

АНАЛІЗ АРХІТЕКТУР МОДЕЛЕЙ ГЛИБОКОГО НАВЧАННЯ ПРИ ДІАГНОСТИЦІ ЗАХВОРЮВАНЬ ВУХА

Вінницький національний технічний університет

Анотація

В даній роботі проведено аналіз різних архітектур моделей глибокого навчання, які можуть бути використані при діагностиці захворювань вуха.

Ключові слова: нейронні мережі, моделі глибокого навчання, архітектура, оториноларингологія

Для виявлення та адекватного медичного лікування хвороб вух велике значення має раннє виявлення, особливо в умовах відсутності експертів та низької точності діагностики.

Як відомо нам усім, слух вважається одним з найважливіших п'яти почуттів, оскільки наші життя в основному залежать від нього. Якщо захворювання вух не виявити рано та не лікувати відповідно, це може залишити негативні наслідки, наприклад, порушення слуху. При оцінці захворювань вух в клініці, звичайна отоскопія або отоендоскопія є важливою складовою фізичного обстеження на першому етапі. Однак отоскопія або отоендоскопія, використана в діагностиці, може легко бути помилковою для фахівців, які не є отоларингологами. Дослідження показують, що середня точність діагностики серед 514 педіатрів стану середнього вуха складала лише 50%. Така низька точність діагностики свідчить, що без допомоги додаткових ресурсів тестування діагнозу захворювання вух буде складним, навіть для експертів. З урахуванням такої ситуації існує велика потреба знайти нову методику діагностики з покращеною точністю.

Використання нейронних мереж та глибокого навчання може суттєво вплинути на якість даного процесу. Глибоке навчання було застосовано в декількох областях медичного зображення, таких як розробка та валідація алгоритму глибокого навчання для виявлення діабетичної ретинопатії на фонограмах сітківки, масштабне глибоке навчання для комп'ютерної допомоги виявленню мамографічних уражень. Більше того, глибоке навчання широко застосовується в класифікації захворювань вух і слуху. У цих застосуваннях глибоких нейронних мереж глибоке згорткове навчання грає дуже важливу роль в розпізнаванні або класифікації зображень. Для навчання моделей глибоких нейронних мереж потрібно дуже мало попередніх професійних знань. Порівняно зі звичайними спектральними класифікаторами, глибокі нейронні мережі більш точні в задачах розпізнавання зображень через мільйони ваг з кількома шарами, але вони мають високі обчислювальні витрати при навчанні моделі.

Для аналізу було обрано найбільш популярні моделі в області комп'ютерного зору, а саме: MobileNet-V3, Inception-V4, DensNet-BC121 та ResNet50.

MobileNet-V3 - це розвиток архітектури MobileNet-V2, глибоких нейронних мереж, спеціально розроблене для мобільних пристроїв та вбудованих систем з обмеженими ресурсами. Ця архітектура була представлена компанією Google і має на меті поєднати високу швидкість та ефективність з високою точністю. Особливостями даної архітектури є:

1. Підтримує множники ширини та роздільної здатності, які дозволяють регулювати обсяг обчислень та точність моделі в залежності від конкретного застосування.
2. Використовує поліпшені блоки стиснення та збудження, які допомагають зосередити увагу мережі на найбільш важливих ознаках та підвищити точність класифікації.
3. Використовує різні активаційні функції, такі як Hard-Swish та Sigmoid, які допомагають покращити швидкість та точність моделі.
4. Використовує методи автоматичного пошуку архітектури нейронних мереж, такі як NAS, для автоматичного вибору оптимальної архітектури для конкретної задачі.

Inception-V4 є одним з варіантів архітектури глибокої нейронної мережі, розробленою командою Google Research. Принцип роботи Inception-V4 ґрунтується на використанні блоків, які включають в себе кілька різних операцій згортки та пулінгу, що виконуються паралельно. Це

дозволяє моделі ефективно використовувати різні шляхи абстракції та вибирати найбільш ефективний підхід для кожного конкретного завдання. Особливостями даної архітектури є:

1. Inception-V4 використовує блоки Inception, які складаються з кількох підблоків. Кожен підблок містить операції згортки та пулінгу з різними розмірами ядер, які працюють паралельно. Це дозволяє моделі ефективно використовувати різні варіанти вибірки ознак.
2. Включає методи зменшення кількості ознак у великих розмірностях даних, що дозволяє зберігати більше інформації за меншої кількості параметрів.
3. Запобігання перенавчанню, для цього використовуються різні техніки регуляризації, такі як дропаут та L2-регуляризація.

