



École Polytechnique Fédérale de Lausanne

---

Classification of cancerous tissue  
from fluorescence lifetime microscopy acquired data  
using deep neural networks

---

by Petr Šícho

Semester Project Report

Moser Christophe  
Project Advisor

Victoria Fay  
Project Supervisor

July 30, 2026

# Acknowledgments

I would like to sincerely thank my supervisor, Victoria Fay, for her guidance, support, and valuable feedback throughout this whole semester project. Her timely responses, inspiring insights and expertise have been instrumental in the progress of this work.

I am also grateful to EPFL, LAPD lab and professor Moser for providing the opportunity and resources to work on this interesting project. Special thanks to the entire team that contributed to this project, including the lab technicians, doctors, surgeons, and other professionals who generously shared their expertise and support.

Finally, I would like to thank my colleagues, friends and family for their encouragement and support during this journey.

*Lausanne, July 30, 2026*

Petr Šícho

# Abstract

Head and Neck Cancer (HNC) affects about one million people worldwide each year. Treatment typically involves surgical resection, deemed complete if histopathology shows clear margins. With a point-of-care method to map tumor margins in real time, we hope to improve tumor removal, increasing life expectancy of HNC patients.

This project focuses on the development of a machine learning algorithm to classify tumor cells based on their endogenous fluorescence lifetime response. Tumor cells exhibit distinct metabolic behaviors compared to healthy cells, resulting in unique fluorescence signatures. The algorithm leverages these specific features to accurately differentiate between cell types, utilizing advanced image analysis techniques.

We used Fluorescence-lifetime imaging microscopy (FLIM) to capture lifetime data from sections of human lymph nodes containing both cancer and healthy cells. Using the gathered data, we then trained a machine learning model to segment new FLIM data to predict whether a given lymph node slice contains cancer cells.[2]

An accuracy of 73% and an Area under curve (AUC) of 0.85 were achieved when testing on previously unseen sections of human lymph nodes, where sections from other parts of the same node were included in the training dataset. These results surpass the accuracies reported in recent related studies. However, the model struggled to generalize to cancer types not represented in the training dataset, achieving an AUC of only 0.55. With access to sufficiently diverse training data, this approach holds potential for adaptation to clinical settings, paving the way for minimally invasive diagnostic tools for margin identification.

# Contents

|                                                        |           |
|--------------------------------------------------------|-----------|
| <b>Acknowledgments</b>                                 | <b>2</b>  |
| <b>Abstract</b>                                        | <b>3</b>  |
| <b>1 Introduction</b>                                  | <b>5</b>  |
| <b>2 Background</b>                                    | <b>6</b>  |
| 2.1 Fluorescence Lifetime Imaging Microscopy . . . . . | 6         |
| 2.2 Machine learning . . . . .                         | 6         |
| <b>3 Data acquisition</b>                              | <b>8</b>  |
| 3.1 Microscopy . . . . .                               | 8         |
| 3.2 Immunostaining . . . . .                           | 9         |
| 3.3 Post processing . . . . .                          | 10        |
| 3.4 Data gathered . . . . .                            | 11        |
| <b>4 Machine learning</b>                              | <b>12</b> |
| 4.1 Data augmentation and processing . . . . .         | 12        |
| 4.2 U-net model . . . . .                              | 13        |
| 4.3 Model evaluation . . . . .                         | 13        |
| 4.4 Model performance metrics . . . . .                | 14        |
| 4.5 Testing various models . . . . .                   | 15        |
| <b>5 Results</b>                                       | <b>16</b> |
| 5.1 Evaluation . . . . .                               | 16        |
| <b>6 Conclusion</b>                                    | <b>20</b> |
| <b>Bibliography</b>                                    | <b>21</b> |
| <b>A Documentation</b>                                 | <b>22</b> |
| A.1 FLIM microscope automation . . . . .               | 22        |

# Introduction

Head and Neck Cancer (HNC) affects about one million people worldwide each year. Treatment typically involves surgical resection, deemed complete if histopathology shows clear (negative) margins. Positive surgical margins indicate residual cancer cells, requiring a second resection to prevent local and distant recurrences, which can lead to patient death. Positive margins increase the risk of local recurrence by 90% and significantly reduce survival expectancy. [4]

Achieving tumor excision with negative margins is the goal in oncology surgery. However, up to 20% of patients with negative intraoperative margins still have positive final margins, due to issues like frozen section sampling errors, analytical mistakes, or resection challenges during reconstruction.

There is a need for a point-of-care method to map tumor margins in 3D and detect residual malignant cells in real time.

The tool we design is a needle-like endoscope utilizing the transmission matrix method for deep imaging with diffraction-limited resolution. The contrast mechanism relies on the endogenous fluorescence of tumor cells, which differ in metabolism from normal cells, affecting their fluorescent lifetime response [6]

Using machine learning (ML) we are able to devise an algorithm that, based on training data gathered beforehand, quickly classify new and previously unseen data from patients. This can help surgeons in making real-time decisions on an imaging modality that the surgeon is not used to seeing. ML algorithms are already established in the medical field for biomedical image classification and are used with success for cancer classification. [2]

# Background

## 2.1 Fluorescence Lifetime Imaging Microscopy

Fluorescence Lifetime Imaging Microscopy (FLIM) is a technique that measures how long molecules remain in an excited state before emitting photons, after being excited by a brief flash of light. When certain molecules in cells absorb light, they enter an "excited state" and release the energy as fluorescence. Return of a specific molecule to a ground state is a stochastic process, for an ensemble of cells with same lifetime, we can measure the decay curve of the emitted fluorescence, which follows an exponential

$$I(t) = I_0 e^{-t/\tau}$$

where  $I_0$  is the initial fluorescence and  $\tau$  represents the average fluorescence lifetime.

