Впровадження розмиття фону в обробку відео в реальному часі створює додаткові проблеми, пов'язані, насамперед, з обчислювальною ефективністю та швидкістю. Обробка в реальному часі вимагає, щоб ці завдання –

сегментація, оцінка глибини і розмиття – виконувалися протягом мілісекунд, щоб уникнути помітного відставання відеопотоку.

Для цього застосовуються різні методи оптимізації. Вони включають використання ефективних архітектур нейронних мереж, які забезпечують баланс між точністю та обчислювальним навантаженням, застосування апаратних прискорювачів, таких як графічні процесори, та оптимізацію алгоритмів для паралельної обробки. Крім того, такі методи, як квантування та обрізання моделі, використовуються для зменшення розміру моделі та обчислювальних вимог без значного впливу на продуктивність.

2.5 Проблеми обробки в реальному часі

Затримка, тобто затримка між входом (відеопотоком), що обробляється, і виходом (розмитим відео), що відображається, є першочерговою проблемою при обробці відео в реальному часі. Мета полягає в тому, щоб мінімізувати затримку для забезпечення безперебійної та інтерактивної взаємодії з користувачем. Висока затримка може порушити потік комунікації, спричинити дезорієнтацію і знизити якість взаємодії.

Кілька факторів впливають на затримку розмиття фону. До них відносяться обчислювальна складність використовуваних алгоритмів, обчислювальна потужність обладнання та ефективність програмної реалізації. Складні операції, такі як нейромережевий висновок, сегментація зображення та оцінка глибини, є обчислювально інтенсивними і можуть спричинити значні затримки, якщо їх не оптимізувати [54].

Для зменшення затримок застосовуються різні підходи. Вони включають оптимізацію архітектури нейронних мереж для швидшого виведення (наприклад, використання полегшених моделей), застосування апаратного прискорення, такого як GPU і TPU, а також оптимізацію коду для

паралельної обробки для ефективного використання декількох ядер сучасних процесорів.

Точність у контексті розмиття фону означає точність, з якою система може відокремити передній план від фону і застосувати ефект розмиття належним чином. Неточності в цьому процесі можуть призвести до таких артефактів, як розмиття частин об'єкта, залишення частин фону нерозмитими або непослідовне розмиття по всьому кадру.

Точність системи залежить від ефективності алгоритмів, що використовуються для сегментації та оцінки глибини. Ці алгоритми повинні бути стійкими до різних викликів, таких як мінливі умови освітлення, складні фони та різні види об'єктів. Тому навчання нейронних мереж відіграє вирішальну роль, вимагаючи різноманітних і всеосяжних наборів даних, які охоплюють широкий спектр сценаріїв.

Крім того, робота в режимі реального часу вимагає, щоб ці алгоритми були не тільки точними, але й швидкими, що часто може бути суперечливою вимогою. Баланс між швидкістю і точністю є ключовим викликом при розробці та впровадженні цих систем.

Фактори навколишнього середовища, такі як умови освітлення, складність фону та якість камери, суттєво впливають на продуктивність систем розмиття фону в реальному часі. Різні умови освітлення можуть впливати на видимість і контрастність об'єкта на тлі, впливаючи на точність сегментації. Складні фони з візерунками або рухомими елементами можуть збити систему з пантелику, що призведе до помилкової сегментації.

Для вирішення цих проблем застосовуються вдосконалені алгоритми, оснащені адаптивними можливостями. Ці алгоритми можуть підлаштовувати свої параметри у відповідь на зміни умов навколишнього середовища, зберігаючи продуктивність у різних сценаріях. Крім того, такі методи, як стиснення динамічного діапазону і шумозаглушення, використовуються для оптимізації відеопотоку для кращої обробки.

Проблеми затримки, точності та впливу навколишнього середовища при обробці відео в реальному часі вирішуються завдяки постійному технологічному прогресу. Розробка більш ефективних нейромережових моделей, вдосконалення апаратних прискорювачів і підвищення алгоритмічної ефективності є постійними зусиллями в цій галузі.

Крім того, перспективним напрямком є інтеграція штучного інтелекту з периферійними обчисленнями, коли деякі завдання обробки даних перекладаються на периферійні пристрої, зменшуючи навантаження на центральні сервери і потенційно зменшуючи затримки. Очікується, що використання технології 5G також відіграватиме значну роль у зменшенні затримок у хмарних системах обробки відео.

