

УДК 004.89

Я.В. Стешенко, студент гр. ПК-11мн, доктор філософії Момот А. С.
КПІ ім. Ігоря Сікорського

ОГЛЯД ЕФЕКТИВНОСТІ АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ ДЛЯ АВТОМАТИЗОВАНОЇ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

Анотація. Проведено огляд різних архітектур нейронних мереж для сегментації зображень. Проаналізовано переваги і недоліки різні моделей для задач сегментації. Розглянуто переваги використання архітектури DoubleU-Net. Запропоновано використання архітектури DoubleU-Net в задачах сегментації зображень, як найкращу за показниками ефективності на сьогоднішній день.

Ключові слова: сегментація, згорткові мережі, глибинне навчання.

ВСТУП

Сегментація зображень — це завдання позначення кожного пікселя досліджуваного об'єкта на зображеннях. Сегментація зображень допомагає зосередитися на певній області та отримати детальну інформацію для більш точного аналізу. Основними проблемами, пов'язаними з сегментацією зображень, є відсутність великої кількості високоякісних зображень для навчання, відсутність стандартного протоколу сегментації та великі варіації зображень.

Згорткові нейронні мережі (CNN) показали найвищу продуктивність для автоматизованої сегментації зображень [1]. Для завдань сегментації однією з початкових архітектур глибинного навчання, навчених для попиксельного прогнозування, є повністю згорткова мережа (FCN).

Узагальнення, тобто здатність моделі працювати в незалежному наборі даних, і стійкість, тобто здатність моделі працювати на складних зображеннях, є ключовими в системах штучного інтелекту. Тому важливо обрати потужну архітектуру, яка є надійною та узагальненою для різних застосувань. Попередньо підготовлені моделі, навчені на наборі даних ImageNet, значно покращили ефективність архітектур CNN [2]. Це дозволяє розробити нові мережі для сегментації та порівняти їх результативність.

ОГЛЯД ПОПЕРЕДНІХ РОБІТ

У статті [3] адаптовано та протестовано архітектуру Mask R-CNN за допомогою трьох попередньо навчених нейромереж: Resnet50, Resnet101 і Inception Resnet (v2). Хоч більш глибинна мережа має підвищувати ефективність сегментації в області природних зображень, але Resnet50 зміг перевершити Resnet101 і Inception Resnet (v2) в умовах обмеженої кількості навчальних даних. Після покращення репрезентативності навчальної вибірки, всі три моделі підвищили ефективність сегментації. Результати підтверджують, що з кращим набором навчальних даних Mask R-CNN є доволі перспективною моделлю.

Мережа U-Net [4] — ще одна популярна архітектура для сегментації зображень. Архітектура U-Net складається з двох частин, а саме шляху аналізу та шляху синтезу. На шляху аналізу вивчаються глибокі ознаки, а на шляху синтезу виконується сегментація на основі вивчених ознак. Крім того, U-Net використовує операцію skip-connections. Skip-connections дозволяє проходити вхідному сигналу через мережу, минаючи нелінійні перетворення шарів. Таким

чином, просторова інформація застосовується до глибшого шару, що значно дає точнішу вихідну карту сегментації. Таким чином, додавання більшої кількості рівнів до U-Net дозволяє мережі дізнатися більше репрезентативних подань даних, що призводить до кращих вихідних масок.

У статті [5] автори провели ретельний аналіз архітектури U-Net з надією знайти потенційні можливості для вдосконалення. Вони запропонували ввести короткі шляхи Res, які вводять деяку додаткову обробку, щоб зробити карти ознак більш однорідними. Крім того, щоб доповнити U-Net можливістю аналізу з різною роздільною здатністю, вони запропонували блоки MultiRes. Автори черпали натхнення з блоків Inception і створили компактну аналогічну структуру, яка є легкою та потребує менше пам'яті. Використовуючи ці модифікації, вони розробили нову архітектуру MultiResUNet.

