

УДОСКОНАЛЕННЯ ПРОЦЕСУ АНОТУВАННЯ МЕДИЧНИХ ЗОБРАЖЕНЬ ДЛЯ ЗАДАЧ СЕГМЕНТАЦІЇ ЗА ДОПОМОГОЮ ТЕКСТОВИХ ПІДКАЗОК

Вінницький національний технічний університет

Анотація

В даній роботі досліджується потенціал моделі Language Segment-Anything для визначення областей інтересу (англ.: ROI – Region of Interest) медичних зображень. Завдяки інтеграції текстових підказок (англ.: text prompts) у структуру Segment Anything Model (SAM) від Meta і розробці зручної для користувача програми з використанням Streamlit, дана робота ілюструє, як такий підхід дозволяє визначати псевдомаски регіонів інтересу без попереднього процесу розмітки даних. Це спрощує процес сегментації зображень, роблячи його швидшим, точнішим та доступнішим для користувачів. Вправність моделі Language Segment-Anything в обробці інструкцій природною мовою та її здатність зменшити потребу в ручному маркуванні даних має значний потенціал у галузі медичної візуалізації.

Ключові слова: аналіз зображень, машинне навчання, глибоке навчання, сегментація зображень, псевдомаски, веб-застосунок.

Abstract

This work investigates the potential of the Language Segment-Anything model for extracting regions of interest (ROI) from medical images. By integrating text prompts into the Segment Anything Model (SAM) framework from Meta and developing a user-friendly application using Streamlit, this work demonstrates how such an approach enables the extraction of pseudomasks of regions of interest without the need for prior data annotation. This simplifies the image segmentation process, making it faster, more accurate, and more accessible for users. The proficiency of the Language Segment-Anything model in handling natural language instructions and its ability to reduce the need for manual data labeling holds significant promise for the field of medical imaging.

Keywords: image analysis, machine learning, deep learning, image segmentation, pseudomasks, web application.

Вступ

У сфері медичного аналізу точність має першочергове значення. Проте точність моделей сегментації може бути знижена через вплив артефактів та шуму, що притаманні медичним зображенням. Такі недосконалості створюють проблеми, які традиційними підходами до сегментації важко подолати, що часто призводить до неоптимальних результатів.

На допомогу може прийти використання різних підходів попередньої обробки даних. Методи попередньої обробки даних відіграють важливу роль у підвищенні точності моделей сегментації медичних зображень, як показано в попередніх роботах авторів, у яких різні фільтри, такі як CLAHE, двосторонній фільтр та інші, використовувалися для покращення якості зображення та зменшення шуму, що сприяє більш точним результатам сегментації [1, 2].

В даній роботі представлено можливості моделі Language Segment-Anything [3], яку створено на основі Segment Anything Model (SAM) від Meta [4], для завдання вилучення псевдомасок регіонів інтересу (ROI) на медичних зображеннях без необхідності попередньої розмітки даних. Даний підхід використовує потужність мовних підказок для вилучення регіонів інтересу на зображеннях із шумом та артефактами.

Однією з основних перешкод у вирішенні задач комп'ютерного зору є необхідність ретельного маркування даних — процес, який потребує багато часу та людських ресурсів [5, 6]. Однак використання моделі Language Segment-Anything зменшує залежність від ручного маркування даних, використовуючи текстові підказки, обходячи необхідність у великих наборах даних з мітками. Це не тільки спрощує процес підготовки даних для вирішення задачі сегментації стоматологічних знімків, але й зменшує фінансовий тягар, пов'язаний з анотацією даних.

Огляд методу створення псевдомасок регіонів інтересу

Використання текстових підказок представляє новий і ефективний підхід для моделі Language Segment-Anything для виділення псевдомасок регіонів інтересу (ROI) із медичних зображень. На відміну від традиційних методів, які значною мірою покладаються на вручну анотовані набори даних для навчання [7], текстові підказки пропонують більш універсальне рішення.

Моделю Language Segment-Anything вправно обробляє інструкції природною мовою, дозволяючи користувачам передавати конкретні критерії або характеристики бажаної псевдомаски ROI за допомогою текстового введення. Такі підказки спрямовують увагу моделі на релевантні елементи зображення, таким чином відфільтровуючи шум та артефакти. Наприклад, підказка може описати бажані анатомічні структури, що присутні на зображенні.