Unlike traditional fluorescence imaging, which measures the brightness of the light, FLIM focuses on the timing, which is unaffected by the concentration of molecules or external factors like light intensity and thus appear to be a better indicator for determining properties of the cells. [6]

Cancer cells exhibit altered metabolism, favoring glycolysis over oxidative phosphorylation even in oxygen-rich conditions, a phenomenon known as the "Warburg effect". This metabolic shift supports rapid growth but also changes the biochemical environment, particularly the states of molecules like NADH involved in energy production.[3] This metabolic fingerprint and other biochemical fluorescence markers can be used in machine learning (ML) models to distinguish normal and cancerous tissue. [2]

## 2.2 Machine learning

Machine learning (ML) can be used to find intricate patterns in data, which can then be used to classify new data. This requires a large amount of data inputs together by human-annotated ground truth. The ML algorithm then finds a transformation that produces an output from input images, that best matches the input-output pairs provided. This transformation can then be applied to new data, that has not been seen by the algorithm before, to infer an unknown output.

We are particularly interested in ML-based segmentation using convolutional neural networks (CNN). We first contract the input data using convolutional, non-linear activation function (RELU) and max-pooling. In each layer, we reduce the size of features captured. An example of U-net CNN can be seen figure 4.2.

The CNN will learn to recognize visual patterns of various sizes in the input data, which can be used to classify the input into distinct categories. In the end, using the small-size features we use convolutions again to upscale the results to the original resolution of the image. This model design is usually referred to as U-net and was used with success for biomedical images segmentation [7] Using sigmoid activation function in the end we predict one scalar value between 0 and 1 for each pixel in the image. This number should reflect the probability that a specific pixel contains cancer or healthy tissue in the original image. The resulting probabilities are called predicted mask for the image.

# Data acquisition

## 3.1 Microscopy

FLIM images were captured using Leica SP8 FLIM Microscope with hybrid photon counting detectors. The cells were excited using a 440 nm pico-second pulsed laser with 40 MHz repetition rate. We detected lifetime and intensity images in two wavebands, specifically between 460nm and 486nm (Waveband 1, W1), and then 517nm to 587nm (Waveband 2, W2). These wavebands correspond to the emission bands of metabolism cofactors NADH and FAD [6] A HC FLUOTAR 25x objective with water medium between the lens and slide was used. This yields us four values for each pixel of the image.

Firstly, the whole sample was imaged with optical microscopy to identify its overall size and orientation. Using prior knowledge about the specific sample, we selected regions of interest (ROI) containing both normal and cancer cells. Normal and stromal cells neighboring the margin of tumor bulks were of particular interest to us as we hope to capture and later identify the borders of the bulk of cancer as precisely as possible. The entirety of the tissue is then divided using a square grid with tiles of  $465 \times 465 \mu\text{m}$  containing  $512 \times 512$  pixels, with 10 percent of overlap between the tiles. Visualization of the grid is provided in figure 3.1.

In previous work [8], tile selection and merging was done manually, leading to imprecision and errors in setting the microscope. To mitigate this, we designed a python-based script to automate the entire image acquisition process on the LEICA SP8 FLIM microscope. The script takes the ROI as an input and automatically scans either a single tile or area of  $3 \times 3$  or  $5 \times 5$  tiles on both wavebands. Adjusting of focus of the microscope has to be done by manually, as this requires manually adjusting the z-axis on the rig every 1 – 2 mm (corresponding to 3 – 5 tiles) in any direction, to ensure the whole image is in focus. More information and documentation on how to use the FLIM automation is available in Appendix A.

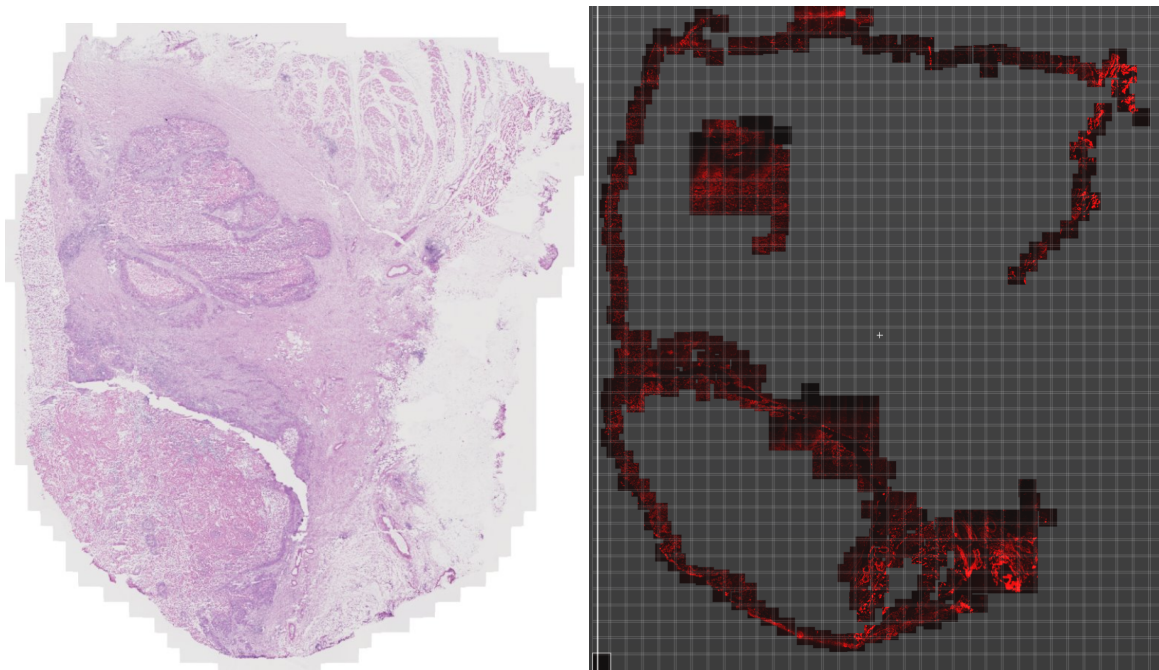


Figure 3.1: Left: H&E stain of a cancer positive lymph node. Right: Subsequent section of the corresponding lymph node, imaged purely in intensity on the FLIM microscope. The ROI of tissue edges allows a 1-1 mapping of the tissue on the immunostaining image for selection of area for imaging.

### 3.2 Immunostaining

Immunohistochemistry (IHC), a form of immunostaining, is a technique in biomedical research and clinical diagnostics for detecting specific antigens within tissue sections. By employing antigen-specific antibodies conjugated to enzymes or fluorescent molecules, IHC enables the visualization of protein location within complex tissue structures. This method is valuable for us as it will help us to precisely identify regions with cancer and its borders.