DenseNet-BC121 - це архітектура глибокої нейронної мережі, яка характеризується високою точністю класифікації зображень та ефективним використанням обчислювальних ресурсів. Ця архітектура базується на ідеї щільних (повнозв'язаних) з'єднань між шарами мережі. Особливостями даної архітектури є:

1. Кожен блок мережі містить кілька підблоків, в кожному з яких кожен шар з'єднаний з кожним попереднім шаром. Це дозволяє передавати інформацію від кожного шару безпосередньо до кожного іншого шару, що поліпшує потік градієнтів та дозволяє ефективно використовувати ознаки на всіх рівнях мережі.
2. Для зменшення кількості параметрів у глибоких мережах використовуються bottleneck шари, які скорочують кількість каналів передачі ознак перед подальшим використанням у більш глибоких шарах.
3. Використання механізму стиснення для зменшення кількості каналів ознак після кожного блоку, що дозволяє зменшити обсяг пам'яті та обчислень у мережі.
4. Використовується нормалізація партії (batch normalization) для стабілізації та прискорення навчання мережі.

ResNet50 - це архітектура глибокої нейронної мережі, розроблена командою Microsoft Research. Ця архітектура відома своєю здатністю до ефективного навчання глибоких нейронних мереж із великою кількістю шарів. Принцип роботи ResNet50 базується на використанні блоків з'єднань залишкових з'єднань (residual connections), які допомагають уникнути проблеми зникнення градієнтів та поліпшують потік градієнтів під час навчання глибоких мереж. Основні характеристики ResNet50 включають:

1. Кожен блок мережі містить пару з'єднань залишкових з'єднань (residual connections), які дозволяють передавати градієнти безпосередньо через блоки, навіть якщо там є багато шарів. Це допомагає уникнути проблеми зникнення градієнтів (vanishing gradients) та дозволяє навчати глибші мережі.
2. Використання згорткових шарів для вилучення ознак з вхідних зображень. Ці шари виконують операції згортки для виявлення різних ознак, таких як краї, текстури, форми тощо.
3. Наявність пулінгових шарів, які дозволяють зменшити розмірність ознак, зберігаючи при цьому важливість.
4. Використовує нормалізацію партії (batch normalization) для стабілізації та прискорення навчання мережі.

Для порівняння даних архітектур було використано фреймворк PyTorch. В якості навчальної бази було використано зображення з відкритого доступу. Результат порівняння продемонстровано в таблиці 1.

Таблиця 1. Порівняння архітектур нейронних мереж

Назва	Точність	Час навчання(с)	Час обробки моделі(с)
MoblieNet-V3	90.854	44.758	0.0857
Inception-V4	91.210	128.27	0.2309
DensNet-BC121	92.112	84.192	0.1159
ResNet50	89.521	81.098	0.1668

Результати порівняння показують, що оглянуті архітектури нейронних мереж відрізняються досить високою точністю, однак MobileNet-V3 потребує найменших ресурсів для роботи та має найбільшу швидкодію. Однак, для остаточних висновків необхідно провести дослідження на більшій вибірці даних, з використанням медичних зображень різної якості.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. John C. Watkinson, Ray W. Clarke. "Scott-Brown's Otorhinolaryngology and Head and Neck Surgery" CRC Press, 2018.
2. Isaac N. Bankman. "Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis" Academic Press, 2008. P.169-208.
3. E. R. Davies. "Computer Vision: Principles, Algorithms, Applications, Learning" Academic Press, 2017.
4. Gulshan, V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA Intern. Med. 316, 2402–2410 (2016).
5. Blomgren, K. & Pitkäranta, A. Is it possible to diagnose acute otitis media accurately in primary health care?. Fam. Pract. 20, 524–527 (2003).
6. Sermanet, P., Eigen, D., Zhang, X. et al. Overfeat: Integrated recognition, localization and detection using convolutional networks. <https://arxiv.org/abs/1312.6229> (2017)
7. Sandler, M., Howard, A., Zhu, M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE conference on computer vision and pattern recognition. <https://arxiv.org/abs/1801.04381> (2018).

Марчук Андрій Юрійович - аспірант кафедри біомедичної інженерії та оптико-електронних систем, Вінницький національний технічний університет, м. Вінниця, andriu4934@gmail.com.

ANALYSIS OF ARCHITECTURES OF DEEP LEARNING MODELS IN THE DIAGNOSTIC OF EAR DISEASES

Abstract. This article analyzes various architectures of deep learning models that can be used in the diagnosis of ear diseases.

Keywords: neural networks, deep learning models, architecture, otorhinolaryngology.

Marchuk Andrii Yuriyovych - Postgraduate student, Department of Biomedical Engineering, Vinnytsia National Technical University, Vinnytsia, andriu4934@gmail.com.