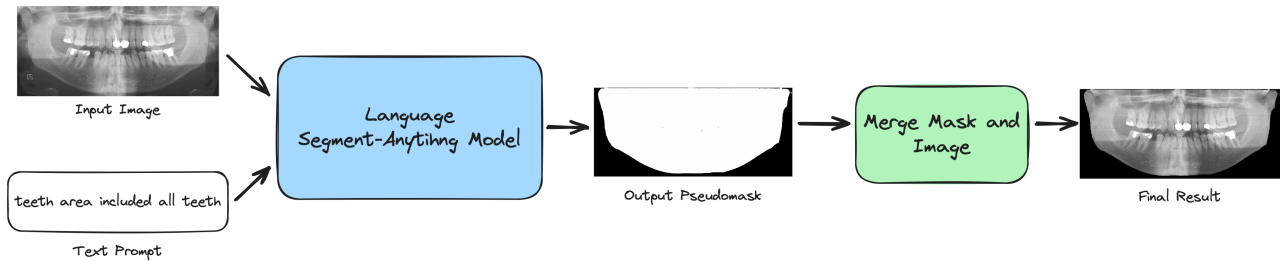


Рисунок 1 – Діаграма роботи методу для створення псевдомасок регіонів інтересу

Оскільки Language Segment-Anything інтерпретує та відповідає на кожну підказку, користувачі можуть коригувати свої інструкції на основі відгуків моделі, точно налаштовуючи сегментацію для досягнення оптимальних та коректних псевдомасок.

Такий текстовий підхід не тільки прискорює процес сегментації, але й підвищує його точність і адаптивність. Використовуючи можливості природної мови, Language Segment-Anything долає обмеження звичайних методів сегментації, прокладаючи шлях для більш ефективного та точного аналізу медичних зображень та створення псевдомасок регіонів інтересу без потреби ручного анотування даних.

Розробка застосунку

Для того, щоб зробити використання моделі Language Segment-Anything для вилучення псевдомасок областей інтересу із медичних зображень ефективнішим, створено простий застосунок використовуючи бібліотеку Streamlit [8]. Цей застосунок надає користувачам зрозумілий та простий інтерфейс, що дозволяє завантажувати зображення, вводити текстові вказівки і отримувати результати виділеної області інтересу, який можна зберегти на локальній пристрій користувача. Таким чином, створення псевдомасок регіонів інтересу стає досить зручним та простим процесом.

Image Segmentation with Text Prompt

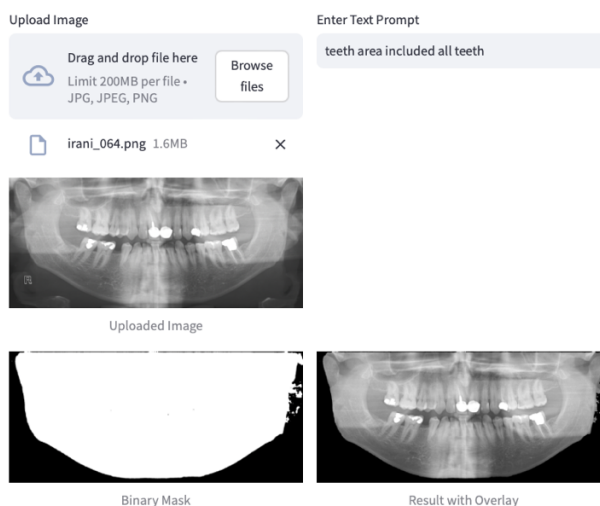


Рисунок 2 – Приклад роботи застосунку

Процес розробки починається з інтеграції можливостей моделі Language Segment-Anything у структуру програми. Використовуючи зручний дизайн Streamlit, програма може надати користувачам зручну взаємодію з моделлю.

Ключові компоненти програми включають наступне:

1. Завантаження зображень: користувачам надається можливість завантажувати медичні зображення безпосередньо в додаток. Віджет завантажувача файлів Streamlit можна використовувати для полегшення цього процесу, дозволяючи користувачам легко вибирати та завантажувати зображення зі своїх локальних пристроїв.