На закінчення, хоча обробка розмиття фону у відеопотоках в режимі реального часу пропонує значні переваги з точки зору конфіденційності та естетики, вона не позбавлена певних проблем. Питання затримки, точності та факторів навколишнього середовища мають вирішальне значення для визначення ефективності цих систем. Вирішення цих проблем вимагає багатогранного підходу, що поєднує досягнення в обчислювальних алгоритмах, апаратних технологіях і мережових можливостях. Очікується, що в міру подальшого розвитку цих технологій ефективність і застосовність розмиття фону в реальному часі у відеозв'язку значно покращиться, що зробить його незамінним інструментом у різних сферах застосування.

3 РОЗРОБКА ТА ДЕМОНСТРАЦІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

3.1 Основні компоненти програми

Розробка програми на мові Python для реалізації розмиття фону у відеопотоках у реальному часі є важливою складовою цієї роботи. Опишемо детально архітектуру, структуру коду та основні функції розробленої програми, з'ясуємо механізми її роботи та застосовані обчислювальні стратегії.

Програма має модульну структуру і складається з кількох ключових компонентів, які працюють у тандемі для досягнення розмиття фону в реальному часі. Основні модулі включають:

- FPSmetric: клас, призначений для вимірювання та відображення кількості кадрів в секунду (FPS) в реальному часі, забезпечуючи важливий зворотній зв'язок щодо продуктивності обробки відео;
- Base: основний двигун (рушій), який обробляє відеопотоки з різних джерел, включаючи вебкамери та відеофайли;
- MPSegmentation: клас, що використовує рішення машинного навчання MediaPipe для сегментації об'єкта від фону в режимі реального часу;
- Main: основний скрипт, який ініціалізує та запускає конвеєр обробки з використанням вищезазначених компонентів.

Клас FPSmetric допомагає відстежувати продуктивність системи обробки відео в реальному часі. Він обчислює кількість кадрів в секунду, вимірюючи інтервал часу між послідовними викликами. Ця інформація має вирішальне значення для оцінки ефективності системи та гарантування того, що обробка відео працює в межах прийнятних порогових значень продуктивності.

Основними функціями класу FPSmetric є:

- розрахунок FPS на основі ковзної середньої;

- відображає FPS на відеокадрі для моніторингу в реальному часі;
- налаштування властивості відображення, такі як положення, шрифт і колір.

Базовий клас `base.py` є основою програми, що відповідає за обробку відеопотоків. Він може обробляти дані з різних джерел, таких як прямі трансляції з вебкамери, відеофайли або окремі зображення. Цей клас є гнучким, що дозволяє інтегрувати користувацькі об'єкти обробки, які можуть бути застосовані до кожного кадру відео.

Основними функціями класу `FPSmetric` є:

- обробка потоків з вебкамери з можливістю відображення та керування в реальному часі;
- обробка та збереження вихідних відеофайлів з настроюваним діапазоном кадрів;
- реалізація користувацьких функцій обробки, що полегшує інтеграцію функцій сегментації та розмиття фону.

Клас `MPSegmentation` використовує рішення `MediaPipe` для сегментації селфі для розділення переднього і заднього плану в реальному часі. Ця сегментація є ключовою для точного застосування ефекту розмиття фону.

Основними функціями класу `FPSmetric` є:

- застосовує розмиття за Гаусом або власне зображення в якості фону;
- сегментація на основі порогових значень для розрізнення переднього і заднього планів;
- реалізація користувацьких функцій обробки, що полегшує інтеграцію функцій сегментації та розмиття фону.

Головний скрипт (`main.py`) організовує весь конвеєр обробки. Він ініціалізує необхідні об'єкти, зокрема модулі FPS-метрики та сегментації, і запускає цикл обробки. Скрипт розроблено універсальним, що дозволяє легко модифікувати його для пристосування до різних джерел відео та вимог до обробки.

Деталі реалізації:

- скрипт інтегрує модулі FPSmetric і MPSegmentation з базовим класом;
- він дозволяє обробляти потоки з вебкамери в реальному часі з метриками продуктивності;
- налаштовувані порогові значення і коефіцієнти розмиття в сегментації для індивідуальної продуктивності.

3.2 Функція розпізнавання людини

Функція розпізнавання людей в умовах розмиття фону в реальному часі є важливим аспектом розробленої програми на мові Python. Розглянемо детально механізми, алгоритми та стратегії реалізації, що використовуються для точного розпізнавання та сегментації людських об'єктів на фоні відеопотоків.

