

CS75 – Exploration on Project requirements and Literatures

V2.0

Overview

This project "Interactive Bio-Image Annotation Tool Based on Large-Scale Basic Vision Models" (CS75) is mainly to develop an advanced bio-image annotation platform to improve the efficiency and accuracy of bio-image analysis.

Main Contents:

Project Goal (Mainly for literature review)

- The tool uses **Napari** for interactive visualization. **Napari needs to support multi-channel images (RGB, multi-channel support for medical images).**
 - What is Napari? [Link to go](#)
 - Main features:
 - Multi-dimensional support: It can efficiently process and visualize 2D, 3D and higher-dimensional image data.
 - Interactive operation: It provides an intuitive user interface and supports operations such as zooming, rotating, and panning of images, which facilitates users to explore data in depth.
 - Plug-in ecosystem: With a rich plug-in library, users can expand Napari's functions according to specific needs to meet the visualization and analysis needs of different fields.
 - Related literature and tutorials:
 - [Napari Image Visualization Tutorial](#)
 - [Using Napari for interactive visualization](#)
 - [Napari Start Guide](#)
- Combined with **Segment Anything Model (SAM)** for automated cell segmentation, based on large-scale vision models to accurately detect targets in multi-dimensional microscope images. **SAM is primarily used for cue-driven segmentation, but needs to be improved to accommodate 3D dimensional data (it currently does not handle 3D data efficiently). We need to find ways to improve on 3D data while limiting video data.**
 - What is SAM? [Link to go](#)
 - Main features:
 - Hint-driven segmentation: SAM accepts multiple forms of hints (such as points, boxes, text, or masks) to indicate the object to be segmented. [Link to go](#)
 - Zero-shot transfer capability: Without specialized training or fine-tuning for specific tasks or datasets, SAM can be directly applied to new image distributions and tasks, demonstrating strong generalization capabilities. [Link to go](#)
 - Large-scale dataset support: SAM is trained on a large-scale dataset (SA-1B) containing more than 110,000 images and 1 billion masks,

ensuring the robustness and wide applicability of the model. [Link to go No.1](#) / [Link to go No.2](#) / [Link to go No.3](#)

- Related Literature and tutorials:
 - [Segment Anything User Manual \(Interactive Data Annotation\)](#)
 - [Segment Anything Model Basic Usage Learning Experience](#)
 - [General Image Segmentation Model: Segment Anything Model \(SAM\) Training Tutorial](#)
 - [LLM Large Model: Segment Anything Model Principle Explanation](#)
 - [Open Source Automatic Segmentation Segment-Anything Cutout AI Usage Tutorial](#)

Researchers can manually adjust the segmentation results generated by AI to achieve an automated + manually corrected annotation method.

- Specific process:
 - Automated segmentation: Use deep learning models (such as U-Net) or other image segmentation algorithms to perform preliminary segmentation of medical images.
 - Manual correction: Experts review and make necessary modifications to the automatic segmentation results to correct possible errors in the model.
 - Iterative optimization: Feed the corrected results back to the model for retraining or fine-tuning to gradually improve the segmentation performance of the model.
- Related Literature and tutorials:
 - [New breakthrough in medical image segmentation: 6 articles give you insight into the most cutting-edge medical AI technology](#)
 - [Research progress on automatic segmentation methods for liver cancer lesions based on electronic computed tomography images](#)
 - [CN109690554B - Method and system for medical image segmentation based on artificial intelligence](#)

Key Features (Group Proposal and Plan, can be changed later)

1. Load and visualize multi-channel biological images.

- a. Implementation steps:
 - i. Data loading: Use appropriate libraries (such as tifffile) to read multi-channel biological imaging data.
 - ii. Data processing: Perform pre-processing on the loaded image data, such as normalization or denoising. **Based on the provided dataset from our stakeholders.**
 - iii. Data visualization: Use Napari's Viewer object to add multi-channel images to the view and set different display parameters for each channel.
 - iv. **Multiple channels: Supports Napari processing of multi-channel allocation of medical images and allows users to adjust the display effects of different channels.**
- b. Required tools:
 - i. Napari: Used to visualize multi-dimensional imaging data.

