

ІНТЕЛЕКТУАЛЬНА СИСТЕМА АВТОМАТИЗОВАНОГО ДЕТЕКТУВАННЯ ПОЗИ ОПЕРАТОРА

Вінницький національний технічний університет;

Анотація

Досліджено поліпшення виявлення та розуміння поведінки людей автомобільними бортовими комп'ютерами за рахунок удосконалення методу розпізнавання пози людини у режимі реального часу.

Ключові слова: розпізнавання пози людини, згорткова нейронна мережа, ResNet, ефективність, точність, оцінка пози.

Abstract

The improvement of the detection and understanding of human behavior by car on-board computers has been investigated due to the improvement of the method of recognizing human posture in real time

Keywords: human pose recognition, convolutional neural network, ResNet, efficiency, accuracy, pose estimation.

Вступ

З кожним роком зростає кількість безпілотних технологій, технологій віртуальної реальності і доповненої реальності, які використовують інформацію про положення тіла людини у просторі. Можливість визначати позу людини на зображенні чи відео у зазначених областях відіграє ключову роль. На сьогоднішній день, не існує готового універсального рішення таких задач, незважаючи на те, що багато великих компаній ІТ індустрії, в тому числі Google, Facebook, NVidia, активно ведуть фундаментальні та прикладні дослідження в області машинного навчання. Значні успіхи були досягненні у розпізнаванні пози людини завдяки використанню нейронних мереж. Одна з найпоширеніших моделей - перцептрон. Однак для вирішення задачі розпізнавання пози людини дана модель не підходить, адже великий розмір вхідних даних (зображення, відео) призводить до значного збільшення кількості синаптичних зв'язків, нейронів у мережі. Як результат, швидкість та обчислювальна складність розпізнавання значно збільшується. Топологія вхідних даних ігнорується у цій моделі нейронної мережі, чітка двовірна структура вхідних зображень не враховується. Для усунення цих недоліків буде використовуватись згорткова нейронна мережа.

Результати дослідження

Для вирішення завдання з розпізнавання пози людини на зображеннях та відео потоках була обрана згорткова нейронна мережа ResNet (Residual Network). Архітектура даної мережі реалізує ідею передачі значень виходу і входу двох послідовно розташованих згорткових шарів для наступних шарів рисунок 1[1].

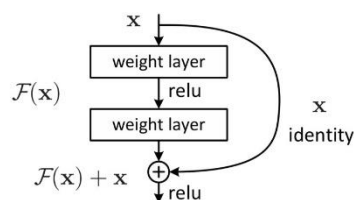


Рисунок 1 - Будівельний блок ResNet

Для кожної частини тіла або суглоба одна головна мережа передбачає показник достовірності, точне розташування та розмір цього суглоба, який називається картою інтенсивності частин (КІЧ). КІЧ [2] мають складну структуру. Вони складаються зі скалярного компонента для достовірності, векторного компонента, який вказує на найближчу частину тіла конкретного типу та іншого скалярного

компонента за розміром суглоба. Більш формально, у кожному вихідному місці (i, j) КІЧ прогнозує достовірність c , вектор (x, y) з розворотом b та шкалу σ і може бути записаний як:

$$p^{ij} = \{p_c^{ij}, p_x^{ij}, p_y^{ij}, p_b^{ij}, p_\sigma^{ij}\}$$

Коли мережа обробляє зображення, повертається теплова карта разом із векторами зсуву, які можна декодувати так, щоб знаходити в зображенні області з високим показником впевненості для ключових точок пози. Карта довіри MIF дуже груба. На рисунку 2(а) показана карта впевненості для лівого плеча на прикладі зображення. Щоб покращити локалізацію цієї карти довіри, з'єднано її з векторіальною частиною MIF, показаною на рисунку 2(б), у карту довіри з високою роздільною здатністю.

