

УДК 004.932

МЕТОДИКА ПСЕВДОМАРКУВАННЯ ДАНИХ ДЛЯ МАШИННОГО НАВЧАННЯ ГЛИБИННИХ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ПРИ РОЗПІЗНАВАННІ СЛАБКОМАРКОВАНИХ ЗОБРАЖЕНЬ ВИРАЗІВ ОБЛИЧЧЯ

Міхалев Кирило Богданович, Петросюк Денис Валерійович
к.т.н., доцент каф. ІС Ніколенко Анатолій Олександрович
Національний університет «Одеська політехніка», Україна

АНОТАЦІЯ. Розглянуто питання розробки та дослідження методики використання слабо маркованих даних для машинного навчання. Показана можливість автоматизувати процес збільшення навчальних вибірок шляхом псевдомаркування тобто додавання раніше немаркованих даних без участі людини.

Вступ. Зі зростанням обчислювальних можливостей сучасних комп'ютерних систем особливо популярними стають рішення щодо розпізнавання та класифікації об'єктів у системах комп'ютерного зору на основі глибокого навчання (Deep Learning, DL) згорткових нейронних мереж (Convolutional Neural Network, CNN). Однак, незважаючи на наявність позитивних прикладів, успішне застосування DL CNN залишається складною проблемою. Це пояснюється тим, що для реалізації технологічних рішень DL CNN з необхідною точністю, крім високовиробничих графічних процесорів (Graphics Processing Unit, GPU) необхідний досить великий набір попередньо розмічених даних. І якщо для вирішення класичних задач розпізнавання об'єктів існують набори даних типу ImageNet, що містять більше 10 мільйонів анотованих вручну зображень, то для вирішення інших актуальних завдань, наприклад, пов'язаних з розпізнаванням виразу особи (міміки) (Facial Expression Recognition, FER) та визначення емоцій (Emotion Detection, ED) такі набори відсутні. Моделювання завдань FER та ED виконують, як правило, на загальнодоступному наборі DISFA, який обмежений відео від 27 суб'єктів – 12 жінок та 15 чоловіків, з кожним із яких записано відео з 4845 кадрами. Що явно замало реалізації DL CNN. Тому актуальними є роботи пов'язані зі збільшенням даних у навчальних наборах, використанням наборів зі слабо розміченими даними, а також трансферним та напівконтрольованим навчанням.[1,2,3]

Метою роботи є розробка методики використання слабо маркованих даних (зображень обличчя людини) для DL CNN для збільшення точності вирішення задач FER и ED

Основна частина роботи.

Псевдомаркування – це процес додавання впевнених передбачуваних даних тесту до тренувальних даних.

Псевдомаркування належить до напівконтрольованого навчання(semi supervised learning). Напівконтрольоване навчання - це підхід до машинного навчання, який поєднує невеликий обсяг помічених даних з великим обсягом немаркованих даних під час навчання. Напівконтрольоване навчання знаходиться між неконтрольованим навчанням (без маркованих даних навчання) та контрольованим навчанням (тільки з позначеними даними навчання). Це особливий випадок слабого нагляду [3,4]

Псевдомаркування складається з 4 кроків.

1. Побудова моделі, використовуючи навчальні дані для даної предметної області.
2. Передбачення мітки для немаркованого набору тестових даних.
3. Додавання найбільш прогнозованих тестових спостережень до навчальних даних
4. Повторне навчання моделі, використовуючи комбіновані дані.

Розглянемо особливості псевдомаркування щодо рішення задачі FER.

На *першому кроці* будуємо модель згорткової нейронної мережі. У якості проміжних шарів були обрані Conv2D-шари в Keras. Цей шар подібний Dense-шару, і містить ваги і зміщення, які піддаються оптимізації (підбору). Conv2D-шар так само містить фільтри («ядра»), значення яких теж оптимізуються. Цей шар створює ядро згортки, яке спільно з входом шару творить тензор(2-вимірні масиви) вихідних даних. Емпіричним шляхом було обрано 3 Conv2D-шари з кількістю

фільтрів 64, 128 та 256, з функцією активації relu та з розмірами ядра (5,5) (5,5) та (3,3), після кожного з них йде шар MaxPooling2D розміром ядра (2,2) , також Dense шари, один з яких розміром фільтра(128) та один останній, що дорівнює кількості емоцій, тобто 7. Схема такої мережі зображена на рисунку 1.

