COMPUTER VISION FOR ANALYSIS OF SCANNING ELECTRON MICROSCOPE (SEM) OF INTEGRATED CIRCUITS

Members:
Chieu Le Yang, Ryler Ng Kai Guan
(NUS High School of Mathematics and Science)

Mentor:
Shen Bingquan (DSO National Laboratories)

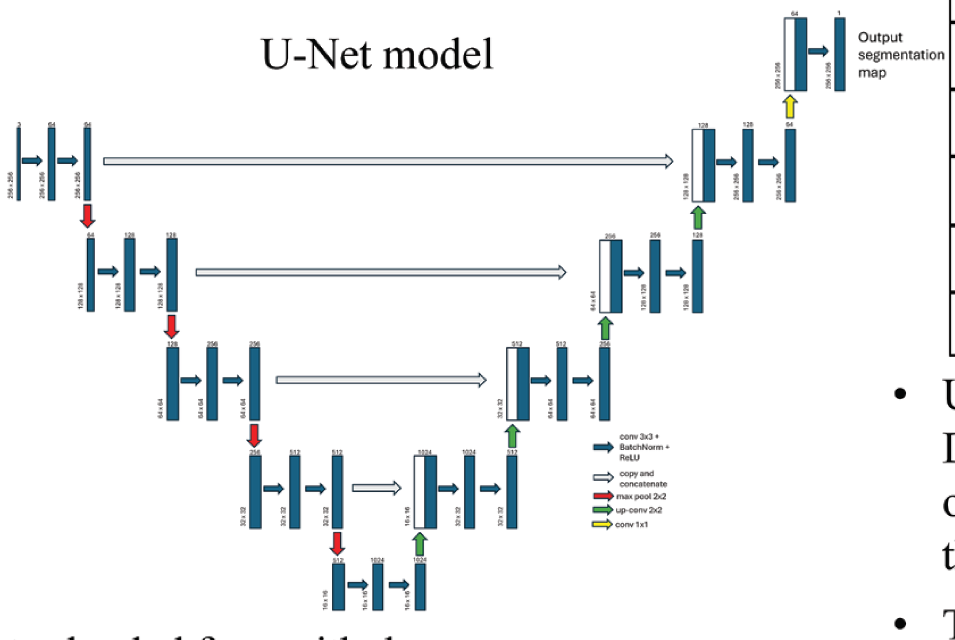
INTRODUCTION

- Integrated circuits (ICs) are essential in technology and national security [3].
- Quality control is critical, yet challenging due to the vast number of ICs produced.
- Solution: This study finetuned these models using images of IC features captured via scanning electron microscopy, optimising parameters like learning rates and data volume. SAM-SEG performed the best with a 94% pixel-wise accuracy, followed by U-NET and RSPrompter lagged slightly. To reduce effort required for labelling, semi-supervised learning was also used. Using 800 labelled and 200 unlabelled images yielded a 91% accuracy, demonstrating the feasibility of reducing labelled data without significant accuracy loss. Key findings highlighted SAM-2's robustness, U-NET's simplicity for small datasets, and semi-supervised learning's potential to balance effort and performance. Future research could explore advanced clustering and larger datasets for further improvements. This work underscores the importance of computer vision in IC quality control and opens pathways for efficient, scalable solutions.

METHODS

- 4 different models were finetuned
    1. SAM-SEG [4]
    2. U-NET [5]
    3. SAM-DET [1][2]
    4. RSPrompter [1][2]
- avr_m1 class of integrated circuits were mainly used

U-Net model



Part 1: Finetuning Models
Loading Models
- U-NET architecture (see right)
- SAM-SEG,SAM-DET and RSPrompter loaded from github

Loading Data
- For SAM-SEG and U-NET: images loaded into tensor format
- For SAM-DET and RSPrompter: images loaded into COCO format
- Prompts obtained from ground truth masks for SAM-SEG model