Авторами статті [6] представлено архітектуру під назвою U-Net++ для більш точної сегментації зображень. Продуктивність U-Net++ пояснюється його вкладеною структурою та переробленими skip-connections, які спрямовані на вирішення двох ключових проблем U-Net: 1) невідома глибина оптимальної архітектури та 2) надмірно обмежувальний дизайн skip-connections. Автори оцінили U-Net++, використовуючи шість різних програм для отримання зображень, і продемонстрували постійне підвищення продуктивності в порівнянні з іншими основними моделями для сегментації.

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МОДИФІКАЦІЙ U-NET

Мережа U-Net все ще вважається основною для різноманітних завдань сегментації зображень. Однак вона має ряд сучасних модифікацій. Наприклад, відома архітектура під назвою DoubleU-Net порівнюється з U-Net у роботах, які використовували той самий набір даних для навчання, щоб продемонструвати її ефективність [7].

Архітектура мережі DoubleU-Net, наведена авторами роботи [7], показана на рис. 1. Результати досліджень показують, що DoubleU-Net досягає значення 0.92 для коефіцієнта Сьоренсена – Дайса (DSC), що вище на 3.91% і mIoU (метрика оцінки якості сегментації) на рівні 0.8611, що вище на 1.14% у порівнянні з базовою моделлю. Ретельний візуальний аналіз результату показує, що DoubleU-Net створює кращі маски сегментації порівняно з початковою мережею.

Модель, яка використовує попередньо навчену мережу ImageNet (наприклад, Resnet101 або VGG-16) як основу, досягає вищого результату при порівнянні з моделлю навчання мережі з нуля. DoubleU-Net досягає DSC на рівні 0.8962 і mIoU на рівні 0.8212 на іншому наборі даних. У порівнянні з U-Net, DoubleU-Net демонструє кращі результати на 5.7% за показником mIoU.

Іншою модифікацією початкової моделі є вже згадана U-Net++. Вона що використовує мережу Resnet101 як базову архітектуру. Однак, архітектура DoubleU-Net має DSC на 1.59% вище, ніж зазначена мережа U-Net++. В той же час, мережа U-Net++ демонструє помітно кращі результати у порівнянні з початковою моделлю U-Net.

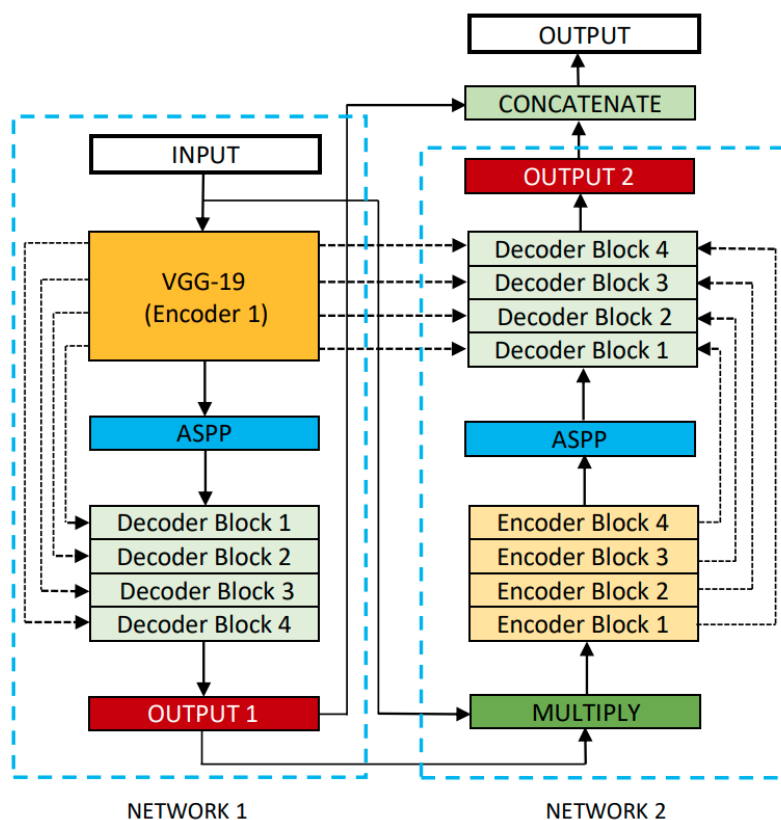


Рис. 1. Архітектура мережі DoubleU-Net

Повні результати порівняння описаних архітектур на наборі даних 2018 Data Science Bowl Challenge наведені у таблиці 1.