After imaging on the microscope, the sample was immunostained at AGORA cancer research center, Lausanne. A DAPI staining, binding to cell nuclei, and E-cad A555 (alexa fluor 555) staining marked potential cancer regions for us. Biologists provided additional annotations to create a binary mask of cancerous areas, as E-Cadherin stains both normal and tumoral epithelial cells non-specifically. An example results from the immunostaining is visible in figure 3.2. The mask was post processed by filling in areas where a region of continuous cancer was identified, while areas determined to be artifacts or unspecific binding were removed from the mask.

After the mask creation, it was necessary to match its position to our captured images. To achieve this, a visualization of the captured lifetimes was created and all the images were reconstructed

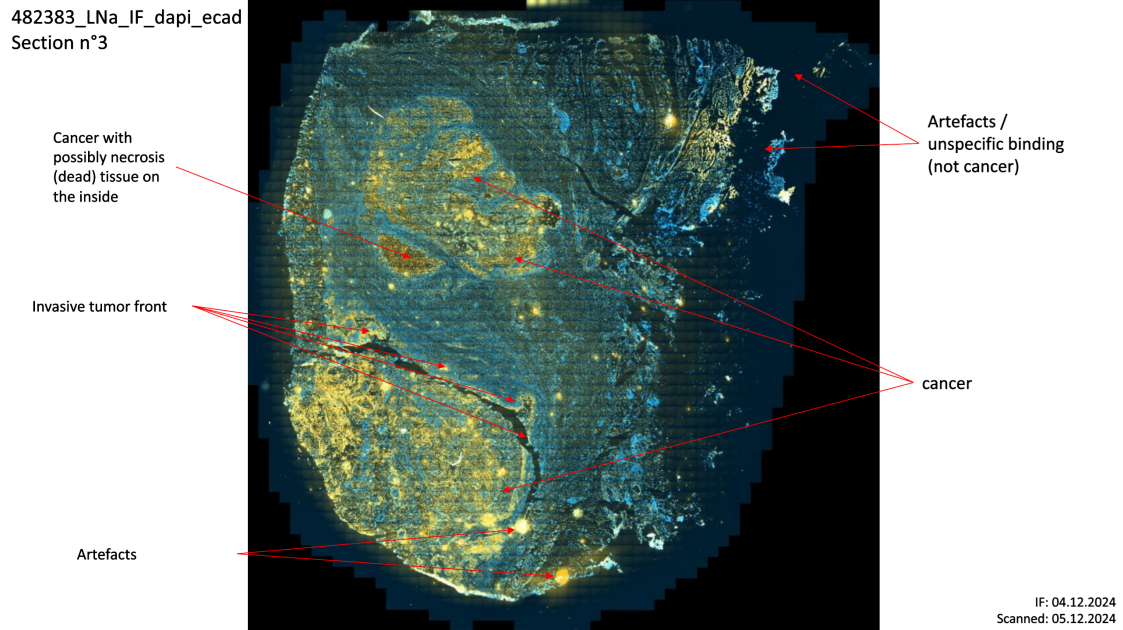


Figure 3.2: Example of immunostaining of a sample with annotations. Blue: DAPI. Yellow: E-Cadherin.

in a grid. Then by matching distinctive features visible on both the immunostaining and in our visualization, such as ridges and ducts, we defined cropping and rotation of the mask to match with the FLIM images. A matched mask-visualization pair is displayed in figure 3.3.

### 3.3 Post processing

To account for differences in excitation across various parts of the tissue, such as variations in focus, laser power, and tissue thickness, we combined the intensity images from two wavelengths into a single channel. This was done by dividing the intensity at W1 by the intensity at W2. By using this approach, we aimed to preserve the relative differences in fluorescence intensity while eliminating any absolute variations that could occur between images. Because the intensity at W2 was almost always higher than at W1, the resulting values typically fell within the range of 0 to 1.

Lifetime however, does not depend on excitation variabilities. Thus, we did not process this data. Typical values present in our samples ranged from 1 to 10 ns. For ML training, it is preferential that all the input values lie in the same range. It improves the convergence of gradient descent helps the model to treat all data with the same importance. To achieve this all lifetime values were divided by the number 10, which brought the majority of them to the 0 – 1 range, the same as for the intensity channel.

