Student: Ayman Asaria, *iBSc Mathematics, Computers and Medicine*
Supervisors: Professor Marinko Sarunic

I would like to recognise that all tables and figures included in this study are original creations. Special appreciation goes to Aleksandra Goch for her generous assistance in recording the videos during retinal surgeries. Code is available from the following public GitHub repository: https://github.com/aa300503/Retinal-Microsurgery-/tree/main

Enhancing Robotic Ophthalmic Surgery: A Machine Learning Perspective on Tool Tracking

INTRODUCTION

- Retinal microsurgery presents distinct challenges due to its demand for **utmost surgical precision.**
- Outcomes are influenced by the surgeon's experience alongside **iatrogenic factors** such as innate tremor and psychological state.
- Robotic systems offer some promise in overcoming human limitations, through **dampening of tremor- but are limited by their requirement of an unobstructed surgical view.**

Research Problem

Current imaging modalities rely heavily on maintaining a clear surgical field, hindered by uncontrollable factors such as tool occlusion, haemorrhages.

The impact on patients

A degraded intraoperative view increases the risk of unwanted tool-tissue interactions, leading to iatrogenic retinal damage

Our Aim

To build an automated solution for semantic segmentation through supervised learning, with the primary objective of tracking the position of the instrument across the surgical field.

METHODS

Data Collection → **Frame Extraction & Annotation** → **Data Augmentation** → **Build U-Net Model** → **Create Ground Truth Masks** → **Tuning hyperparameters** → **Final Evaluation**

Nine retinal surgeries were recorded using a Proveo 8 microscope.



The VGG Video Annotator was used to mark free bounding boxes around the instrument tooltip depicted in every frame of the surgical footage. Free bounding boxes were used to encompass the entirety of the instrument tooltip, ensuring its precise localisation across varying bearings within the frames. A JSON file was generated with the X and Y coordinates for each bounding box. A ground truth mask was where a value of 1 was assigned to the pixels enclosed by the bounding box encompassing the tool and a value of 0 to all other pixels within a matrix.

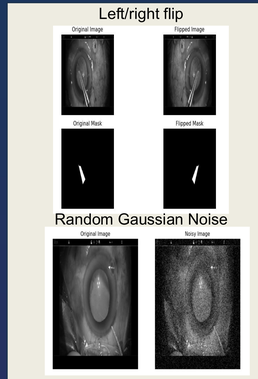


Figure 1 | Data augmentation techniques employed and their corresponding effects on the ground truth binary segmentation mask and the image.

- 80% data used for algorithm training.
- 20% data held out for validation.
- A baseline U-Net model was trained for 30 epochs using the default activation function.

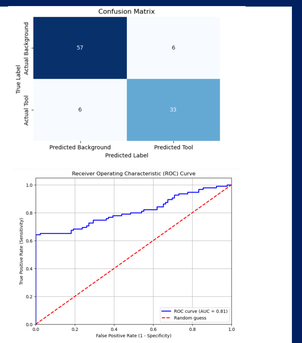
A comparison is performed using eighteen significant U-Net variants on the same annotated dataset to obtain an understanding of the effect of each hyperparameter on segmentation performance.

Hyperparameters tuned:
1. Epoch Number
2. Activation Function
3. Class Weighting

Evaluation of the 18 tuned models on the held-out validation set. Evaluation metrics: **Mean IOU, AUROC curve, Sensitivity, Specificity.**

RESULTS

Figure 2 | ROC Curve and confusion matrices for the original and final, tuned model. Final, tuned model outperforms baseline model, where the ROC curve illustrates the trade-off between true positive rate (sensitivity) and false positive rate (1-specificity) as discrimination threshold values. The confusion matrix shows the ratio of true and false positives and negatives, allowing us to evaluate segmentation performance.



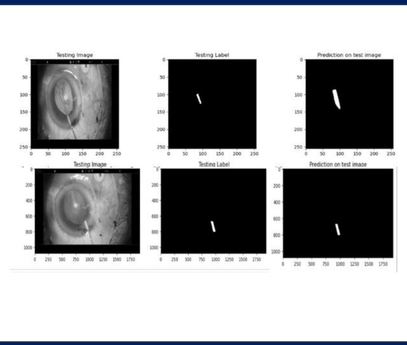


Figure 3 | Visual analysis of results binary ground truth and predicted segmentation mask shown for baseline model (top row) and final, tuned model (bottom row). IOU score for the final, tuned model was 84%.

CONCLUSIONS

Limitations → Future directions

Single Centre Dataset
Limits **generalisability** and introduces **cross-over bias**

→ **Single Centre Dataset**
Multi-institution dataset
Benefits of widening scope: account for variations in surgical practice and techniques.

All images contained instruments
Model's performance may be compromised when tested with images lacking them.

→ **Diversify dataset**
Insert images lacking surgical instruments to enhance the model's applicability to surgical scenarios where the instrument may enter and exit the frame

Summary

- The integration of robotics into this field holds great promise in surmounting human limitations but is challenged by the impaired intraoperative visualisation and the need for manual repositioning in the area of interest.
- We propose an efficient ML approach for supervised segmentation of the tooltip in RM, achieving an IOU of 84%.
- Tooltip tracking has the potential to provide real-time surgical guidance, facilitating the incorporation of robotics into RM and thus enhancing the safety of retinal surgeries.
- Moving forward, our focus will be on refining the model and conducting further training on a more diverse dataset.