Розпізнавання людей в обробці відео передбачає ідентифікацію та сегментацію людських об'єктів на фоні їхнього оточення. Це завдання є важливим для додатків з розмиттям фону, оскільки воно визначає область кадру, яка повинна залишатися у фокусі, а решта розмивається. Точне розпізнавання людей є складним завданням через такі фактори, як мінливі умови освітлення, різноманітні зовнішності людей та складні фони.

Алгоритмічний підхід до розпізнавання людей у цьому проєкті використовує методи машинного навчання, зокрема моделі глибокого навчання, завдяки їхній високій точності та ефективності в задачах сегментації зображень. Обраною моделлю для цієї мети є згортова нейронна мережа (CNN), оптимізована для обробки в реальному часі.

Лістинг 3.1 Метод `__call__` для обчислення frame:

```
def __call__(self, frame: np.ndarray = None) -> ndarray | None | float:
    self._prev_frame_time = self._frame_time
```

```

self._frame_time = time.time()
if not self._prev_frame_time:
    return frame

self._fps_list.append(1 / (self._frame_time - self._prev_frame_time))
self._fps_list = self._fps_list[-self._range_average:]
fps = float(np.average(self._fps_list))
if frame is None:
    return fps

cv2.putText(frame, str(int(fps)), self.position, self.fontFace, self.fontScale,
self.color, self.thickness,
            self.lineType)
return frame

```

Модель обрано на основі балансу між точністю та обчислювальною ефективністю. Вона попередньо навчається на великих наборах даних для розпізнавання людських фігур у різних сценаріях. Модель CNN обробляє кожен кадр відео, щоб класифікувати пікселі на дві категорії: людина і фон. Ця класифікація базується на вивчених особливостях, таких як форми, краї та текстури.

Реалізація розпізнавання людей інтегрована в клас MPSegmentation програмного застосунку. Цей клас використовує рішення машинного навчання MediaPipe, які включають надійні моделі для сегментації людини:

- інтеграція з MediaPipe пропонує ефективне і точне рішення для сегментації людей, клас MPSegmentation використовує цю можливість для виконання сегментації в режимі реального часу;

- завдяки параметрам кастомізації реалізація дозволяє налаштовувати поріг сегментації та вибір моделі, ця гнучкість забезпечує адаптивність до різних операційних середовищ;

- процес сегментації оптимізовано для роботи в режимі реального часу, він передбачає мінімальну затримку, забезпечуючи плавність та інтерактивність відеопотоку;
- стійкість моделі до різних умов освітлення підвищується завдяки інтенсивному навчанню та коригуванню порогу сегментації в реальному часі;
- здатність моделі CNN розрізняти людей на складному фоні досягається завдяки навчанню на різноманітних наборах даних;
- модель призначена для обробки динамічних змін у положенні та русі людини, зберігаючи точну сегментацію протягом усього відео.

Функція розпізнавання людини в розробленому застосунку відіграє ключову роль в ефективності функції розмиття фону. Використовуючи передові моделі глибокого навчання та оптимізуючи їх для роботи в режимі реального часу, програма успішно відокремлює людей від фону у відеопотоках. Ця функція має важливе значення для збереження фокусу на людині під час застосування ефекту розмиття до фону, тим самим підвищуючи конфіденційність і візуальну привабливість у різних додатках, таких як відеоконференції та створення контенту.

3.3 Реалізація розмиття фону: технічні деталі

Розмиття фону включає два основні технічні кроки: сегментацію об'єкта зйомки від фону, як описано в попередньому розділі, і застосування ефекту розмиття до сегментованого фону. Процес має бути достатньо ефективним, щоб працювати в режимі реального часу без шкоди для якості відеопотоку.

Основним методом, що використовується для розмиття фону, є Гаусове розмиття, широко застосовуване в обробці зображень. Воно забезпечує природний ефект розмиття, подібний до об'єктива камери.

Гаусове розмиття застосовується до всього кадру, а потім зверху накладається оригінальне, нерозмите зображення людини. Такий підхід гарантує, що розмивається лише фон, а об'єкт залишається у фокусі.

Функціонал розмиття фону інтегровано у клас `MPSegmentation`.

