Баловсяк С.В.1, докт. техн. наук, доцент

Стець С. Ю.2, аспірант

Чернівецький національний університет імені Юрія Федьковича, м. Чернівці

1Кафедра комп'ютерних систем та мереж, доцент

2Кафедра програмного забезпечення комп’ютерних систем, аспірант

**Використання модуля Inception для підвищення точності розпізнавання зображень У згорткових нейронних мережах**

Розпізнавання зображень є складним завданням, оскільки форма об’єктів на зображеннях може бути різноманітною. Об’єкти на зображеннях також відрізняються за розмірами, положенням, ракурсом та умовами освітлення. Тому для розпізнавання зображень об’єктів є ефективними штучні нейронні мережі (ШНМ), зокрема, згорткові нейронні мережі (ЗНМ). Проте, при використанні ЗНМ однією з ключових проблем є збільшення кількості параметрів моделі нейронної мережі, що може призвести до перенавчання та зменшення точності розпізнавання. Тому в роботі використано модуль Inception для підвищення точності розпізнавання зображень в ЗНМ з архітектурою GoogLeNet. Модуль Inception [1, 2] застосовується для підвищення точності класифікації зображень, зменшення кількості параметрів та прискорення роботи ЗНМ.

У модулі Inception використовуються різні розміри фільтрів та шарів підсумування для отримання ознак зображення. Такий модуль складається з кількох згорткових та пулінгових шарів із різними розмірами та кількістю фільтрів [3]. Це дозволяє ЗНМ виявляти об'єкти різних розмірів на зображенні та зменшити кількість параметрів моделі.

Архітектура Inception розглядається як набір Inception-модулів, які повторюються кілька разів (рис. 1). Із збільшенням кількості модулів збільшується глибина нейромережі та покращується її абстрактність.

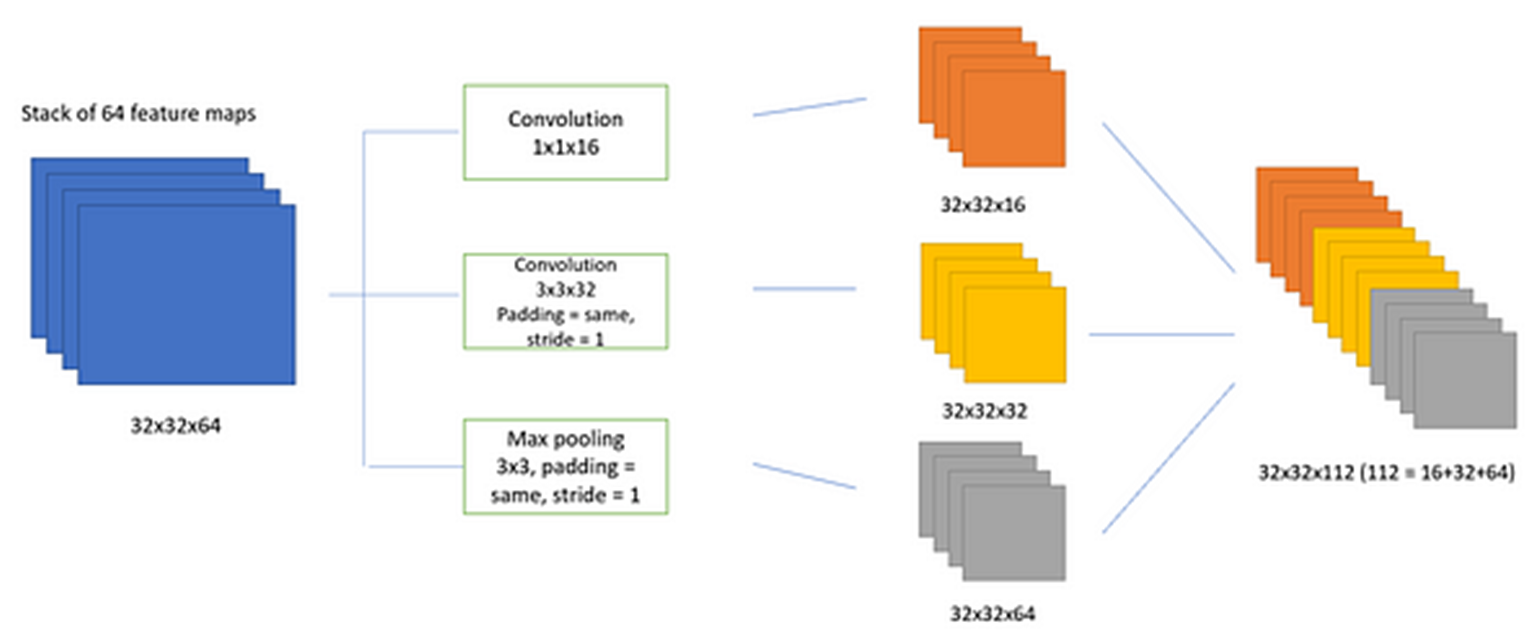


Рис 1. Приклад модуля Inception

В модулі Inception вхідні дані проходять кілька шляхів обробки, які називаються гілками. Кожна гілка містить свій власний набір фільтрів для обробки вхідних даних. Іншими словами, Inception-модуль є з'єднанням кількох фільтрів із різними розмірами ядер. Такий підхід дозволяє мережі вибирати оптимальний розмір ядра для кожної частини зображення.

Функція inception\_module (рис. 2) реалізує Inception модуль, запропонований в роботі "Going Deeper with Convolutions" від команди Google [1]. Цей модуль є ключовим елементом архітектури GoogLeNet для ЗНМ, призначених для розпізнавання зображень.



Рис.2 Реалізація Inception модуля у мові програмування Python.

Внутрішній механізм роботи Inception модуля полягає в тому, що на входи до функції inception\_module (яка обробляє тензор inputs) додається до чотирьох різних типів згорткових фільтрів з розмірами 1 × 1, 3 × 3, 5 × 5 елементів та оператор Max pooling з ядром 3 × 3 елементів. Кожний з цих типів фільтрів виконує згорткові операції над вхідним тензором inputs. Результатом цих операцій є три тензори, які містять відповідно виходи фільтрів з розмірами 1  × 1, 3 × 3 та 5 × 5 елементів, та четвертий тензор, який містить відповідно вихід оператора Max pooling. Далі ці чотири тензори конкатенуються в один тензор за допомогою функції concatenate з параметром axis=-1. Отриманий результат є вихідним значенням функції inception\_module. Цей код можна використовувати для створення згорткових нейронних мереж, заснованих на архітектурі GoogLeNet, а також для будь-яких інших задач, де використання Inception модуля є доцільним.

Вищеописаний підхід з використанням модуля Inception дозволив досягти вищої точності при класифікації зображень за допомогою ЗНМ. Це було підтверджено при розпізнаванні зображень поширених наборів даних, зокрема, набору ImageNet. Також перевагою модуля Inception є зменшення часу навчання ЗНМ, що особливо важливо для «глибоких» нейромереж із великою кількістю шарів.

Література

1. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, 1(7), 1-9. – Режим доступу до ресурсу: <https://arxiv.org/pdf/1409.4842.pdf>
2. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826). – Режим доступу до ресурсу: <https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.pdf>
3. Alto, V. (2020). Understanding the Inception Module in GoogLeNet. Режим доступу до ресурсу: <https://valentinaalto.medium.com/understanding-the-inception-module-in-googlenet-2e1b7c406106>