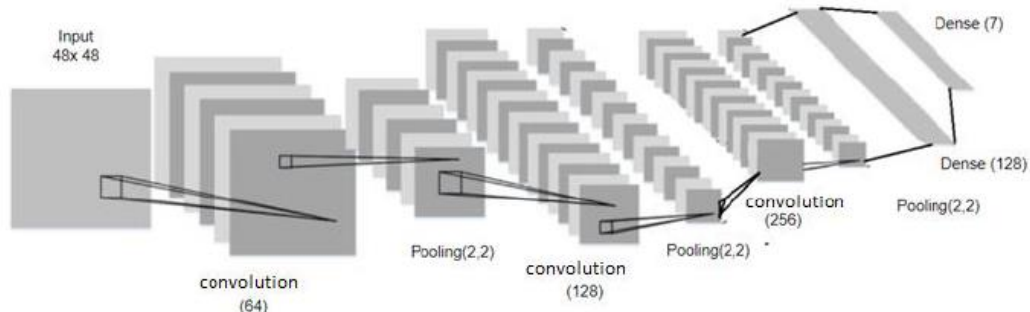


Рисунок 1 – Схема згорткової нейронної мережі

У якості маркованого набору даних для навчання використовується датасет з зображеннями 7 основних емоцій людини: гнів, огида, страх, щастя, сум, здивованість, нейтральність. Другий датасет містить вже немарковані зображення, тобто ми не маємо апріорної інформації про емоції людей на зображеннях із датасета. Перший датасет – це архів під назвою fer2013. Архів містить csv файл, в якому дані представлені у вигляді рядків з трьома стовпцями [5]. Розділяємо маркований набір даних на дві групи (тренувальну та тестову - еталонну вибірки у пропорції 80 на 20 відсотків відповідно). Навчаємо згорткову нейронну мережу на тренувальній вибірці та тестуємо на тестовій еталонній вибірці

На *другому кроці* відбувається передбачення мітки для немаркованого набору тестових даних з використанням згорткової нейронної мережі, навченої на першому кроці. Таким чином отримаємо псевдомаркований набір даних

На *третьому кроці* обираються дані із псевдомаркованого набору даних, в яких модель була найбільш впевнена - задається поріг точності класифікації (передбачення). Обрані таким чином данні додаються до початкової тренувальної вибірки тим самим збільшується її розмір

На *четвертому кроці* повторюється навчання згорткової нейронної мережі на оновленій тренувальній вибірці. Перевірка навченої згорткової нейронної мережі виконується з використанням даних тестової - еталонної вибірки.

Кроки з другого по четвертий спрямовані на збільшення обсягу маркованих даних та виконуються ітераційно доки не буде зменшуватися якість тестування за даними тестової - еталонної вибірки.

На рисунку 2, а,б,в показані характеристики точності розпізнавання емоцій в залежності від ітерації навчання за методикою псевдомаркування.

	precision	recall	f1-score		precision	recall	f1-score		precision	recall	f1-score
здивованість	0.71	0.79	0.75	здивованість	0.77	0.74	0.76	здивованість	0.76	0.72	0.74
страх	0.64	0.34	0.44	страх	0.69	0.39	0.50	страх	0.59	0.49	0.53
огида	0.51	0.29	0.37	огида	0.51	0.41	0.46	огида	0.35	0.51	0.41
щастя	0.91	0.86	0.88	щастя	0.86	0.92	0.89	щастя	0.91	0.87	0.89
сум	0.72	0.71	0.71	сум	0.76	0.69	0.72	сум	0.68	0.79	0.73
злість	0.64	0.70	0.67	злість	0.74	0.64	0.68	злість	0.69	0.66	0.67
нейтральність	0.66	0.78	0.72	нейтральність	0.71	0.78	0.75	нейтральність	0.73	0.66	0.70
точність	0.76			точність	0.78			точність	0.76		

а

б

в

Рисунок 2 – Характеристики точності розпізнавання емоцій за методикою псевдомаркування на першій (а), другій (б) та третій ітерації (в) навчання.