Створюється мапа довіри з високою роздільною здатністю $f(x, y)$ із згортковим ненормалізованим ядром Гаусса N шириною p_σ над регресуючими цілями з карти інтенсивності частин, зваженого рівнем:

$$f(x, y) = \sum_{ij} p_c^{ij} N(x, y | p_x^{ij}, p_y^{ij}, p_\sigma^{ij})$$

Це рівняння підкреслює безмержевий характер локалізації. Вивчається просторова міра σ суглоба, як частина поля. Приклад показаний на рисунку 2(в). Отримана карта локалізованих суглобів використовується для генерації поз та оцінки місця новостворених суглобів.

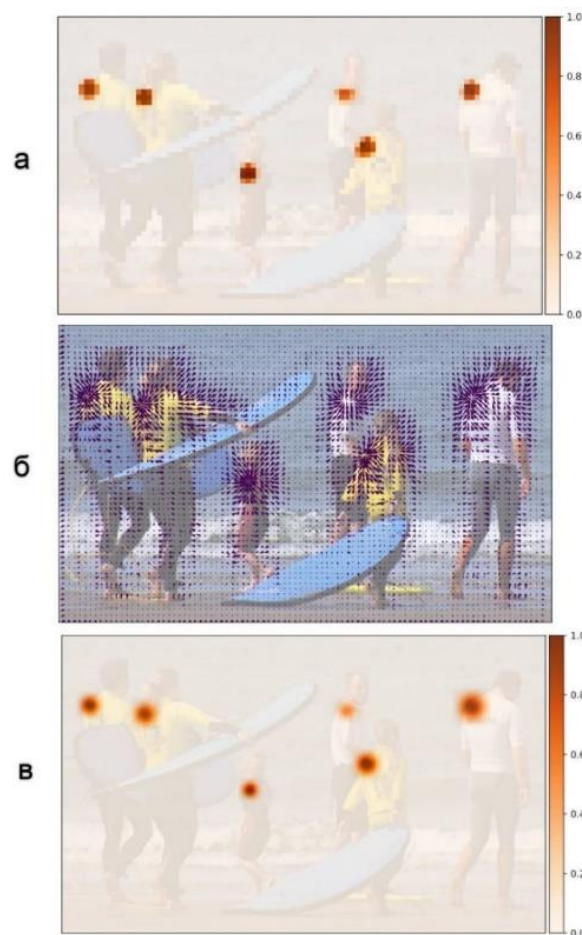


Рисунок 2 - Візуалізація компонентів КІЧ для лівого плеча. Це один із 17 складових КІЧ. Карта достовірності показана в (а), а векторне поле показано в (б). Посднання карти достовірності, компонентів вектору і масштабу показані у (в).

Висновки

Застосування згорткової нейронної мережі ResNet дозволяє ефективно виявляти позу людини на зображеннях та відео. Використання таких технологій дозволяє створювати теплові карти та векторні компоненти, що сприяють точній локалізації ключових точок пози. Результатом є можливість генерації поз та оцінки місця суглобів у реальному часі, що може мати значимість у сферах віртуальної реальності, медицини та спорту.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition
Режим доступу: <https://arxiv.org/abs/1512.03385>
2. L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. V. Gehler, and B. Schiele. Deepcut: Joint subset partition and labeling for multi person pose estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4929–4937, 2016

Черната Дмитро Вадимович — студент групи 2АКІТ-22м, Факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: chetn1944@gmail.com

Науковий керівник: **Гришук Тетяна Вікторівна** — к.т.н., доцент кафедри комп'ютерних систем управління, Вінницький національний технічний університет, м. Вінниця

Chernata Dmytro Vadymovych — student of group 2AKIT-22m, Faculty of intelligent information technologies and automation, Vinnytsia National Technical University, Vinnytsia, e-mail: chetn1944@gmail.com

Supervisor: **Gryshchuk Tetiana Viktorivna** — Cand. Sc., Associate Professor of the Department of Computer Control Systems, Vinnytsia National Technical University, Vinnytsia