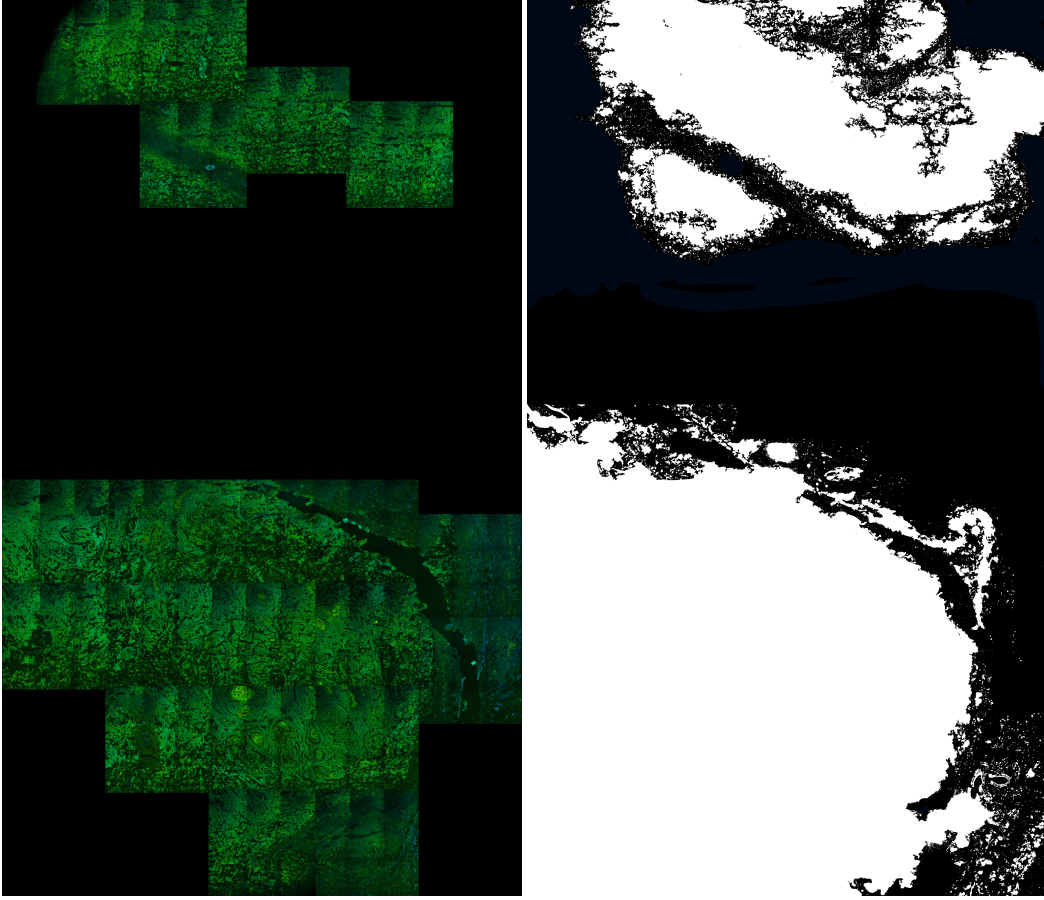


Figure 3.3: Left: A visualization of FLIM images capturing part of lymph node 492383\_LNa. Right: A corresponding mask for the same lymph node mask, with cancer pixels in white.

In rare cases where the photon count was close to 0, lifetime values could not be determined. Assuming that tissues are spatially homogeneous in small areas, the missing values were replaced with the mean of their neighbors. Potential errors in learning arising from these uncertain lifetime values are negligible, as this issue occurred in fewer than 0.1% of the pixels.

### 3.4 Data gathered

For the imaging we used five lymph nodes of different patients, three of them with cancer and two were normal. We used three slices from one of the lymph nodes and one from the rest, so in total seven tissue slides were imaged on the FLIM. The samples were selected to include enough diversity in order to minimize the risk of the algorithm misidentifying healthy tissues as cancer or vice versa. In total we gathered 815 images of  $465 \times 465 \mu\text{m}$  at different positions for each of the two wavelengths.

# Machine learning

## 4.1 Data augmentation and processing

We start by cropping the microscope images in regions of 64x64 pixels to mimic the field of view of the endoscope. This caused some artifacts in the predicted mask, as the borders of these subareas were visible due to sudden change of predicted cancer probability. An example of this is visible in figure 4.1. To mitigate this and also augment the size of the dataset, we added a second cropping step shifted by 32 pixels from the origin. While this did not remove all the artifacts from the mask, it seemed to help with the performance of the classification, improving the test performance by a 5% in our first run of the U-net model.

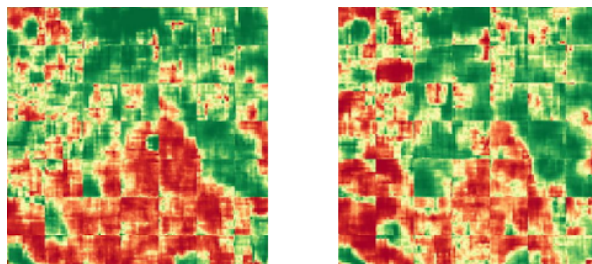


Figure 4.1: Artifacts in the predicted mask. Left: An original image. Right: The same area predicted after training with additional cropping.

The dataset contained more normal pixels than cancerous. It is important that the number of pixels in both masks is comparable, so the neural net does not arbitrarily prefer one label. To balance it, we applied selective data augmentation to regions containing cancer. Originally, the number of tiles in the dataset with predominantly cancer was less than 30%. These tiles were flipped and/or rotated by 90°, so in the end cancer was present in 69% of the input images and made up about half of the pixels.

One sample of lymph node was selected as a testing sample and was not used at all during training and model selection. The rest of the gathered images were randomly shuffled and 90% of them were used for training, while the rest was kept apart for validation.

## 4.2 U-net model

The U-net architecture was selected as a first choice for CNN model to train. It is a model for image segmentation, widely used for biomedical imaging, and consists of a fully convolutional network with a symmetric encoder-decoder structure. The encoder progressively reduces spatial dimensions using convolutional blocks with  $3\times 3$  convolutions, ReLU activations, and  $2\times 2$  max-pooling, doubling the feature channels at each step. The decoder mirrors this structure by up-sampling the feature maps using transposed convolutions or up-convolutions, halving the number of feature channels at each stage. Skip connections link corresponding layers of the encoder and decoder, concatenating feature maps to preserve spatial details lost during down-sampling. This enables precise localization while leveraging deep context from the encoder. The final layer is a  $1\times 1$  convolution to map the feature maps to the desired number of output classes, producing pixel-wise segmentation. U-Net is well-suited for tasks with limited data due to its efficient use of context and localization through the skip connections. [7] Visualization of the model architecture can be seen in figure 4.2.

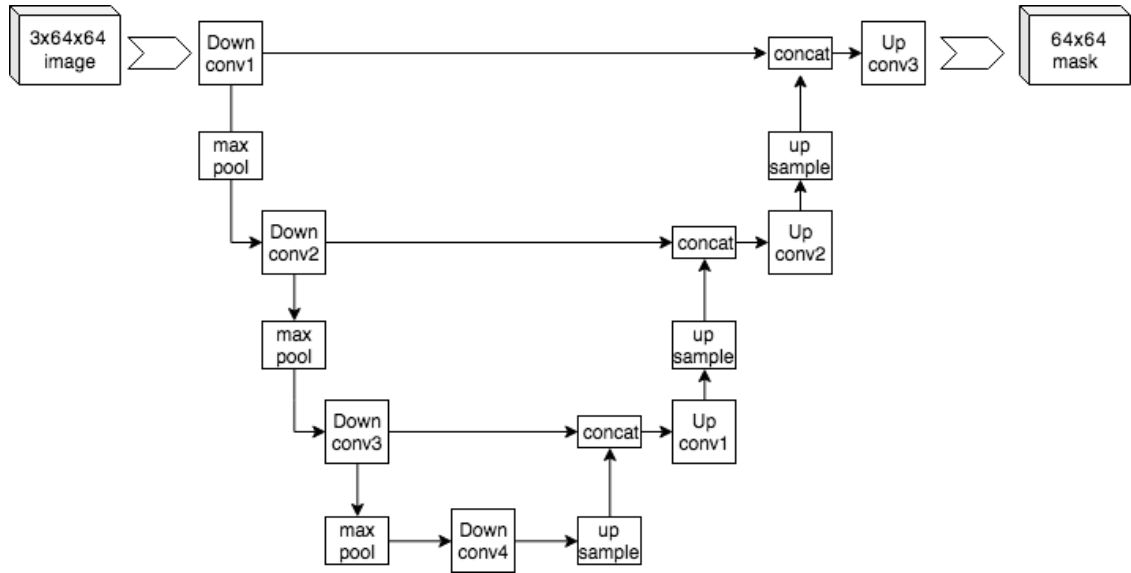


Figure 4.2: U-net CNN for 3 channel 64-by-64 pixel images producing a mask of 64-by-64 pixels. Edited from [1]

