

КЛАСИФІКАЦІЯ ОБЛИЧ В МАСКАХ З ВИКОРИСТАННЯМ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Вінницький національний технічний університет

Анотація

В ході проведеного дослідження проаналізовані особливості класифікації облич в масках на основі згорткових нейронних мереж. Розроблено архітектуру мережі, що дозволяє витримати високу якість розпізнавання та задовільну швидкість обробки.

Ключові слова: згорткові нейронні мережі

Abstract

In the given research the peculiarities of face classification in masks on the basis of convolutional neural networks are analyzed. The network architecture has been developed, which allows to maintain high recognition quality and satisfactory processing speed.

Keywords: convolutional neural networks

Вступ

Тренування та порівняння мереж проводилось за набором даних в якому містяться зображення людей в масках і без масок, з ресурсу kaggle.com [1]. Після попередньої обробки зображення були приведені до одного розміру та вирізана приблизно однакова частини голови, щоб спростити подальшу роботу мережі. Результативний набір даних складався з 3788 фотографій.



Рисунок 1 - Приклади зображень після обробки з мітками “good” та “bad”.

Тренування

В ході дослідження для обробки даних застосовувалося три різні підходи:

а) звичайна нейронна мережа;
б) мережа Xception підготовлена за допомогою методу “transfer learning” [2] на наборі даних “imagenet” [3] (навчання для згорткової частини Xception заблоковане, щоб прискорити час навчання, натомість замість верхівки мережі додано новий шар нейронів);

в) згорткова нейронна мережа [4] спроектована спеціально до цієї задачі [5].

У всіх трьох мереж приблизно однакова кількість тренуваних параметрів. Функція втрат - бінарна перехресна ентропія [6].

Результати

Звичайна нейронна мережа з завданням не впоралась, показавши невисокі результати (див. табл. 1). Точність мережі Xception нижча за 0.9, тобто вона помиляється в одному випадку з 10, що також не є прийнятними показниками. Показники Xception значно поліпшувались

відповідно до кількості тренуваних параметрів але час тренування збільшувався в десятки раз. Також вона займає значно більше місця і вимагає значно більшого часу на обробку даних після тренування. Спеціально спроектована під задачу, згортова нейронна мережа дала високі результати і найкращу точність серед трьох.

Таблиця 1 - результати тренування

	Параметри (кількість змінних в мережі)		Кращий результат	
	Треновані	Загалом	Точність	Значення функції втрат
Звичайна НМ	196 641	196 641	0,82	2,76
Хсерption	131 201	20 992 681	0,88	1,23
Згортова НМ	176 929	176 929	0,97	0,08

Висновки

Основною складністю дослідження була обробка реальних фотографій. Серед фотографій були не тільки фото с людьми в маска і без масок, а також фото з частково надітими масками, частково закритим обличчям чи складними кутами повороту голови. Боротьбу з цими ускладненнями значно спростила аугментація зображень [7].

Підхід “transfer learning” дає гарні результати значно зменшуючи витрати часу на проектування мережі, та при зменшенні кількості тренуваних параметрів пришвидчуючи сам процес тренування. Тим не менш мережа спеціально підготовлена під задачу, значно перевершує результати популярних нейромережових архітектур.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Mikolaj Witkowski. Medical Masks Dataset (Pictures of people wearing medical masks) – [Електронний ресурс] – Режим доступу: <https://www.kaggle.com/vtech6/medical-masks-dataset>
2. Barret Zoph, Deniz Yuret, Jonathan May, Kevin Knight. Transfer Learning for Low-Resource Neural Machine Translation. – Cornell University Library, 2016 – [Електронний ресурс] – Режим доступу: <https://arxiv.org/pdf/1604.02201>
3. ImageNet – [Електронний ресурс] – Режим доступу: <http://www.image-net.org/>
4. 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
5. Yevhen Shemet. Medical Masks Dataset Crop faces and classify – [Електронний ресурс] – Режим доступу: <https://www.kaggle.com/yevhene/medical-masks-dataset-crop-faces-and-classify>
6. Yaoshiang Ho, Samuel Wookey. The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling. – Cornell University Library, 2020 – [Електронний ресурс] – Режим доступу: <https://arxiv.org/abs/2001.00570>
7. Hankook Lee, Sung Ju Hwang, Jinwoo Shin. Rethinking Data Augmentation: Self-Supervision and Self-Distillation. – Cornell University Library, 2019 – [Електронний ресурс] – Режим доступу: <https://arxiv.org/abs/1910.05872>

Шемет Євген Олександрович - асистент кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця, yevhene@vntu.edu.ua

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