Finetuning/Training
- Independent Variables: Learning rate/Number of images
- Learning rate: varied from 1e-4 down to 1e-8 by factor of 10 each time
- No of images: increased from 200 to 1000 in intervals of 200
- Constant Variables: Number of epochs was kept at 5

Testing
- 50/50 train-test split used
- No of test images used is the same as no of images used for training
- Pixel-wise accuracy was used to compare all models
- Images were printed out for visualisation purposes

Part 2: Semi-supervised Learning
- SAM-SEG selected for semi-supervised learning
- Varying labelled/unlabelled splits were used, ranging from 700/300 to 1000/0
- Labelled data used to train model first
- Model used to predict pseudo masks for unlabelled data
- Model trained on unlabelled data and pseudo masks

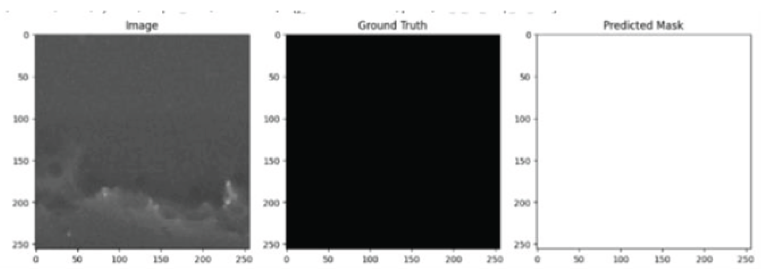
RESULTS & DISCUSSION

- Table 1: Pixel accuracy of models trained via supervised learning on 1000 images

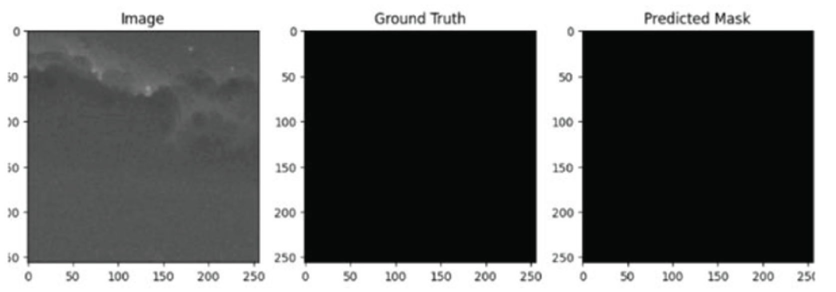
U-NET	SAM-SEG	SAM-DET	RSPrompter
0.9312	0.9407	0.8477	0.9276

- SAM-SEG, U-NET, and RSPrompter achieved high accuracy, but SAM-SEG performed best on challenging images (complex images with ambiguous object-like features not included in the ground truth) compared to the other models. U-NET's simpler architecture and insufficient training for RSPrompter likely caused their limitations when dealing with such challenging images. SAM-DET had lower accuracy due to overfilled segmentation masks.

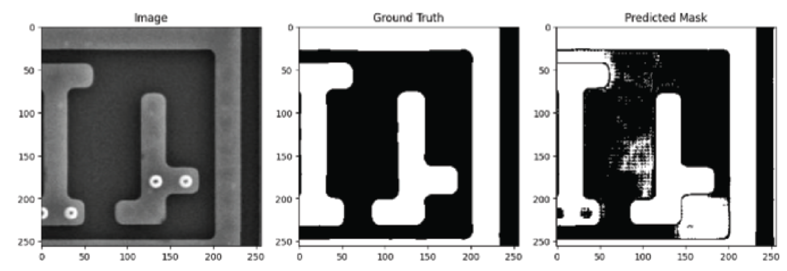
RESULTS & DISCUSSION (Continued)