Лістинг 3.2 Клас `MPSegmentation`:

```
class MPSegmentation:
    def __init__(
        self,
        bg_blur_ratio: typing.Tuple[int, int] = (35, 35),
        bg_image: typing.Optional[np.ndarray] = None,
        threshold: float = 0.5,
        model_selection: bool = 1,
        bg_images_path: str = None,
        bg_color: typing.Tuple[int, int, int] = None,
    ) -> None:
        self.mp_selfie_segmentation = mp.solutions.selfie_segmentation
        self.selfie_segmentation =
self.mp_selfie_segmentation.SelfieSegmentation(model_selection=model_sele
ction)
        self.bg_blur_ratio = bg_blur_ratio
        self.bg_image = bg_image
        self.threshold = threshold
        self.bg_color = bg_color
        if bg_images_path:
            self.bg_images = [cv2.imread(image.path) for image in
stow.ls(bg_images_path)]
            self.bg_image = self.bg_images[0]
```

Повний лістинг коду наданий у додатку А, однак наведений вище фрагмент дає чітке уявлення про те, як розмиття фону реалізовано у програмі.

Оптимізація продуктивності:

- розмиття умовно застосовується лише до тла, визначеного маскою сегментації. Таким чином зменшується кількість непотрібних обчислень;
- розмір ядра для гаусового розмиття можна регулювати, що дозволяє знайти компроміс між інтенсивністю ефекту розмиття та обчислювальним навантаженням.

Інтеграція з обробкою в реальному часі:

- ефект розмиття оптимізовано таким чином, щоб він не впливав суттєво на частоту кадрів в секунду (FPS) відеопотоку;
- застосунок дозволяє динамічно налаштовувати інтенсивність розмиття під час виконання, пристосовуючись до різних уподобань користувача та умов навколишнього середовища.

3.4 Виконання в реальному часі та аналіз ефективності

Основна мета виконання в реальному часі – забезпечити безперебійну та інтерактивну роботу процесу розмиття фону під час трансляції відео в реальному часі. Це вимагає від системи миттєвої обробки відеокадрів зі збереженням високої точності розпізнавання людей та ефективного розмиття фону.

Продуктивність системи в режимі реального часу оцінюється шляхом вимірювання частоти кадрів в секунду (FPS). Вищий показник FPS свідчить про більш плавну обробку відео. Затримка, тобто затримка між захопленням і відображенням відеокадру, ретельно відстежується. Низька затримка має вирішальне значення для взаємодії в реальному часі.

Щоб забезпечити надійність, система тестується в умовах різного освітлення – від слабкого до яскравого – щоб оцінити ефективність сегментації

людей і чіткість ефекту розмиття. Тести проводяться на фонах від однотонних до складних і динамічних сцен, щоб оцінити точність сегментації і якість розмиття. Оцінюється здатність системи підтримувати стабільну якість сегментації та розмиття для об'єктів, що рухаються, щоб забезпечити стабільність у різних сценаріях руху.

Аналіз продуктивності передбачає кількісну оцінку ефективності та результативності системи в тестових умовах. Основні показники включають:

- точність сегментації, що вимірюється здатністю системи правильно розрізняти об'єкт і фон у різних сценаріях;
- якість розмиття, що оцінюється природністю та рівномірністю ефекту розмиття, що застосовується до фону;
- FPS і затримка, що вимірюються в кожній тестовій ситуації для оцінки здатності системи обробляти дані в реальному часі;
- стійкість в умовах різного освітлення: система продемонструвала ефективну роботу в різних умовах, хоча незначні коливання точності сегментації спостерігаються в умовах екстремально низької освітленості;
- обробка складних фонів ефективно виділяє об'єкт на складному фоні з високим рівнем точності в більшості сценаріїв;
- система надійно відстежує рухомі об'єкти, зберігаючи точну сегментацію і послідовне розмиття фону без значного падіння частоти кадрів в секунду.

На основі аналізу продуктивності застосовуються кілька способів для підвищення ефективності системи:

- коригування порогів сегментації та коефіцієнтів розмиття для забезпечення балансу між точністю та швидкістю обробки;
- використання прискорення графічного процесора, де це можливо, для покращення FPS та зменшення затримок.

ВИСНОВКИ

У рамках кваліфікаційної роботи було розроблено та впроваджено систему розмиття фону у відеопотоці у реальному часі за допомогою нейромережових технологій. Результати роботи не тільки демонструють практичну застосовність передових методів комп'ютерного зору, але й створюють основу для подальшого вдосконалення та ширшого застосування цієї технології.

