

ОГЛЯД ПІДХОДІВ ДО ПІДРАХУНКУ ВІДВІДУВАЧІВ У МАГАЗИНАХ РОЗДРІБНОЇ ТОРГІВЛІ

Вінницький Національний Технічний Університет

Анотація: У цій статті була описана технологія підрахунку людей, що використовує ШІ, яка може допомогти супермаркетам та іншій роздрібній торгівлі боротися з конкуренцією з боку онлайн-конкурентів.

Ключові слова: штучний інтелект, датчики, лічильники, відеодатчики, глибоке навчання, нейронні мережі, CNN, CSRNet, computer vision.

Abstract: This article described the technology of counting people using AI, which can help supermarkets and other retailers fight competition from online competitors.

Keywords: artificial intelligence, sensors, counters, video sensors, deep learning, neural networks, CNN, CSRNet, computer vision.

Вступ

Цифрова трансформація роздрібної індустрії триває вже кілька років. Завдяки вдосконаленню технологій роздрібні магазини можуть збільшити прибуток і підвищити рівень задоволеності своїх клієнтів. Одним із таких методів, які використовуються для покращення галузі, є застосування технології підрахунку людей. Системи підрахунку людей використовують камеру та програмне забезпечення для підрахунку кількості людей, які входять до магазину. Сам пристрій може представляти собою кілька підсистем, включаючи інфрачервоні датчики, ручні клікери або тепловізійні системи.

Термолічильник та інфрачервоний лічильник

Теплові лічильники використовують тепло тіла людини для вимірювання кількості відвідувачів. Користуючись алгоритмами глибокого навчання створюють зображення за допомогою інфрачервоного випромінювання так само, як типова камера формує зображення, використовуючи видиме світло. На входах встановлені термодатчики, які визначають температуру тіла людини та підраховують її.

Інфрачервоні датчики ToF схожі на теплові лічильники за конструкцією.

Ці датчики працюють за допомогою алгоритму глибокого навчання, надсилаючи сигнал об'єктам під датчиком, потім записують час, необхідний для того, щоб сигнал досяг свого походження, і, нарешті, реєструють відлік від різниці в часі.

Відеолічильник

Датчики цього виду є найбільш точними, тому що використовують методи комп'ютерного зору для підрахунку натовпу. Раніше у деяких ситуаціях проблемою для таких лічильників була анонімність, але сучасні датчики нового покоління використовують перекриття обличчя людей для анонімності даних, якщо це необхідно.

На початковому етапі дослідження дослідники нанесли позначки на голову кожної людини, щоб тренувати ШІ шляхом повторення. Після цього рівень точності був покращений завдяки глибокому навчанню. Цей навчальний процес передбачав використання кількох сотень тисяч зразків зображень натовпу. Оскільки

отримати таку величезну кількість реальних зображень було б неможливо, було створено і використано комп'ютерне 3D-модель натовпу, щоб допомогти збільшити обсяг даних для навчання.

Щоб ще більше покращити зручність використання, у таких датчиках розроблено легку модель глибокого навчання, яка дозволяє виконувати обробку за допомогою лише ЦП, не покладаючись на графічний процесор. Це сприяє зниженню експлуатаційних витрат і енергоспоживання, що є однією з найбільших переваг технології.

Методи комп'ютерного зору для підрахунку натовпу

Наразі існує чотири методи, які використовуються для підрахунку кількості людей у натовпі:

1. Методи на основі виявлення
2. Регресійні методи
3. Методи оцінки щільності
4. Методи на основі CNN

Архітектура методу навчання CSRNet (базовано на CNN)

CSRNet використовує VGG-16 як зовнішній інтерфейс через його здатність до навчання. Вихідний розмір VGG становить $\frac{1}{8}$ -ю від початкового розміру вхідного сигналу. CSRNet також використовує dilated convolutions на серверній частині.

Основна концепція використання dilated convolutions полягає в тому, щоб збільшити ядро без збільшення параметрів.