- ii. `tifffile`: Used to read image files in TIFF format.

2. Deep learning-based cell segmentation (automatic segmentation).

a. Implementation steps:

- i. **Model loading: Use `sam_model_registry` to register and load the pre-trained SAM model.**
- ii. **Image processing: Call the `SamPredictor.set_image` method to convert the input image into an embedding representation that the model can process. And it needs to support scrolling and looping so that users can load different pictures.**
- iii. **Prompt input: Provide prompt information such as points and boxes as required to guide the model to segment specific areas.**
- iv. **Mask generation: Call the `predictor.predict` method to generate the segmentation mask of the target area.**

b. Tools required:

- i. Deep learning framework: Such as TensorFlow or PyTorch.
- ii. Training dataset: High-quality annotated biological image data from our project stakeholders.
- iii. **OpenCV: for image processing and preprocessing.**

3. Interactively edit cell masks (manually adjust AI results).

a. Implementation steps:

- i. Mask loading: Load the mask generated by automatic segmentation into Napari. **Multi-channel blending is implemented on Napari, allowing users to adjust the visualization effects of different layers.**
- ii. Interactive editing: Manually modify the mask to correct errors using Napari's drawing tools.
- iii. Save modifications: Save the edited mask for further analysis or model retraining.

b. Required tools:

- i. Napari: Provides interactive editing capabilities.

4. Export annotated data for further analysis.

a. Implementation steps:

- i. Data format selection: Determine the format of the exported data (such as CSV, JSON, or TIFF).
- ii. Data export: Use the corresponding library to export the labeled data to the required format.
- iii. Data validation: Ensure that the exported data is complete and in the expected format. **Additional support for RGB files and PNG image formats for mask visualization.**

b. Required tools:

- i. **Pandas: Processing tabular data.**
- ii. **`tifffile`: Processing the export of image data.**

5. The tool will be developed as a Napari plugin to ensure its compatibility with existing bio-image analysis pipelines.

- a. Implementation steps:
 - i. Plugin template usage: Use the plugin template provided by Napari to create a new plugin project.
 - ii. Function implementation: Implement the required functions in the plugin, such as data loading, processing, and visualization.
 - iii. Plugin testing: Test the compatibility and stability of the plugin in different environments. **Initially local device test, and then do the test by supporting server connection to realize remote data loading.**
 - iv. Plugin publishing: Publish the plugin to the Napari plugin library for our stakeholders to use.
- b. Required tools:
 - i. Napari plugin template: Accelerate plugin development.
 - ii. Python development environment: Used to write and test plugin code.

Development Scope

1. Implement segmentation models (based on SAM).

- a. Implementation steps:
 - i. [Model loading](#): Use sam_model_registry to register and load the pre-trained SAM model.
 - ii. [Image processing](#): Call the SamPredictor.set_image method to convert the input image into an embedding representation that the model can handle. **It is necessary to optimize SAM or find other models to adapt to 3D dimensional data.**
 - iii. [Prompt input](#): Provide prompt information such as points and boxes as needed to guide the model to segment specific areas.
 - iv. [Mask generation](#): Call the predictor.predict method to generate the segmentation mask of the target area.
- b. Tools required:
 - i. Python: Used to write and run code.
 - ii. PyTorch: Used to load and operate deep learning models.
 - iii. [OpenCV](#): Used for image processing and preprocessing.

2. Develop UI interfaces to support annotation optimization.

- a. Implementation steps:
 - i. **Create plugin template: Initialize UI using Napari plugin API.**
 - ii. **Implement interactive features: Enhance Napari interactive experience, such as dynamic data navigation, multi-channel visualization adjustment. Design UI to support image loading, segmentation and interactive operations.**
 - iii. **Test and debug: Test UI stability in different systems.**
- b. Required tools:
 - i. Napari: A platform for image visualization and plugin development.
 - ii. Python: Used to write plugin code.
 - iii. [Related plugins](#): Such as napari-plugin-manager, which is used to manage and install other plugins.

