Reusing existing trained AI models in WIPP

Antoine Meyer, Mylene Simon, Timothee Kheyrkhah, Peter Bajcsy

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Overview

Problem

- Lower the barriers for reusing trained AI models available in AI models repositories

Challenges

- Models are spread on lot of different repositories and formats
- For non-technical researchs, not trivial to setup and execute AI models
- Re-train models from scratch take lot of time and money

Motivations

- Accelerate application of AI models to scientific problems
- Save hours of computation time by reusing pre-trained models from diverse repositories

Already existing APIs or tools to download, inference (and train) AI models

- Transformers (Hugging Face API)*
- Cellpose API**
- BioImage.IO Core (Python libraries)
- SAM2 repository***

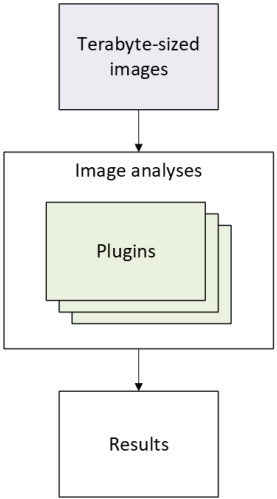
*https://aclanthology.org/2020.emnlp-demos.6/
**https://www.biorxiv.org/content/10.1101/2020.02.02.931238v1
***https://arxiv.org/abs/2408.00714

Our approach

- Leverage WIPP
- Use AI model card
- Using existing API of AI repositories

Web Image Processing Pipelines (WIPP)

- Purposes
  - Measurements based on terabyte-sized images
  - Algorithmic plugin platform
- Goal
  - Lower the bar to execute image analyses

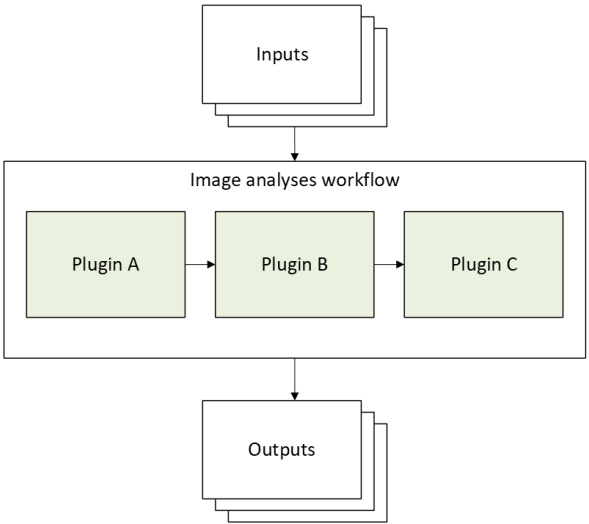


WIPP workflow

- Sequence of plugins

WIPP plugin

- Piece of code taking inputs/outputs and executing code



Goal: Automatic documentation for AI models trained inside WIPP

Work: Make AI model card proposal and development for integration

```
1       public class AiModelCard {
2         private String                version;
3         private String                name;
4         private Date                  creationDate;
5         private String                framework;
6         private Map<String, String> trainingData;
7         private Map<String, String> trainingParameters;
8         [8 additional fields]
9       }
```

Feedback: Not relevant, AI users test directly, they don't read the documentation

2. WIPP plugins for AI models reuse
   - Public AI repositories
   - Unlock these repositories
   - Evaluate results

Many public AI models on lots of public AI repositories

AI repositories	Image classification models	Segmentation + MG* models
Hugging Face	15,593	1,160 + 176
BioImage.IO	1	4 + 32
Cellpose	×	21
SAM2	×	8
PyTorch Hub	20	5

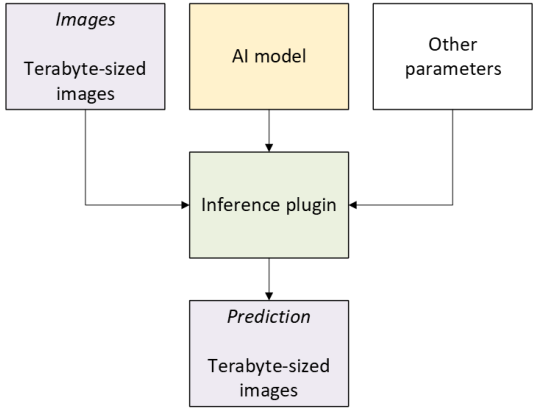
Table: Number of models per repository

*MG: Mask Generation

Goal: Access external AI models in WIPP

Work:
Plugins for Hugging Face, BioImage.IO, Cellpose and more public repositories

Question:
How to assess the relevance of results?



