

Вплив параметрів згорткових нейронних мереж на якість розпізнавання людини за фотопортретом

Дорогий Я. Ю.

Національний технічний університет України «КПІ», cisco.rna@gmail.com

The purpose of the study - to develop guidelines for choosing the parameters of convolutional neural networks for solving the problem of human face recognition.

ВСТУП

В [1] проаналізовано ряд згорткових нейронних мереж з різними параметрами. Мета проведеного дослідження – аналіз параметрів згорткових нейронних мереж щодо покращення якості розпізнавання людини за фотопортретом з використанням згорткових нейронних мереж.

ПАРАМЕТРИ НЕЙРОННИХ МЕРЕЖ

Існуючі рекомендації щодо вибору параметрів мають евристичний характер та дозволяють лише оцінити деякі значення, але не знайти їх точно. В серії експериментів, була проведена спроба знайти оптимальну конфігурацію згорткової нейронної мережі для розпізнавання обличч людей з бази обличч *ORL Faces* [2].

Проведені експерименти дали змогу визначити певний зв'язок між багатьма параметрами нейронної мережі і їх впливом на процес навчання.

Кількість шарів згортки-субдискретизації більше 3-х робить мережу настільки великою, що її навчання займає більше 200 епох, тому ці мережі не були розглянуті в дослідженні. Цілком можливо, що такі мережі можуть бути в подальшому використані для ідентифікації більш складних об'єктів в складніших умовах, наприклад кольорові зображення людей високої чіткості, 3D-зображення, розпізнавання складних об'єктів тощо.

Кількість карт ознак має значний вплив на розмір нейронної мережі, а також на її загальне функціонування. Кількість карт ознак в кожному шарі повинна обиратись індивідуально виходячи з поставленої перед згортковою мережею задачі. Для гнучкого налаштування мережі, що впливає на розмір і потужність краще підходить параметр щільності зв'язків між шарами нейронів.

Вища щільність вікон дозволяє мережі вмістити у пам'яті більше ознак зображень, тобто мережа може запам'ятати більше зображень. Недостатня щільність зв'язків може мати як негативний, так і позитивний характер. У цьому разі мережа може мати гарну здатність до узагальнення, але мати недостатню потужність для збереження можливості ідентифікації великої кількості ознак на зображенні.

Крім того, щільність зв'язків сильно впливає на швидкість навчання – згорткова нейронна мережа з високою щільністю зв'язків може навчатись досить довго. На відміну мережа з низькою щільністю може стабілізуватись за меншу кількість епох навчання, але ніколи не давати оптимальну похибку на виході.

Експериментально було визначено, що нейронна мережа може навчатися і давати прийнятну похибку у стабільному режимі, якщо щільність зв'язків знаходиться у діапазоні від 0.3 до 0.8.

Більша кількість вікон згортки дозволяє мережі вмістити у пам'яті більше ознак зображень, тобто мережа може запам'ятати більше зображень. Але за умови невеликої

тестової множини висока кількість вікон згортки може призвести до поганої узагальнюючої здатності нейронної мережі, поганого виділення ознак.

Недостатня кількість вікон згортки може мати як негативний, так і позитивний характер. У цьому разі мережа може мати гарну здатність до узагальнення, але мати недостатню потужність для збереження можливості ідентифікації великої кількості ознак на зображенні.

Крім того, кількість вікон згортки сильно впливає на швидкість навчання – згорточна нейронна мережа з високою кількістю вікон згортки може навчатись досить довго. На відміну мережа з низькою кількістю вікон згортки може стабілізуватись за меншу кількість епох навчання, але ніколи не давати оптимальну похибку на виході.

Було визначено, що великий вплив на навчання згорточної мережі має вибір початкових значень вагових коефіцієнтів. Великі початкові коефіцієнти можуть не дати мережі стабілізуватись за прийнятний час. Оптимальними є вагові значення в межах від 0.05 до 0.2 за модулем.

Висока швидкість навчання може значно прискорити процес навчання, але може і дестабілізувати нейронну мережу або призвести до неякісного навчання під час якого нейронна мережа не отримає гарної здатності до узагальнення образів. Для даної конфігурації мережі та вхідних даних оптимальним є значення у межах від 0.2 до 0.3 з обов'язковим застосуванням алгоритму віджигу з константою рівною 50-100 епохам.

Крім вище зазначених параметрів величезний вплив на навчання та подальшу роботу згорточної нейронної мережі справляє вибір правильних значень розміру вікна згортки та взаємного перетину цих вікон. Експериментально було встановлено наступні залежності:

- великі розміри вікна згортки роблять неефективним пошук просторових ознак малого розміру, що у великій кількості наявні на зображеннях людських обличч;
- також великі розміри вікна значно ускладнюють будову згорточної мережі та призводять до непропорційного збільшення потужності нейронної мережі, що справляє негативний вплив на процес навчання мережі;
- вікна згортки малого розміру дозволяють добре шукати примітивні ознаки першого порядку, але лише малого розміру;
- малі значення розміру вікна згортки можуть зменшити потужність мережі до тієї межі, коли згорточна нейронна мережа не може адекватно ідентифікувати обличчя. У цьому разі може бути потрібно збільшити щільність зв'язків між шаром та додати більше карт ознак до відповідного шару;
- оптимальне значення розміру взаємного перекриття сусідніх вікон згортки лежить в межах 50-60% від розміру вікна, що цілком співпадає з даними отриманими в [3] та [4].

ВИСНОВКИ

Використання наведених рекомендацій щодо вибору параметрів згорточної мережі надало змогу побудувати класифікатор для розпізнавання людини за фотопортретом з помилкою 1,5-2% (рис.1). Цей результат є кращим серед представлених для цього класу нейронних мереж робіт в інформаційному середовищі Інтернет.

ЛІТЕРАТУРА

- Дорогой Я.Ю. Компактные ячеистые сверточные нейронные сети для локализации лица человека // Збірник тезисів міжнародної науково-практичної конференції “Обробка сигналів і негауссівських процесів”. - Черкаси, 2007.
- [1] Samaria F., Harter A. Parameterization of a Stochastic Model for Human Face Identification // Proceedings of 2nd IEEE Workshop on Applications of Computer Vision. - Sarasota FL, December 1994. - P.1345.
- Lecun Y., Bottou L., Bengio Y., Haffne P. Gradient-Based Learning Applied to Document Recognition // Proc. IEEE. - 1998. - P.59-67.

Lawrence S., Giles C. Lee, Tsoi Ah Chung, Andrew D Face Recognition: A Convolutional Neural Network Approach // IEEE Transactions on Neural Networks: Special Issue on Neural Networks and Pattern Recognition. - 2000. - P.1015-1036.