3. Support dataset export (for subsequent analysis).

- a. Implementation steps:
 - i. Data format selection: Determine the format of the exported data, such as TIFF, PNG or custom format. **Additional support for TIFF, PNG or other common formats avoids relying solely on .npy.**
 - ii. Export function implementation: Add data export function in the plug-in to allow users to save processed images or annotation results locally.
 - iii. Metadata preservation: Ensure that relevant metadata information (such as annotation time, operator, etc.) is also saved during the export process to facilitate subsequent analysis.
 - iv. Compatibility testing: Verify whether the exported data can be correctly read and processed by other commonly used analysis tools.
- b. Required tools:
 - i. Napari: Used for image processing and export function implementation.
 - ii. Image processing library: Such as PIL or OpenCV, used for image format conversion and processing.

4. Possible future extensions:

- a. Fine-tune SAM to specific data
- b. Support multi-category annotation
- c. Integrate other deep learning visual models

Expected deliverables

1. Python code base (implements interactive annotation tools).
 - a. **The code should be modularized to facilitate subsequent maintenance and expansion.**
2. Technical report (shows the development process and example results).
 - a. **Need to show the performance comparison results of the tools on different datasets.**
3. README document (guides users on how to use the tool).
 - a. **Need to include examples of using the plugin and provide tutorials for multi-channel data processing.**

Required technologies

1. Deep learning
 - a. **SAM performs automatic segmentation, optimization, or finds new models for 3D adaptation.**
2. Python & libraries List (Based on Jay CHEN's demo)
Basic scientific computing library
numpy==1.26.4
scipy==1.14.1
pandas==2.2.2
matplotlib==3.9.2

```
seaborn==0.13.2
```

```
# Machine learning & deep learning
```

```
scikit-learn==1.5.1
```

```
torch
```

```
torchvision
```

```
torchaudio
```

```
tensorflow-intel
```

```
tensorboard==2.17.1
```

```
# FastAPI / Flask API service
```

```
fastapi==0.115.9
```

```
Flask==3.1.0
```

```
uvicorn==0.34.0
```

```
# Dependent tools
```

```
requests==2.31.0
```

```
tqdm==4.66.5
```

```
protobuf==4.25.3
```

```
pydantic==2.10.6
```

```
typing_extensions==4.12.2
```

```
PyYAML==6.0.2
```

```
# Other important tools
```

```
joblib==1.4.2
```

```
six==1.16.0
```

```
tabulate==0.9.0
```

3. Data visualization

- a. Implementation of interactive annotation, scrolling and loop browsing.

Fields involved

1. Artificial intelligence
2. Data science/analysis
3. Bioinformatics/biomedicine

Conclusion

The core of this project is to use **deep learning** and **interactive visualization tools** in the field of **biomedical image analysis** to create an efficient, intuitive, and adjustable cell segmentation and annotation platform suitable for researchers or biomedical engineers.

Appendix (Main key points and updates from 1st stakeholders meeting)

✓ Multi-channel image support (optimized multi-channel display of RGB and medical images)

- ✓ 3D data processing issues (SAM cannot process 3D efficiently at present, needs improvement)
- ✓ Improved data storage format (support PNG and other easy-to-read formats, avoid using only NPY)
- ✓ Scrolling & looping (optimize user interaction and improve data loading experience)
- ✓ Optimize interactive annotation tools (reduce manual drawing time, support more annotation methods, such as points, boxes, automatic segmentation, etc.)
- ✓ Real-time model fine-tuning (currently the tool lacks this function, needs improvement to improve model flexibility)
- ✓ Richer export options (ensure that the exported data can be easily used by non-programmers)
- ✓ Support server connection (improve the feasibility of remote data loading)