## 4.3 Model evaluation

For model evaluation we need to set apart one whole section of a lymph node that would not be used at all during training, to ensure we can gauge the models prediction ability on new data. For that good balance of cancer and normal tiles in the node is preferred.

One of the lymph nodes (157111) seems to be quite different from the other samples we got. Areas with cancer are very dense, with lots of nuclei, unlike the cancer in our other lymph nodes. When this sample was chosen as the test one, the CNN we trained on the rest of the data did not generalize at all, misclassifying most of the cancer as normal cells. Improving the dataset by acquiring images of cancer section with similar features would help the model to generalize to this type of cancer.

We imaged multiple sections from different parts of one lymph node (482381/482383), which contain similar cell structure while being different enough to avoid data repeat. Training on one section and keeping the other for testing allows to verify the classification works on similar cancer types.

The last choice is a lymph node (465666) of which we have only one slice, and spacial features seems to match the other nodes in the training dataset well. In the end we trained all the models twice, choosing either the section 482381 or 465666 for testing. We could judge the models baseline performance on similar cancer types and determine how well it generalizes to other cancer types.

## 4.4 Model performance metrics

In evaluating the performance of segmentation models, various metrics capture distinct aspects of their effectiveness. For our evaluation we chose accuracy, dice coefficient and Area Under the ROC Curve (AUC).

Accuracy measures the proportion of correctly classified pixels relative to the total pixels, considering pixels with cancer probability > 50% as cancer and the rest as normal. While it provides a broad sense of the model's performance, it can be sensitive to class imbalances, leading to skewed view of the model performance. That is why additional metrics are used to better gauge the model. Furthermore, in cancer detection we usually care more about false negatives, e.g. not detecting cancer, so a threshold other than 50% could be more practical.

The Dice coefficient, also known as the Sørensen–Dice similarity coefficient, is a more specialized metric for evaluating the overlap between the predicted segmentation and the ground truth mask. It is calculated as  $Dice = \frac{2 \times TP}{2 \times TP + FP + FN}$ , where TP, FP, and FN are the numbers of true positive, false positive, and false negative pixels, respectively. The Dice coefficient ranges from 0 to 1, where 1 indicates perfect overlap between the predicted and ground truth masks, and 0 indicates no overlap. This metric emphasizes the balance between precision (correctly predicted positives relative to all predicted positives) and recall (correctly predicted positives relative to all actual positives), making it particularly effective in evaluating the accuracy of the predicted shapes and boundaries of segmented regions.

The Area Under the Receiver Operating Characteristic Curve (AUC) provides another perspective

by evaluating the model's ability to distinguish between positive and negative pixels across a range of decision thresholds. The ROC curve plots the true positive rate (sensitivity) against the false positive rate for varying thresholds. AUC is the area under this curve and serves as a measure of the model's discriminatory power. A perfect model achieves an AUC of 1, indicating flawless distinction between positive and negative pixels, while an AUC of 0.5 suggests random guessing. AUC is particularly useful for understanding how the model performs under different threshold settings, and might help us decide on what classification threshold to choose, balancing true positive rate and false positive rate for our specific use case.

## 4.5 Testing various models

As the U-net model gave satisfying results, we designed ten variants of this architecture. The modifications to the basic U-Net model introduce targeted architectural enhancements aimed at optimizing its performance for lymph node cancer classification. Batch Normalization (`basic_bn`) was added to improve convergence and stabilize training, while Dropout (`basic_do`) helps mitigate overfitting by introducing regularization. The `deep_bn` model increases depth by adding an extra downsampling and upsampling block with Batch Normalization to capture finer details, while `shallow_net` reduces the depth for a lighter, faster model. Filter adjustments include `wide_filters`, which doubles the number of filters for potentially higher feature extraction capacity, and `reduced_filters`, which halves the filters to create a lightweight alternative.

Other variants introduce novel architectural elements. `Basic_res` incorporates residual connections to improve gradient flow and address vanishing gradient issues. `Increased_kernel` employs larger convolutional kernels (5x5) to capture broader spatial features, while `dilated_conv` substitutes standard convolutions with dilated ones in intermediate layers to expand the receptive field without increasing the model's depth. Finally, `dropout_augment` combines Dropout and Batch Normalization, aiming to balance regularization with training stability. Together, these variants explore a range of structural adjustments to optimize classification accuracy, computational efficiency, and robustness.

All these models were trained and evaluated in the same way as the basic U-net model.

# Results

## 5.1 Evaluation

The performance of the U-Net variants was evaluated based on their test accuracy, Dice coefficient, and Area Under the ROC Curve (AUC). Table 5.1 summarizes these metrics, providing a detailed comparison of the models. Results are using node 482381 for testing.

Table 5.1: Performance metrics for U-Net model variants.

| Model Variant    | Test Accuracy | Dice Coefficient | ROC AUC |
|------------------|---------------|------------------|---------|
| Simple Model     | 0.73          | 0.73             | 0.85    |
| Simple_bn        | 0.71          | 0.72             | 0.86    |
| Simple_do        | 0.71          | 0.70             | 0.84    |
| Deep_bn          | 0.72          | 0.71             | 0.88    |
| Wide_filters     | 0.71          | 0.72             | 0.86    |
| Reduced_filters  | 0.71          | 0.70             | 0.82    |
| Simple_res       | 0.71          | 0.72             | 0.85    |
| Shallow_net      | 0.64          | 0.67             | 0.80    |
| Increased_kernel | 0.71          | 0.69             | 0.88    |
| Dropout_augment  | 0.69          | 0.71             | 0.87    |
| Dilated_conv     | 0.70          | 0.68             | 0.87    |

The updated results reveal that the baseline simple U-net Model performed well with a Dice coefficient of 0.73 and a ROC AUC of 0.85. The best-performing models in terms of ROC AUC were Deep\_bn and Increased\_kernel, both achieving 0.88. This indicates that increasing model depth and kernel size can effectively enhance the model's ability to distinguish between positive and negative pixel classes.

