

УДК 004.9:621.397.43

**РОЗПІЗНАВАННЯ КРАДІЖОК У МАГАЗИНАХ НА ЗАПИСАХ
КАМЕР ВІДЕОСПОСТЕРЕЖЕННЯ ЗА ДОПОМОГОЮ
КОМБІНОВАНОЇ МОДЕЛІ CNN-RNN**

Сидоренко Б.Ю.

Науковий керівник – д-р техн. наук, проф. Кіріченко Л.О.
Харківський національний університет радіоелектроніки, каф. ПМ
м. Харків, Україна

тел. +38 (057) 702-14-36, email: bohdan.sydorenko1@nure.ua

The recognition of human activities through surveillance has numerous applications across various fields. This article presents a proposed approach to identify shoplifting in camera-recorded video data using a neural classifier that combines two neural networks, specifically, convolutional and recurrent networks. The hybrid architecture consists of two parallel streams: initial and processed video fragments (histogram of oriented gradients and optical flow). The convolutional network extracts features from each frame of the video fragment, while the recurrent network processes the temporal information from sequences of frames as features to classify the activity.

Розпізнавання дій людини є важливим завданням сучасного відеоспостереження. З роками кількість відеокамер у громадських місцях зростає, тому це ускладнює завдання відеоспостереження. Мережі камер відеоспостереження генерують і передають величезні обсяги даних, що робить автоматичну обробку всієї інформації надзвичайно важливою.

Обробка відеоспостереження є важливим інструментом для виявлення крадіжок у магазинах, оскільки автоматичний аналіз великих обсягів відєоданих може забезпечити виявлення незаконних дій у режимі реального часу. Для цього використовуються різні алгоритми машинного навчання. Такі моделі навчаються на всеохоплюючих наборах даних про крадіжки в магазинах, що дозволяє їм ідентифікувати моделі злодіїв і класифікувати їхні дії на основі певних ознак.

Згорткові нейронні мережі є потужним інструментом для класифікації зображень і значно вдосконалили обробку відео за останні роки. Однак слід зазначити, що класифікація відео вимагає врахування як просторових, так і часових характеристик об'єктів.

У нашому дослідженні використовується набір даних UCF-Crime [1] як вхідні дані для наших експериментів. Набір даних включає 1900 відео різної тривалості, що становить 128 годин реальних кримінальних дій, таких як насильство, арешт, підпал, напад, дорожньо-транспортні пригоди, крадіжка зі зломом, вибух, бійка, пограбування, стрілянина, крадіжка в магазині та вандалізм. Зокрема, набір даних включає 28 відеозаписів з роздрібного магазину та камер відеоспостереження, які містять крадіжки в магазині.

Для навчання нашої нейронної мережі ми штучно збільшили кількість

екземплярів, розділивши кожне відео на 32 фрагменти тривалістю 3 секунди. Отриманий набір даних із 896 фрагментів відео був розділений на два класи: 155 відео з крадіжками в магазинах і 741 відео без крадіжок у магазинах. Оскільки відеозапис містить інформацію про час і простір, при аналізі відеофрагментів необхідно враховувати обидва види інформації. Ми вважаємо, що згорткові та рекурентні нейронні мережі є найкращими архітектурами для виконання цього завдання. Тому ми вибрали комбінацію цих мереж для класифікації відео.

В якості вхідних даних використовувалися маркіровані відеокліпи однакової тривалості. Кожна послідовність кадрів була позначена як «0» (не крадіжка) або «1» (крадіжка). Позначений набір кадрових послідовностей використовувався як навчальна вибірка. Характеристики були отримані для кожного об'єкта за допомогою гібридної нейронної мережі для навчання класифікатора, який потім використовувався для класифікації нових об'єктів. Для підвищення якості моделі ми вирішили застосувати попередню обробку фрагментів відео. А саме, ми використовували комбінацію гістограм орієнтованих градієнтів і оптичних потоків. Швидкий і багатообіцяючий спосіб представлення зображень для класифікації – це використання функцій HOG. Ці функції широко використовувалися для ідентифікації пішоходів і залишаються надійною технікою виділення ознак [2]. Позірний рух об'єктів у візуальній сцені, викликаний рухом камери, об'єкта або обох, можна описати як оптичний потік [3].

Це дослідження мало на меті розробити класифікатор для ідентифікації випадків крадіжок у магазинах у відеоданих з камер спостереження. Для досягнення поставленої мети було запропоновано гібридний класифікатор нейронних мереж, що включає згорткові та рекурентні нейронні мережі. Класифікатор продемонстрував високу точність класифікації 94%, що на кілька відсотків вище, ніж точність моделей, представлених у попередніх відповідних дослідженнях. Крім того, наш навчений класифікатор демонструє високу продуктивність, що дозволяє використовувати його в програмах реального часу. Подальші дослідження будуть зосереджені на практичній реалізації запропонованої моделі в торгових центрах.

Список використаних джерел:

1. *Papers with code – UCF-crime dataset*. URL <https://paperswithcode.com/dataset/ucf-crime>.

2. Torrione, P. A., Morton, K. D., Sakaguchi, R., & Collins, L. M. *Histograms of oriented gradients for landmine detection in ground-penetrating radar data*. 52 (3):1539–1550. ISSN 1558-0644. doi:10.1109/TGRS.2013.2252016.

3. Shah, S. T. H. & Xuezhai, X. *Traditional and modern strategies for optical flow: an investigation*. 3 (3) : 289. doi:10.1007/s42452-021-04227-x