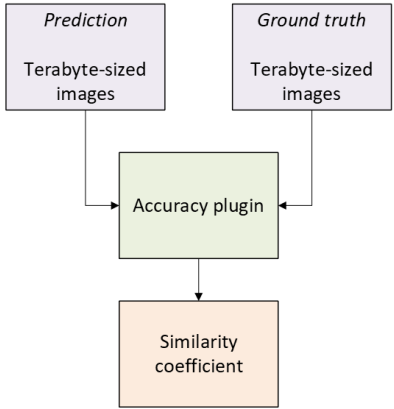
Goals

- Mesure accuracy of external AI results
- Select the most accurate/fastest one

Work:
Plugin to compute the Dice-Sørensen coefficient*

$$Dice = \frac{2 * TP}{2 * TP + FP + FN}$$

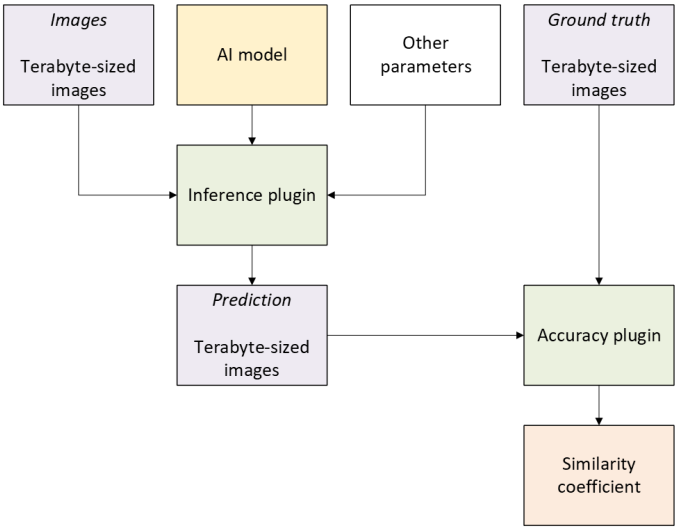
*Statistic used to gauge the similarity of two samples



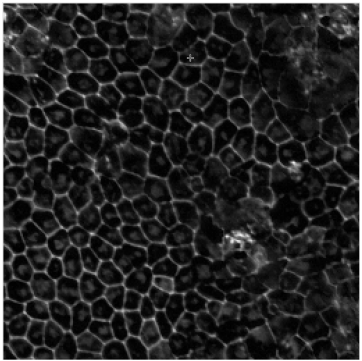
3 Benchmarks
- 2-steps workflow
- Data 'cell boundary'
- Accuracy for 'cell boundary'
- Data 'nuclei segmentation'
- Accuracy for 'nuclei segmentation'

WIPP 2-steps workflow

1. Inference the model
2. Compute the accuracy

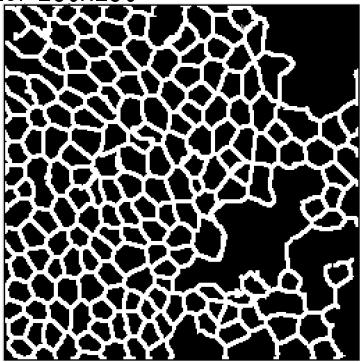


Name: Retinal Pigment Epithelium
Number: 1032 images (822/210)



Image

Type: Cell microscopy
Size: 256x256



Mask

Source: https://doi.org/doi:10.18434/T4/1503229

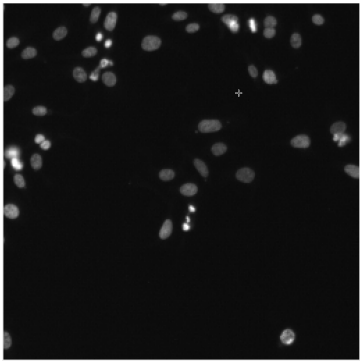
Compute time: Few minutes on WIPP
Task: Segments cell edges