The Simple\_bn, Wide\_filters, and Dropout\_augment variants all demonstrated competitive performance, with ROC AUC values of 0.86 and 0.87, suggesting the benefits of these architectural adjustments for improving classification confidence. However, the Shallow\_net variant exhibited the lowest performance across all metrics, with a Dice coefficient of 0.67 and ROC AUC of 0.80, highlighting the limitations of reducing model complexity for this task.

ROC curves are displayed below for the simple U-net model, the best-performing model , in

figure 5.1. Predicted mask for the whole area, with comparison to the ground-truth is visible in figure 5.3.

We then tested the model on a more distinct lymph node, 465666. For this type of cancer, the model did not generalize well, achieving an accuracy of approximately 54% across all models and an ROC ranging from 0.50 to 0.55. This suggests that the training data provided to the model was not sufficiently similar to the cancer cells present in this image, though some level of similarity was observed. Interestingly, the model predicted all the borders of blood vessels and lymph ducts as cancer. The E-cadherin staining marks all the ducts as well, and these regions had to be excluded from the cancer mask before training. This observation suggests that the model may have learned to predict E-cadherin staining, identifying distinctive features in the ducts and veins that trigger both the staining and the model's predictions. An example of this phenomenon is illustrated in Figure 5.2.

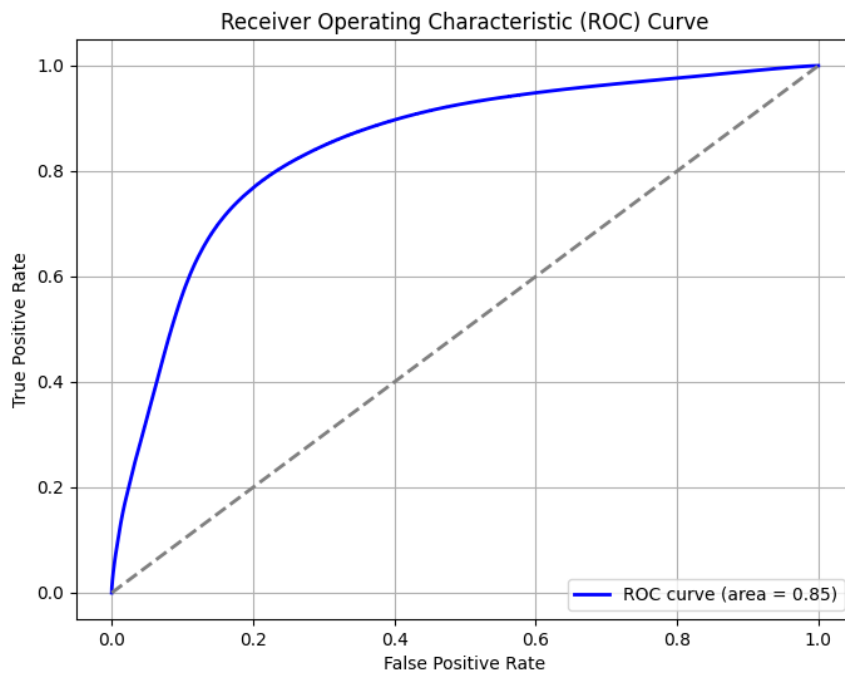


Figure 5.1: ROC for the simple U-net model

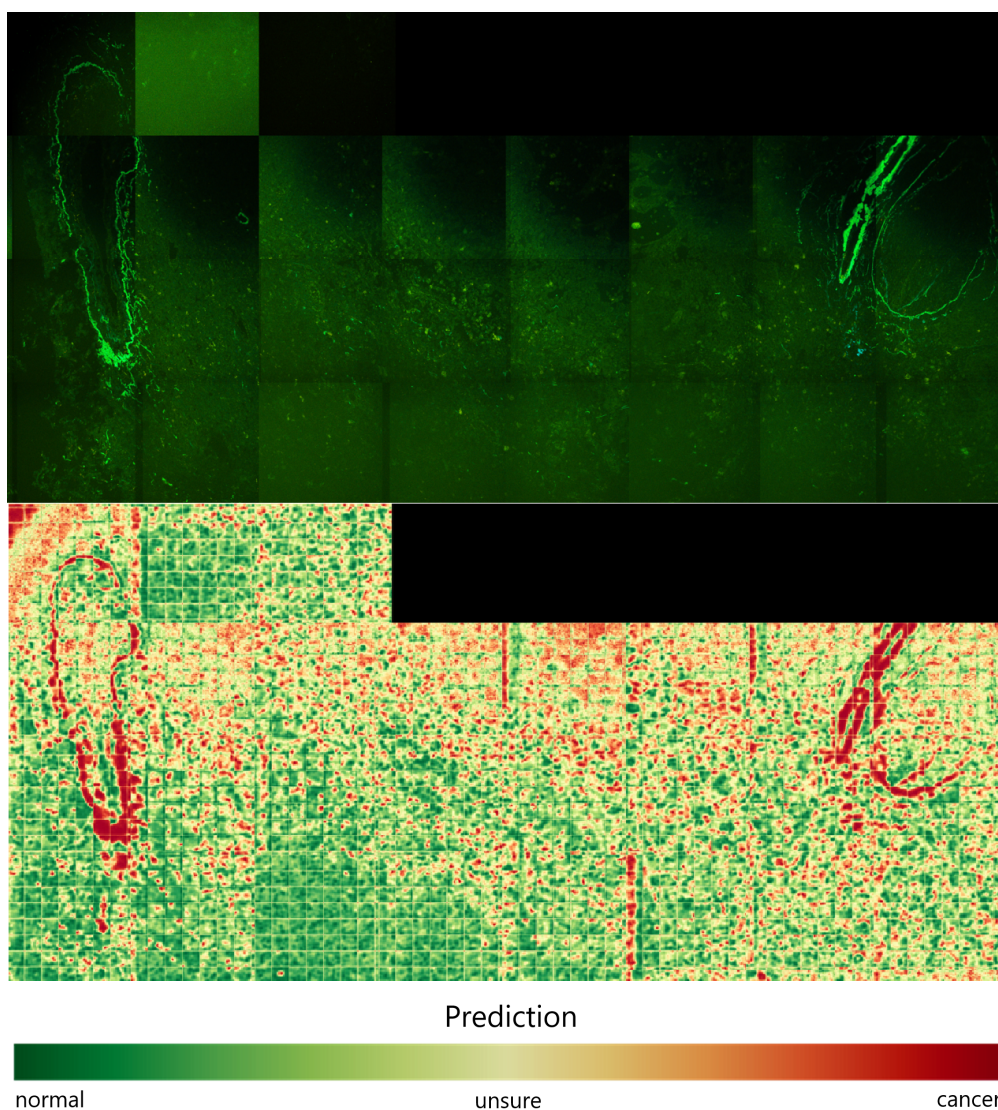


Figure 5.2: Top: visualization of part of lymph node 465666 with two veins/ducts, all cells are normal in this part of the tissue. Bottom: cancer probability prediction of our model



Figure 5.3: Prediction of cancer probability with superimposed cancer ground-truth mask in white

# Conclusion

When tested on previously unseen section of a human lymph nodes with sections of different part present in the training dataset the model achieved an accuracy of 0.73 and an AUC of 0.85, outperforming results reported in related studies [5][8]. This high performance is likely thanks to the similarity between the training and testing data, which did not account for the full cellular diversity and structural differences present in human lymph nodes. When tested on more different data, the model did not generalize well, with accuracy and AUC both dropping to about 0.55. This was due to our dataset being quite limited, consisting of only five lymph nodes, three of which were cancerous, making it very difficult for our model to generalize to the wide range of lymph nodes found in HNC patients.

We tested various model architectures for the classification and found out there is no significant difference between them. Rather than trying more intricate models, we think it could be beneficial to extract more features from lymph nodes that could help the machine learning, for example measuring on more wavebands or adding other FLIM features.

