

ПЕРЕВАГИ АРХІТЕКТУРИ SPINENET

Прочухан Д. В.

Відокремлений структурний підрозділ Харківський комп'ютерно-технологічний фаховий коледж Національного технічного університету «Харківський політехнічний інститут», м. Харків

Переваги згорткових нейронних мереж наступні: менша кількість вагових коефіцієнтів, що налаштовуються, порівняно зі звичайними нейронними мережами, можливість використання широкого спектру технологій і методів глибокого навчання, незалежність результату дослідження від перешкод. Недоліки згорткових нейронних мереж наступні: велика тривалість часу навчання, необхідність у великій кількості набору даних для навчання, існування великої ймовірності перенавчання, багато параметрів мережі, що варіюються, вибираються дослідниками емпірично і впливають на результат. Однією з ключових особливостей згорткових нейронних мереж є перетворення вихідного зображення у послідовність проміжних ознак меншої розмірності. Зазначена особливість успішно застосовується під час вирішення завдань класифікації зображень. Однак у завданнях детектування та локалізації об'єктів на зображенні її використання пов'язано з певними обмеженнями. Для їх усунення застосовуються згорткові нейронні мережі виду «кодувальник-декодувальник» (Convolutional Encoder-Decoder Neural Network), які мають архітектуру «пісочний годинник». В архітектурах зазначеного типу містяться два структурні компоненти – декодувальник і кодувальник. Декодувальник розташовується над кодувальником. Декодувальник необхідно застосовувати для вирішення завдань класифікації. Кодувальник використовується для вирішення задач локалізації. Кодувальник є головним компонентом згорткової нейронної мережі. Тому кодувальник містить більше параметрів, ніж декодувальник. Головний недолік архітектури типу «кодувальник-декодувальник» – неможливість ефективно генерувати ознаки різних масштабів, необхідні для успішного вирішення завдання одночасного детектування та локалізації об'єктів на зображенні. Зазначений недолік був усунений у згортковій нейронній мережі SpineNet. Модель SpineNet дозволяє виявляти ознаки різних масштабів завдяки блокам певних розмірів. Розміри блоків підбираються за допомогою механізму нейронного пошуку архітектури (Neural Architecture Search, NAS). Кожен блок може підключатися лише до батьківського блоку. Упорядкування блоків має значення. Кожен блок має можливість налаштування певного рівня масштабування та типу. Рівні масштабування проміжних блоків можуть змінюватись. Архітектура SpineNet може містити звужені та залишкові блоки. Використання архітектури SpineNet як базової моделі дозволяє досягти високого показника середньої точності, який становить 52,1%. У даній моделі використовується менша кількість параметрів та операцій за секунду у порівнянні з аналогами. Архітектуру SpineNet було успішно застосовано для вирішення задач детектування об'єктів на зображеннях.