

ПРОБЛЕМНІ АСПЕКТИ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Вінницький національний технічний університет

Анотація

В ході проведеного дослідження проаналізовані проблемні аспекти класифікації зображень на основі згорткових нейронних мереж, а саме витік даних та невідповідність зон зображення, на які реагує мережа, зонам що містять важливу інформацію. Розглянуто підходи до візуалізації та інтерпретації результатів роботи згорткових нейронних мереж.

Ключові слова: згорткові нейронні мережі, візуалізація, карта виразності, LIME

Abstract

In the given research, the problem aspects of the convolutional neural networks image classification, namely the data stream and the discrepancy of the image zones responding to the network, to the zones containing the important information were analyzed. Approaches to visualization and interpretation of the results of convolutional neural networks work are considered.

Keywords: convolutional neural networks, visualization, saliency map, LIME

Вступ

Зважаючи на огляд публікацій з даної тематики, для задачі класифікації зображень Згорткові Нейронні Мережі (ЗНМ) вважаються перспективним і сучасним підходом [1]. Але результати роботи ЗНМ важко інтерпретувати та впевнитись, що мережа реагує саме на ті зони зображення, що містять необхідну інформацію для класифікації. Можливі проблемні аспекти розглянуто на прикладі задачі класифікації вантажних автомобілів на зерновому терміналі за допомогою ЗНМ.

Візуалізація

Натренована ЗНМ коректно класифікує зображення з високою точністю. Але при тренуванні мережі міг статись витік даних, наприклад десь на зображенні міг бути індикатор або ознака що однозначно вказує на певний клас. Для того, щоб перевірити коректність класифікації важливо оцінити зони зображення, які для натренованої мережі мають найбільшу вагу. Для візуалізації розглядається два підходи (Рисунок 1):

- Saliency Map - побудова карти виразності, шляхом аналізу активацій [2];
- LIME - техніка, розбиття зображення на зони та послідовною обробкою за допомогою ЗНМ [3].

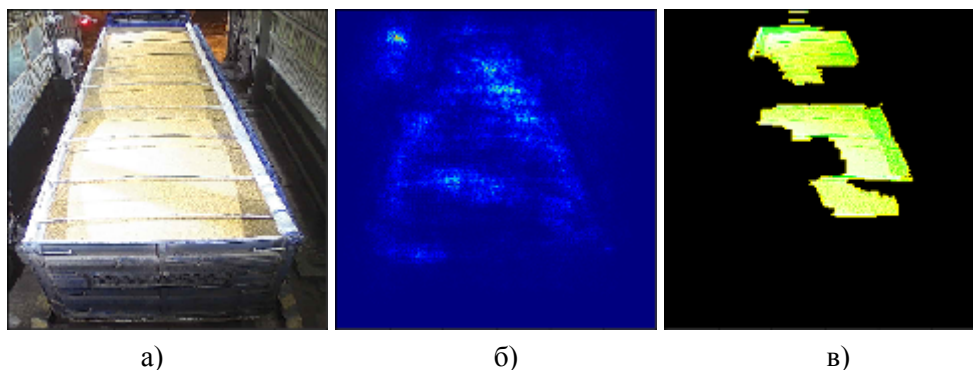


Рисунок 1 - Аналіз зон, на які реагує ЗНМ: а) Оригінал; б) Saliency Map; в) LIME.

На Saliency Map (Рисунок 1б) - чітко видно контури кузова та підсвічені зони, що відповідають за вміст. За допомогою техніки LIME обрано зони, що відповідають за вміст кузова (Рисунок 1в). Загалом видно, що ЗНМ першочергово реагує на вміст кузова, що і є очікуваним результатом.

Візуалізація на наборі даних з “виразними” ознаками

Для того щоб впевнитись що візуалізація працює саме з зонами, на які реагує ЗНМ, було введено так звану “виразну” ознаку, підказку для нейромережі, що буде імітувати витік даних. Для кожного класу зображення обрано колір та на зображення додано прямокутник обраного кольору (Рисунок 2).

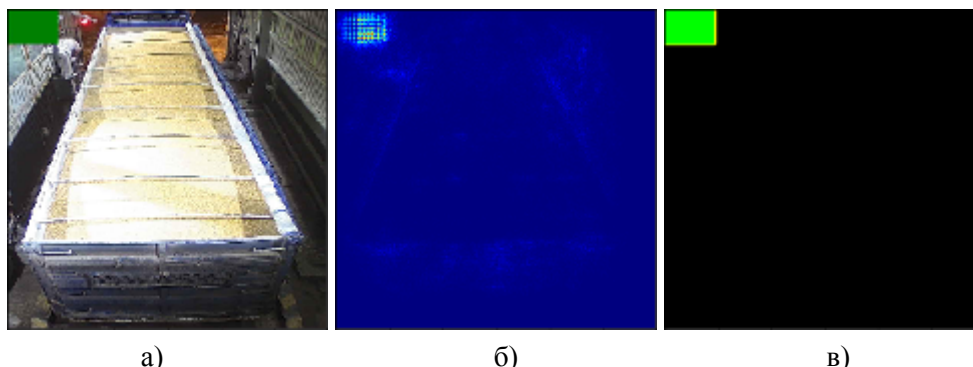


Рисунок 2 - Аналіз зон, на які реагує нейромережа на зображенні з “виразними” ознаками:
а) Оригінал; б) Saliency Map; в) LIME.

На Saliency Map (Рисунок 2б) - видно нечіткі контури кузова та яскраво підсвічена зона прямокутника. За допомогою техніки LIME обрана лише зона прямокутника (Рисунок 2в). Загалом видно, що ЗНМ реагує лише на прямокутник в кутку зображення, який імітує витік даних, при цьому повністю ігнорує вміст кузова, що не є коректним.

Висновки

Проблема витоку даних є важливою проблемою, яку важко діагностувати. Saliency Map і LIME важливі інструменти для аналізу коректності роботи ЗНМ. Вони дають можливість оцінити зони, на які мережа реагує в першу чергу і зони, які вона повністю ігнорує, та проаналізувати їх.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Kozemiako, V. P., Kolesnytskyj, O. K., Lischenko, T. S., Wojcik, W., & Sulemenov, A. (2013). Optoelectronic spiking neural network. Paper presented at the Proceedings of SPIE - the International Society for Optical Engineering, 8698 doi:10.1117/12.2019340 Retrieved from www.scopus.com
2. Karen Simonyan, Andrea Vedaldi, Andrew Zisserman. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. – Cornell University Library, 2014 – [Електронний ресурс] – Режим доступу: <https://arxiv.org/abs/1312.6034>
3. Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin. “Why Should I Trust You?” Explaining the Predictions of Any Classifier. – Cornell University Library, 2016 – [Електронний ресурс] – Режим доступу: <https://arxiv.org/abs/1602.04938>

Шемет Євген Олександрович - асистент кафедри Комп'ютерних Наук, Вінницький Національний Технічний Університет, Вінниця, yevhene@vntu.edu.ua

Шемет Тетяна Олегівна - Вінниця, tanyshak@gmail.com

Яровий Андрій Анатолійович - доктор технічних наук, професор, професор кафедри Комп'ютерних Наук, Вінницький Національний Технічний Університет, Вінниця

Shemet Yevhen O. - Department of the Computer Science, Vinnytsia National Technical University, Vinnytsia, yevhene@vntu.edu.ua

Shemet Tetiana O. - Vinnytsia, tanyshak@gmail.com

Yarovyuy Andriy A. - Dr. Sc. (Eng.), Professor, Head of the Computer Science, Vinnytsia National Technical University, Vinnytsia