Розроблена система успішно реалізує розмиття фону в реальному часі, демонструючи практичне застосування нейронних мереж в обробці відео. Це досягнення має значну цінність для покращення користувацького досвіду на різних платформах відеозв'язку в реальному часі.

Здатність ефективно розмивати фон у відеопотоках в режимі реального часу вирішує найважливіші проблеми конфіденційності та підвищує естетичну якість цифрової комунікації, особливо в професійних умовах, таких як віддалена робота та онлайн-освіта.

Система готова до інтеграції в існуючі платформи відеоконференцій і потокового мовлення, пропонуючи значне поліпшення конфіденційності користувачів і якості відео. Завдяки своїй адаптивності вона підходить для різних умов, від низької освітленості до динамічного фону.

Хоча конкретні деталі організаційних впроваджень не доступні в рамках цього проекту, готовність технології передбачає її потенціал для прийняття різними організаціями, що займаються цифровими комунікаціями та рішеннями для обробки відео.

Подальші дослідження і розробки можуть бути зосереджені на оптимізації алгоритмів сегментації для ще більш складних сценаріїв, таких як надзвичайно динамічний фон або низькоконтрастне середовище. Вивчення більш досконалих методів апаратного прискорення було б корисним для підвищення продуктивності системи, особливо з точки зору зменшення затримок і збільшення FPS. Додавання більшої кількості функцій, керованих

користувачем, таких як регульовані рівні розмиття та чутливість сегментації, може зробити систему більш універсальною та зручною для користувача.

Таким чином, ця робота є черговим кроком у застосуванні нейронних мереж і методів комп'ютерного зору для обробки відео в реальному часі. Практична цінність розробленої системи в поєднанні з її потенціалом для подальшого вдосконалення підкреслює її актуальність і корисність в сучасному все більш цифровому і візуально керованому комунікаційному ландшафті. Впровадження такої технології не тільки покращує користувацький досвід і конфіденційність, але й відкриває нові шляхи для досліджень і розробок у галузі цифрової обробки відео.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
2. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NIPS)*.
3. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.
4. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection.
5. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234-241.
6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.
7. Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.
8. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
9. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
10. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets,

Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

11. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature Pyramid Networks for Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

12. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

13. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

14. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International Conference on Machine Learning (ICML)*.

15. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

16. Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

17. Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations (ICLR)*.

18. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., & Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems (NeurIPS)*.

19. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., & Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine

Learning. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI).

20. MediaPipe. (2020). MediaPipe Solutions.

21. Thoma, M. (2016). A Survey of Semantic Segmentation.

22. Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene Parsing through ADE20K Dataset. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

23. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

24. Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) Challenge. International Journal of Computer Vision.

25. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

26. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. European Conference on Computer Vision (ECCV).

27. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision.

28. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

29. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems (NIPS).

30. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., Hudáková M., and Gorokhovatskyi O. (2024) Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set, *IEEE Access*, vol. 12, pp. 73376-73385.

31. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features, *Indonesian Journal of Electrical Engineering and Computer Science*, 33 (1), 113-125.

32. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., and Al-Dhaifallah M. (2023) Statistical data analysis models for determining the relevance of structural image descriptions, *IEEE Access*, 11, 126938-126949.

33. Gorokhovatsky, V. (2014), *Structural Analysis and Intellectual Data Processing in Computer Vision*, SMIT, Kharkiv.

34. Gadetska, S.V., Gorokhovatsky, V.O. (2018) Statistical Measures for Computation of the Image Relevance of Visual Objects in the Structural Image Classification Methods. *Telecommunications and Radio Engineering*, Vol. 77 (12), pp. 1041–1053.

35. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Intelligent classification of biophysical system states using fuzzy interval logic, *Telecommunications and Radio Engineering*, 78(14), pp. 1303–1315.

36. Gorokhovatsky V.A. (2016) Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions. *Telecommunications and Radio Engineering*, Vol. 75, No 14. – P. 1271–1283.

37. Tvoroshenko I., Gorokhovatskyi V., Kobylin O., and Tvoroshenko A. (2023) Application of deep learning methods for recognizing and classifying culinary dishes in images, *International Journal of Academic and Applied Research*, 7(9), pp. 57-70.