Таблиця 1. Оцінка результатів сегментації на наборі даних 2018 Data Science Bowl Challenge для різних архітектур

Архітектура	Перенесення навчання	DSC	mIoU
U-Net	Resnet101	0.7573	0.9103
U-Net++	Resnet101	0.8974	0.9255
DoubleU-Net	VGG-19	0.9133	0.8407

З результатів оцінки якості різних архітектур видно, що DoubleU-Net здатна створити кращу маску сегментації навіть для складних зображень. Загальний результат оцінки ефективності показує, що архітектура добре працює для різних наборів даних сегментації зображень. Таким чином, наведені вище результати свідчать про надійність запропонованої моделі. Перенесення навчання з попередньо навченої на наборі ImageNet мережі значно покращує результати кожної розглянутої архітектури, що компенсує відсутність достатньої кількості навчальних даних для інших завдань.

Таким чином, DoubleU-Net являє собою нову основну модель для вирішення завдання сегментації зображення. DoubleU-Net є більш узагальненим і може використовуватися у різних галузях.

ВИСНОВКИ

У цій статті оглянуто декілька архітектур та запропоновано використання архітектури CNN під назвою DoubleU-Net, як найкращу серед описаних. Продуктивність DoubleU-Net значно краща порівняно з початковою моделлю U-Net на різних тестових наборах даних, що підтверджує її універсальність.

Крім того, розглянута архітектура DoubleU-Net є гнучкою, що дає можливість легко інтегрувати до неї інші блоки. Результати сегментації можна покращити шляхом подальшої інтеграції різних відомих блоків CNN та використання методів постобробки результатів, таких як умовне випадкове поле та поріг Оцу.

Актуальним завданням для майбутніх досліджень є створення ефективної моделі для завдань сегментації зображень з дефектами. Пропонується зосередитися на спрощенні архітектури, зберігаючи її здатність створювати якісні маски сегментації. Обмеженням DoubleU-Net є те, що вона використовує більше параметрів порівняно з U-Net, що призводить до збільшення часу навчання. У майбутньому дослідження мають більше зосереджуватися на розробці спрощених архітектур із меншою кількістю параметрів, зберігаючи при цьому свої можливості.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Momot A. Deep learning automated data analysis of security infrared cameras / A. Momot, I. Skladchykov // Slovak international scientific journal. – 2021. – №52. – pp. 13–16.
- [2] Momot A. Automation of ultrasound breast cancer images classification using deep neural networks / A. Momot, R. Galagan, M. Zaboluiieva // Sciences of Europe. – 2022. – №96. – pp. 38–41.
- [3] A survey on deep learning in medical image analysis / [G. Litjens, T. Kooi, B. Bejnordi та ін.]. // Medical image analysis (MedIA). – 2017. – №42. – С. 60–88.
- [4] Ronneberger O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox. // In Proceedings of International Conference on Medical image computing and computer-assisted intervention (MICCAI). – 2015. – С. 234–241.
- [5] A Polyp detection and segmentation using mask r-cnn: Does a deeper feature extractor cnn always perform better? / [H. Qadir, Y. Shin, J. Solhusvik та ін.]. // In Proceeding of International Symposium on Medical Information and Communication Technology (ISMICT). – 2019. – С. 1–6.
- [6] Unet++: Redesigning skip connections to exploit multiscale features in image segmentation / Z.Zhou, M. Siddiquee, N. Tajbakhsh, J. Liang. // IEEE Transactions on Medical Imaging. – 2019.
- [7] Ibtehaz N. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation / N. Ibtehaz, M. Rahman. // Neural Networks. – 2020. – №121. – С. 74–87.

Наук. керівник – доктор філософії, ст.викл. Момот А.С.