However it clearly shows that our approach is viable for classifying cancer cells in human lymph nodes. Provided with more training data with higher variety, we hope to achieve similar performance applicable to more general range of lymph nodes. In conclusion, we think that CNN based classification of FLIM images is an efficient method that could be in the future used in clinical setting reliably.

# Bibliography

- [1] Wikimedia Commons and Mehrdad Yazdani. *Example architecture of U-Net for producing k 256-by-256 image masks for a 256-by-256 RGB image*. 2019. URL: [https://commons.wikimedia.org/wiki/File:Example\\_architecture\\_of\\_U-Net\\_for\\_producing\\_k\\_256-by-256\\_image\\_masks\\_for\\_a\\_256-by-256\\_RGB\\_image.png](https://commons.wikimedia.org/wiki/File:Example_architecture_of_U-Net_for_producing_k_256-by-256_image_masks_for_a_256-by-256_RGB_image.png).
- [2] Duran-Sierra E., Cheng S., Cuenca R., Ahmed B., Ji J., Yakovlev V.V., Martinez M., Al-Khalil M., Al-Enazi H., and Cheng Y.-S.L. et al. “Machine-Learning Assisted Discrimination of Precancerous and Cancerous from Healthy Oral Tissue Based on Multispectral Autofluorescence Lifetime Imaging Endoscopy”. In: *Cancers* 13.19 (2021).
- [3] Matthew G. Vander Heiden, Lewis C. Cantley, and Craig B. Thompson. “Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation”. In: *Science* 324.5930 (2009), pp. 1029–1033.
- [4] Li MM, Puram SV, Silverman DA, Old MO, Rocco JW, and Kang SY. “Margin Analysis in Head and Neck Cancer: State of the Art and Future Directions.” In: *Annals of surgical oncology*, 26(12), 4070–4080. (2019).
- [5] Yuan N, Hassan MA, Ehrlich K, Weyers BW, Biddle G, Ivanovic V, Raslan OAA, Gui D, Abouyared M, Bewley AF, and et al. “Early Detection of Lymph Node Metastasis Using Primary Head and Neck Cancer Computed Tomography and Fluorescence Lifetime Imaging”. In: *Diagnostics* 14(18) (2024), p. 2097.
- [6] Datta R, Heaster TM, Sharick JT, Gillette AA, and Skala MC. “Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications.” In: *Journal of biomedical optics*, 25(7), 1–43. (2020).
- [7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. arXiv: 1505.04597 [cs.CV]. URL: <https://arxiv.org/abs/1505.04597>.
- [8] Fay Victoria, Tolstonog Genrich, Simon Christian, Psaltis Demetri, and Moser Christophe. “Analysis of imaging modalities for classification of tumoral vs. normal tissues using an FD-FLIM based MMF endoscopy probe”. In: *EPJ Web Conf.* 309 (2024), p. 04004. DOI: 10.1051/epjconf/202430904004. URL: <https://doi.org/10.1051/epjconf/202430904004>.

# Documentation

## A.1 FLIM microscope automation

The automation script works by moving the mouse in a similar fashion that a human operator would. For the script to work properly the LEICA FLIM software has to be open in the same fashion as in the image bellow (figure A.1), window to the left occupying about 75% of the screen, with the FLIM tab open and measurements selected. The script will firstly verify if this is the case and notify the user to move the window if not in the current position. The script will ask the user for the initial position of the microscope and the name of the square of this position. It is suggested to manually move the microscope to the up-left corner of the grid, to the first square and then name first square as (0,0)