38. Gorokhovatskyi V., Tvoroshenko I. (2023) Identification of visual objects by the search request. *Int. scientific symp. «Intelligent Solutions-S»*. Computational

intelligence. Decision making theory: proceedings of the international symposium, September 28, 2023, Kyiv-Uzhorod, Ukraine, 25-27.

39. Pomazan V., Tvoroshenko I., and Gorokhovatskyi V. (2023) Handwritten character recognition models based on convolutional neural networks, *International Journal of Academic Engineering Research*, 7(9), pp. 64-72.

40. Gadetska, S.V., Gorokhovatskyi, V. O., Stiahlyk, N. I., Vlasenko, N.V. (2021) Statistical data analysis tools in image classification methods based on the description as a set of binary descriptors of key points. *Radio Electronics, Computer Science, Control*, №4, pp. 58-68.

41. Gorokhovatskyi, V., Vlasenko, N. (2021). Редукція опису зображення у складі множини дескрипторів на основі метричного критерію інформативності. *Advanced Information Systems*, 5(4), pp. 10-16.

42. Гороховатський В., Творошенко І., Сидоренко Д. (2021) Класифікація зображень із використанням кластерного подання, Міжн. наук. симпозиум «Інтелектуальні рішення-С». *Обчислювальний інтелект. Теорія прийняття рішень* (Вересень 29, 2021). Київ – Ужгород, С. 44-45.

43. Gorokhovatskyi, O., Peredrii, O., Gorokhovatskyi, V., Vlasenko, N. (2023) Explanation of CNN Image Classifiers with Hiding Parts. In: J. Benoit-Pineau, R. Bourqui, D. Petkovic, G. Quenot (eds), *Explainable Deep Learning Artificial Intelligence*, pp. 125-146, Academic Press, 346 p.

44. Tvoroshenko, I., & Zarivchatskyi, R. (2020). Analysis of existing methods for searching object in the video stream, in *Proc. VI Int. Sci. Practic. Conf. «About the problems of science and practice, tasks and ways to solve them»*, Milan, pp. 500-505.

45. Gorokhovatskyi V.A., Zamula A.A. (2016) Employment of Intelligent Technologies in Multiparametric Control Systems. *Telecommunications and Radio Engineering*. Vol. 75, No 19, p. 1775–1785.

46. V. Gorokhovatsky, Y. Putyatin and V. Stolyarov (2017) Research of Effectiveness of Structural Image Classification Methods using Cluster Data Model, *Radio Electronics Computer Science Control*, vol. 3, no. 42, pp. 78-85.

47. Gorokhovatsky V.A., Putyatin Ye.P. (2009) Image Likelihood Measures of the Basis of the Set of Conformities. *Telecommunications and Radio Engineering*, 68 (9), pp. 763-778.

48. Gorokhovatsky, V.A., Putyatin, Y.P. (2008) Structural recognition of images on the basis of voting models of attributes of typical points, *Data recording, storage and processing*, 10(4), 75-85.

49. V. A. Gorokhovatskiy, (2011), Compression of descriptions in the structural image recognition, *Telecommunications and Radio Engineering*, vol. 70, no. 15, pp. 1363–1371, doi: 10.1615/TelecomRadEng.v70.i15.60.

50. Гороховатський В.О., Пупченко Д.В., Солодченко К.Г. (2018) Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення. *Системи управління, навігації та зв'язку*. С. 93–98.

51. Gadetska S., Gorokhovatskiy V., Stiahlyk N., Vlasenko N. (2022) Aggregate Parametric Representation of Image Structural Description in Statistical Classification Methods. In *CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2022)*, 3137, pp. 68-77.

52. Gorokhovatskiy V.A. (2018) Image Classification Methods in the Space of Descriptions in the Form of a Set of the Key Point Descriptors. *Telecommunications and Radio Engineering*, 77 (9), pp. 787-797.

53. Гороховатський В.О., Гадецька С.В., Стяглик Н.І., Власенко Н.В. (2020) Класифікація зображень на підставі ансамблю статистичних розподілів за класами еталонів для компонентів структурного опису. *Радіоелектроніка, інформатика, управління*, №4, с. 85–94.

54. Gorokhovatsky, V.O. and Gadetska, S.V. (2019) Determination of Relevance of Visual Object Images by Application of Statistical Analysis of Regarding Fragment Representation of their Descriptions, *Telecommunications and Radio Engineering*, 78 (3), pp. 211–220.