2. Текстова підказка: після завантаження зображення користувачам пропонується ввести текстові інструкції або підказки із зазначенням бажаних характеристик регіонів інтересу, які вони бажають виділити. Це поле введення служить засобом для користувачів, щоб повідомити свої вимоги до моделі Language Segment-Anything.

3. Обробка та вивід: після отримання зображення та текстової підказки, програма викликає модель Language Segment-Anything для аналізу зображення та визначення відповідних регіонів на основі наданих інструкцій. Витягнуті псевдомаски області інтересу потім відображаються користувачеві в інтерфейсі програми, пропонуючи візуальне представлення сегментованої області.

4. Інтерактивний зворотний зв'язок: щоб покращити взаємодію з користувачем і полегшити ітераційне вдосконалення процесу сегментації, програма включає інтерактивні елементи. Користувачі мають можливість налаштувати текстові підказки на основі вихідних даних, дозволяючи ітеративно уточнювати сегментацію, доки не буде досягнуто задовільних результатів.

Інкапсулюючи функціональні можливості моделі Language Segment-Anything у зручній програмі Streamlit, користувачі отримують доступ до потужного інструменту для легкого та ефективного вилучення областей інтересу з медичних зображень. Інтуїтивно зрозумілий інтерфейс та інтерактивні функції дають змогу користувачам ефективно використовувати можливості запропонованого підходу для вилучення псевдомасок регіонів інтересу, покращуючи робочі процеси аналізу медичних зображень.

Висновки

Підсумовуючи, інтеграція моделі Language Segment-Anything у процес анотування медичних зображень та підбір відповідних текстових підказок може суттєво спростити вирішення задач сегментації за допомогою більш швидкого та зручного визначення регіонів інтересу. Використовуючи текстові підказки та інтуїтивно зрозумілий інтерфейс Streamlit, процес вилучення стає не тільки точнішим, але й доступнішим для користувачів. Завдяки поєднанню обробки мови та методів аналізу зображень Language Segment-Anything пропонує універсальне рішення, яке адаптується до проблем, що пов'язані із зашумленими та навантаженими артефактами медичними зображеннями. Більше того,

здатність такого підходу оптимізувати процес сегментації без необхідності маркування великих даних демонструє зміну парадигми у використанні ресурсів та ефективності.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Komenchuk O. V. and Mokin O. B., "Analysis of Methods for Preprocessing of Panoramic Dental X-Rays for Image Segmentation Tasks," *Visnyk of Vinnytsia Politechnical Institute*. 2023. Vol. 170, no. 5. P. 41–49. URL: <https://doi.org/10.31649/1997-9266-2023-170-5-41-49>.
2. H. Abdi, S. Kasaei, and M. Mehdizadeh, "Automatic segmentation of mandible in panoramic x-ray," *J. Med. Imaging (Bellingham)*, vol. 2, no. 4, 044003, 2015. [Online]. Available: https://www.academia.edu/36038975/Pre-Processing_of_Dental_X-Ray_Images_Using_Adaptive_Histogram_Equalization_Method.
3. "Language Segment-Anything," GitHub, [Online]. Available: <https://github.com/luca-medeiros/lang-segment-anything>
4. "Segment Anything," GitHub, [Online]. Available: <https://github.com/facebookresearch/segment-anything>
5. Yi Li, Zhanghui Kuang, Liyang Liu, Yimin Chen and Wayne Zhang, "Pseudo-Mask Matters in Weakly-Supervised Semantic Segmentation," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021. P. 6964-6973. URL: <https://doi.org/10.48550/arXiv.2108.12995>
6. Abdi and S. Kasaei, "Panoramic Dental X-rays With Segmented Mandibles," 2020. [Online]. Available: <https://data.mendeley.com/datasets/hxt48yk462/2>
7. Jie Hu, Chen Chen, Liujuan Cao, Shengchuan Zhang, Annan Shu, Guannan Jiang and Rongrong Ji, "Pseudo-label Alignment for Semi-supervised Instance Segmentation," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2023. P. 16337-16347. URL: <https://doi.org/10.48550/arXiv.2308.05359>
8. "Streamlit," GitHub, [Online]. Available: <https://github.com/streamlit/streamlit>

Коменчук Олег Вікторович – аспірант кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: komenchuk77@gmail.com

Мокін Олександр Борисович – д-р техн. наук, професор, професор кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: abmokin@gmail.com