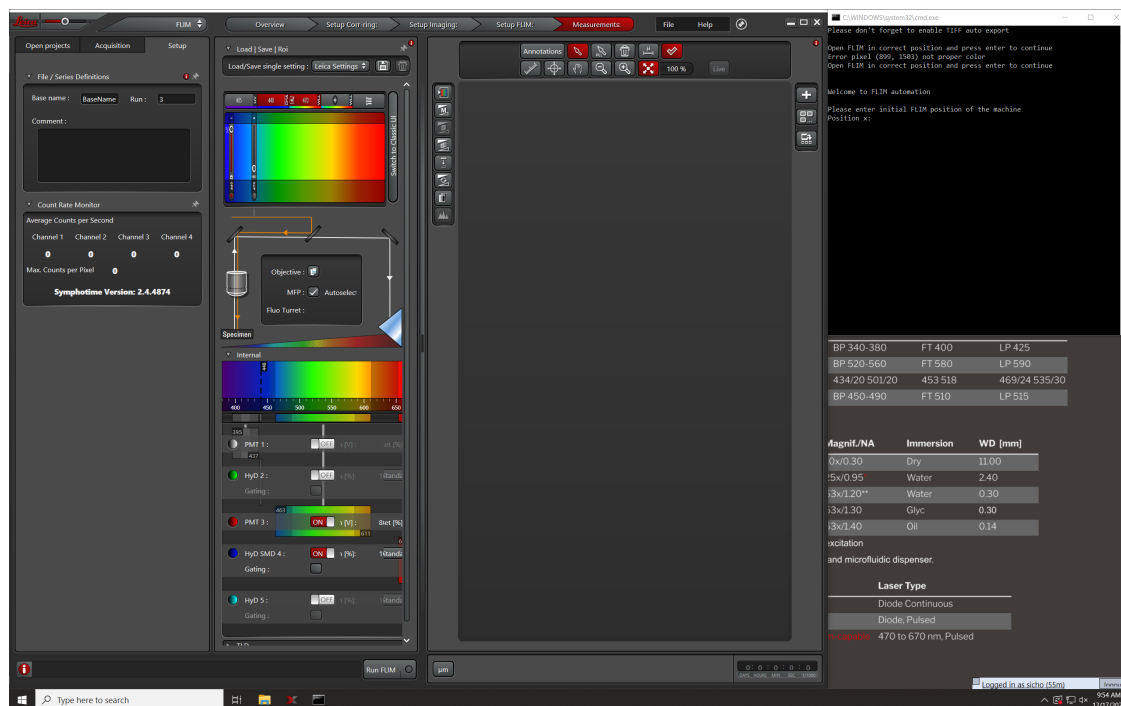


Figure A.1: Initial setting and window position of the FLIM automation

After correct window position is verified, please change back from FLIM to optical (TSC SP8) in

the dropout section and ensure that the left part of the window looks the same as in the picture A.2, with the 'acquisition' and setting tab titled 'XY: 512x512 | 400 Hz | 1.00 | 1.00 AU' expanded, while the 'Z-Stack' tab hidden away. If this is the case, you can switch back to FLIM and to measurements and begin with measuring. Before that, don't forget to enable TIF auto export in the Symphotime software.

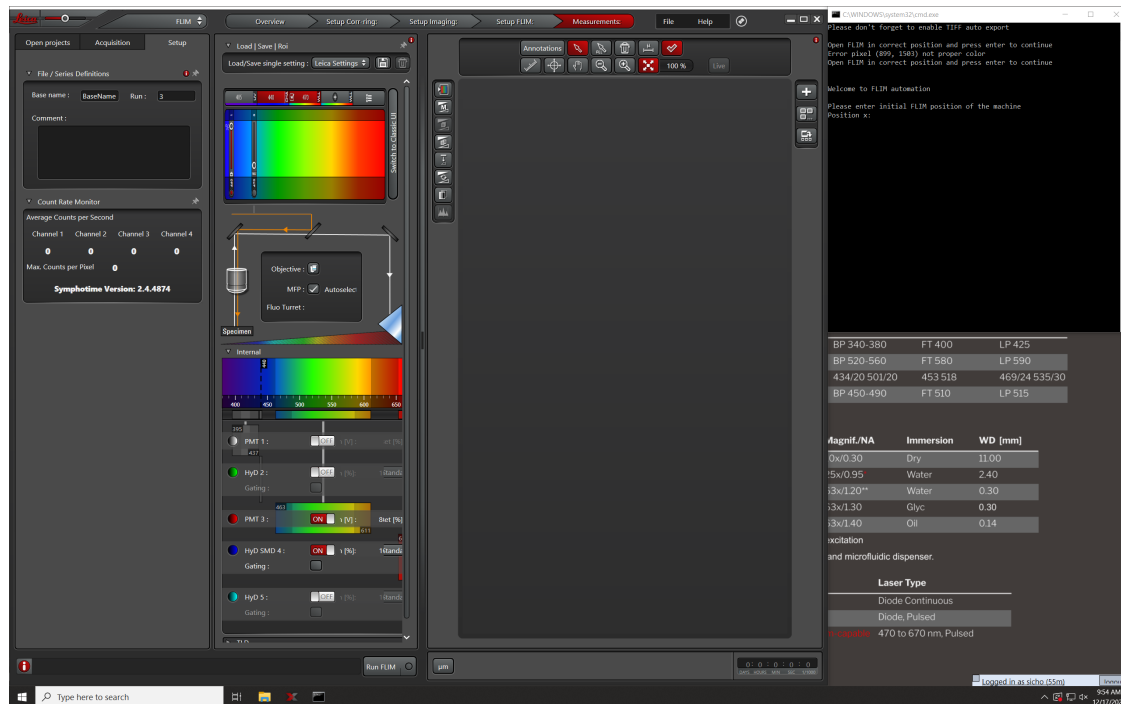


Figure A.2: Optical settings

The script can be then controlled by text commands. For moving write 'up', 'down', 'left' or 'right' and confirm by Enter. Then type by how many squares you want to move in the selected direction. The position will change, but microscope will stay in position. If you want to physically move the microscope type 'move'. This will move the stage to the position written in the script. Moving the stage with 'move' command can be useful for adjusting focus before scanning. If you are ready to scan, type 'scan'. Then write either '1', '3' or '5' indicating the size of area you want to scan. The scan will by default only scan tiles it has not scanned before, so if your areas overlay on duplicate scan will be performed. In case you want to scan area again ignoring any tiles that have been scanned before use the command 'rescan'. Scanned tiles will be automatically named with their tile names and waveband. It is imperative that mouse is not moved during scanning or moving, otherwise this will interrupt the programme. If you move the mouse by accident, set the LEICA window manually to its default position as in figure A.1 and resume work with the script.