Repository	Model	Accuracy
WIPP	unet-cnn*	95.11% ± 0.78%
BioImage.IO	10.5281/zenodo.5869899	89.30% ± 0.84%
Hugging Face	facebook/sam-vit-huge	86.01% ± 2.50%
SAM2	facebook/sam2.1-hiera-large	80.18% ± 5.02%
Cellpose	cyto3	78.51% ± 2.35%

Table: Models accuracy after inference on data 'cell boundary'

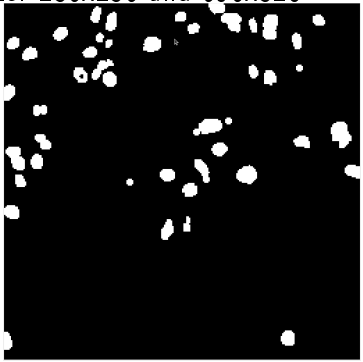
*Trained with data, then inferenced (in WIPP)

Name: Data Science Bowl 2018
Number: 497 images (447/50)



Image

Type: Cell microscopy
Size: 256x256 and 696x520



Mask

Source: https://bbbc.broadinstitute.org/BBBC038/

Compute time: Few minutes on WIPP
Task: Segments nuclei of cells

Repository	Model	Accuracy
BioImage.IO	10.5281/zenodo.5764892	93.73% ± 3.98%
WIPP	Stardist 2D paper DSB 2018	90.67% ± 4.42%
Cellpose	cyto3	82.31% ± 17.25%
Cellpose	nuclei	81.00% ± 21.00%
SAM2	facebook/sam2.1-hiera-small	48.18% ± 32.41%
BioImage.IO	10.5281/zenodo.5869899	29.47% ± 8.32%
Hugging Face	facebook/sam-vit-huge	21.63% ± 15.37%

Table: Models accuracy after inference on data 'nuclei segmentation'

4 Conclusion
- New WIPP plugins
- Fine-tuning plugin

Inference plugins

- wipp/wipp-huggingface-maskgeneration-inference
- wipp/wipp-bioimage-io-inference-plugin
- wipp/wipp-sam2-inference-plugin
- wipp/wipp-cellpose-inference-plugin

Accuracy plugin

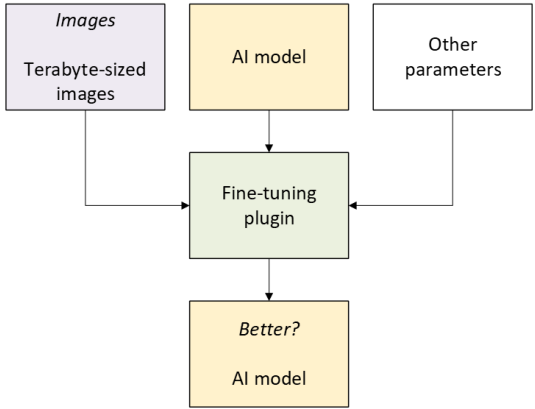
- wipp/wipp-dice-segmentationaccuracy

Goal

- Improve already existing AI models

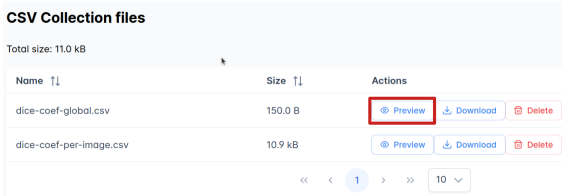
Work

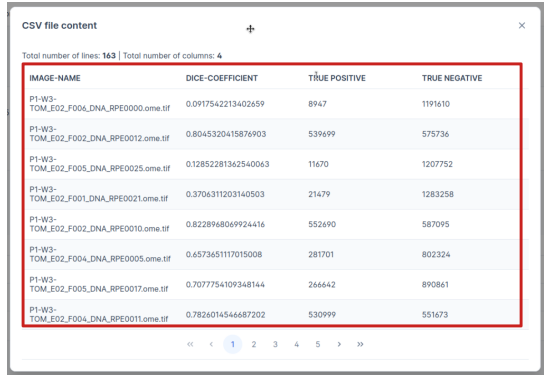
- Identify if models can be retrain/fine-tune
- Plugin to fine-tune an AI model



Thanks !

Issue: Impossible to view CSV content without downloading
Solution: Content directly in the user interface





Issue: Difficult to navigate between different plugins in a workflow
Solution: Allows user to open one modal per plugin