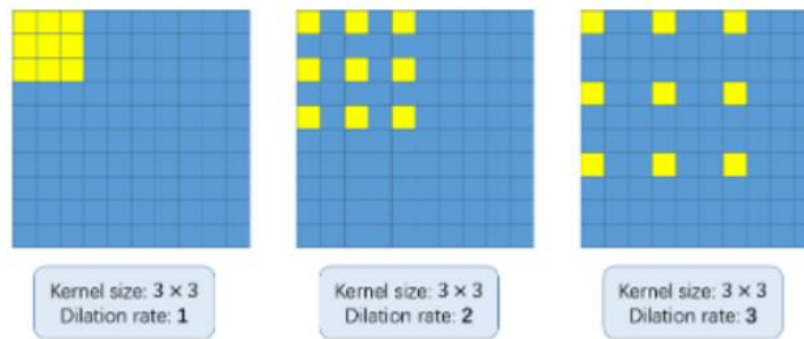


Рис. 1 – Концепція dilated convolutions

Математична модель CSRNet

Припустимо, що є input $x(m,n)$, filter $w(i,j)$ і швидкість розширення r . Результатом $y(m,n)$ буде:

$$y(m, n) = \sum_{i=1}^M \sum_{j=1}^N x(m + r \times i, n + r \times j) w(i, j)$$

Узагальнення даного рівняння, використовуючи ядро ($k*k$) зі швидкістю розширення r . Ядро збільшується до:

$$([k + (k-1)*(r-1)] * [k + (k-1)*(r-1)])$$

Таким чином, для кожного зображення було встановлене істинне положення об'єкту (ground truth). Голова кожної людини на заданому зображенні розмита за допомогою ядра Гауса. Усі зображення обрізано на 9 фрагментів, і розмір кожного фрагмента становить $\frac{1}{4}$ від початкового розміру зображення.

Перші 4 фрагменти розділені на 4 частини, а інші 5 фрагментів обрізаються випадковим чином.

Стохастичний градієнтний спуск використовується для навчання CSRNet як наскрізної структури. Під час навчання фіксована швидкість навчання встановлюється на $1e-6$. Функція втрат приймається як евклідова відстань, щоб виміряти різницю між основною істиною та оціненою картою щільності. Це представлено як:

$$L(\Theta) = \frac{1}{2N} \sum_{i=1}^N \| Z(X_i; \Theta) - Z_i^{GT} \|_2^2$$

де N — розмір навчальної групи. Метрикою оцінки, що використовується в CSRNet, є MAE і MSE, тобто середня абсолютна помилка і середня квадратична помилка. Їх дають:

$$MAE = \frac{1}{N} \sum_{i=1}^N | C_i - C_i^{GT} |$$
$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^N | C_i - C_i^{GT} |^2}$$

Дана модель спочатку передбачить карту щільності для даного зображення. Значення пікселя буде 0, якщо немає людей. Попередньо визначене значення буде призначено, якщо цей піксель відповідає людині. Отже, обчислення загальних значень пікселів, що відповідають людині, на виході дасть кількість людей на цьому зображенні.

Висновок

У даній статті оглянуто використання лічильників людей у індустрії роздрібною торгівлі, їх доцільність та види. Описано їх види для наочності функціоналу кожного із них. Було описано методи реалізації алгоритмів глибокого навчання для відеодатчиків. Також було приведено приклади алгоритмів ШІ, які можуть бути використані для роботи відеодатчиків.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Crowd Counting. Analytics Vidhya [Електронний ресурс] : [Веб-сайт]. – Режим доступу: <https://www.analyticsvidhya.com/blog/2019/02/building-crowd-counting-model-python/> – Назва з екрана
2. Infrared Sensor – How it Works, Types, Applications, Advantage & Disadvantages. [Електронний ресурс] : [Веб-сайт]. – Режим доступу: <https://electricalfundablog.com/infrared-sensor/> – Назва з екрана
3. People Counter. Canon [Електронний ресурс] : [Веб-сайт]. – Режим доступу <https://asia.canon/en/business/people-counter/product> – Назва з екрана
4. Thermal Counter. EvolvePlus [Електронний ресурс] : [Веб-сайт]. – Режим доступу <https://www.evolveplus.com.au/solutions/people-counters/overhead-thermal-counters> – Назва з екрана

Барановська Анастасія Юріївна – студентка групи ІІСТ-19Б, кафедра автоматизації та інтелектуальних інформаційних технологій, Факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, м.Вінниця, e-mail: 01-19-051.stud@vntu.edu.ua

Маслій Роман Васильович – к.т.н., доцент кафедри Автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м.Вінниця, e-mail: maslij.r.v@vntu.edu.ua

Baranovska Anastasiia Yuriivna – student of IIST-19B group, Department of Automatization and Intellectual Informational Technologies, Faculty of Computer Systems and Automatics, Vinnytsia National Technical University, Vinnytsia, e-mail: 01-19-051.stud@vntu.edu.ua.

Maslij Roman Vasylovych – Associate Professor of Automation and Intelligent Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: maslij.r.v@vntu.edu.ua