Як можна побачити після першої ітерації (рис. 2,а), найбільш точно нейромережева модель розпізнає емоцію щастя, злості, нейтральності та здивованості. А загальна точність моделі склала 76%. На другій ітерації (рис. 2,б) згідно з методикою тренувального набору даних було додано 363 зображень, в яких перша модель була впевнена на 99% та отримано наступні результати розпізнавання емоції здивованості (з 71 до 77 %), страху (з 64 до 69 %), суму(з 72 до 76 %), злості з (з 64 до 74 %), нейтральності (з 66 до 71%). Не змінилася точність у емоції огиди. При цьому точність розпізнавання емоції щастя, що є найбільш поширена в тренувальній вибірці знизилась (з 91 до 86 %). Тим самим, загальна точність моделі стала на 2 відсотка краще порівняно з попередньою і в підсумку ми маємо загальну точність 78%. Порівнюючи дані, які отримані на третій ітерації із попередньою, ми можемо побачити, що в цілому результати погіршилися. Тільки емоції щастя та нейтральності отримали приріст точності (з 86 до 91 % та з 71 до 73%) відповідно. На третій ітерації (рис. 2,в) загальна точність моделі впала на 2 % і тепер склала 76%. З цього можна зробити висновок, що використання запропонованої методики псевдомаркування, а саме додавання до тренувальної вибірки даних, в яких модель впевнена навіть більш ніж на 99 відсотків, не завжди веде до зростання загальної точності моделі та окремих класів, представлених у вибірці.

Висновки. В роботі розглянуто проблему дослідження та розробки методики використання слабо маркованих даних для машинного навчання. Цей підхід дозволяє автоматизувати процес збільшення навчальних вибірок шляхом додавання до них раніше немаркованих даних без участі людини.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Petrosiuk, D. V., Arsirii, O. O., Babilunha, O. Ju. & Nikolenko, A. O. "Deep Learning Technology of Convolutional Neural Networks for Facial Expression Recognition". Applied Aspects of Information Technology. Publ. Nauka i Tekhnika. Odessa: Ukraine. 2021; Vol.4 No.2: 192–201. DOI: <https://doi.org/10.15276/aait.02.2021.6>
2. Pseudo Labeling - QDA [Електронний ресурс]. – Режим доступу: URL: <https://www.kaggle.com/code/cdeotte/pseudo-labeling-qda-0-969/notebook>.
3. Introduction to Pseudo-Labeling : A Semi-Supervised learning technique [Електронний ресурс]. – Режим доступу: URL: <https://www.analyticsvidhya.com/blog/2017/09/pseudo-labeling-semi-supervised-learning-technique/>.
4. Pseudo-Label : The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks [Електронний ресурс]. – Режим доступу: URL: https://www.researchgate.net/publication/280581078_Pseudo_Label_The_Simple_and_Efficient_Semi-Supervised_Learning_Method_for_Deep_Neural_Networks
5. fer2013 – Режим доступу: URL: <https://drive.google.com/file/d/1pHLzOBostGFp5NIIVBL3u2F9n73kCCxZ/view?usp=sharing>

UDC 004.932

METHOD OF PSEUDO-LABELING DATA FOR MACHINE LEARNING OF DEEP CONVOLUTIONAL NEURAL NETWORKS IN RECOGNITION OF WEAKLY LABELED IMAGES OF FACIAL EXPRESSIONS

Kyrylo Mihaljov, Denys Petrosjuk

Ph.D., associate professor, Department IS Anatolii Nikolenko

Odesa Polytechnic National University, UKRAINE

ANNOTATION The issue of developing and researching methods of using weakly labeled data for machine learning is considered. The possibility of automating the process of increasing training samples by means of pseudo-labeling, i.e. adding previously unlabeled data without human participation, is shown.