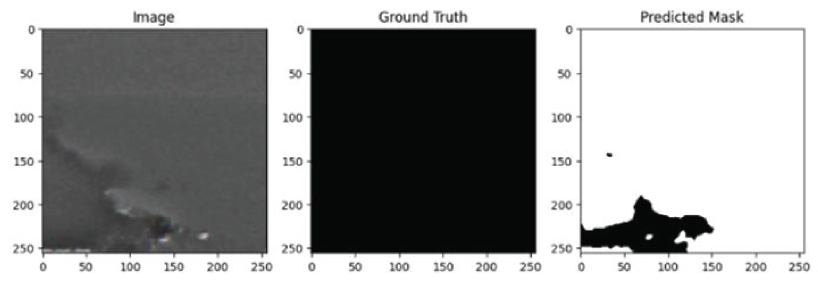
A prediction by RSPrompter



A prediction by SAM-SEG



A prediction by SAM-DET



A prediction by U-NET

- Table 2: Effect of learning rate on pixel-wise accuracy. 1000 images were provided and 5 epochs were run

Learning rate	U-NET	SAM-SEG	SAM-DET	RSPrompter
1e-4	0.8814	0.9407	0.7259	N/A
1e-5	0.9312	0.8562	0.6096	0.8166
1e-6	0.9060	0.6231	0.3871	0.6537
1e-7	0.6609	0.4682	0.4321	0.6405
1e-8	0.6291	0.4679	0.4443	0.5743

- U-NET and SAM-SEG performed best at 1e-5 and 1e-4, respectively, while SAM-DET and RSPrompter improved with higher learning rates, suggesting that their optimal rates exceed 1e-4. However, higher learning rates would cause crashes and thus could not be used.
- Table 3: Effect of the amount of training images provided on pixel-wise accuracy.

Number of images provided	U-NET	SAM-SEG	SAM-DET	RSPrompter
200	0.9119	0.7103	0.3799	0.6996
400	0.9050	0.8173	0.6206	0.8875
600	0.9281	0.8219	0.5735	0.8865
800	0.9454	0.8523	0.5940	0.8583
1000	0.9397	0.8562	0.6096	0.7946

- SAM-SEG accuracy increased with more data, plateauing at ~85%. U-NET performed well even with smaller datasets, while RSPrompter and SAM-DET occasionally performed better with 400 images, possibly due to dataset similarity.
- Table 4: Effect of ratio of labelled to unlabelled data used for training on accuracy.

Labelled	Unlabelled	SAM-SEG
700	300	0.9011
800	200	0.9199
900	100	0.8888
1000	0	0.9407

Using SAM-SEG, semi-supervised training with 800 labelled and 200 unlabelled images achieved 91% accuracy, highlighting its potential to reduce annotation effort. SAM-SEG still performs quite well despite requiring less labelled images (images with provided ground truth masks for the model to compare against).

CONCLUSION

In conclusion, we fine-tuned four models (U-ET, SAM-SEG, SAM-DET, and RSPrompter) under various conditions, achieving 94% accuracy with SAM-SEG. Semi-supervised learning proved effective, reducing the need for labelled data and saving effort. Future work could explore the impact of more labelled and unlabelled data and the use of clustering to select representative training sets.

References

1. Chen, K. (n.d.). RSPrompter: Learning to prompt for remote sensing instance segmentation. GitHub. Retrieved December 22, 2024, from https://github.com/KyanChen/RSPrompter
2. Chen, K., Liu, C., Chen, H., Zhang, H., Li, W., Zou, Z., & Shi, Z. (2023). RSPrompter: Learning to prompt for remote sensing instance segmentation based on visual foundation model. arXiv. https://arxiv.org/pdf/2306.16269
3. Fortune Business Insights. (n.d.). Integrated circuit market size, share and trends. Retrieved December 22, 2024, from https://www.fortunebusinessinsights.com/amp/integrated-circuit-market-106522
4.Sagieppel. (n.d.). Fine-tune & train segment anything 2 in 60 lines of code. GitHub. Retrieved December 22, 2024, from https://github.com/sagieppel/fine-tune-train_segment_anything_2_in_60_lines_of_code
5. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-NET: Convolutional networks for biomedical image segmentation. arXiv. https://arxiv.org/pdf/1505.04597