Mueller matrix polarimetry for fresh ex-vivo cancerous tissue segmentation

Paulo Sampaio[a], Davide Scandella[a], C. H. Lucas Patty[a], Pablo Márquez-Neila[a], Heather DiFazio[a], Martin Wartenberg[b], Federico Storni[c], Brice-Olivier Demory[a,d], Daniel Candinas[c], Aurel Perren[b], and Raphael Sznitman[a]

[a]University of Bern, ARTORG Center, Bern, Switzerland
[b]University of Bern, Institute of Tissue Medicine and Pathology, Bern, Switzerland
[c]Bern University Hospital, Department of Visceral Surgery and Medicine, Bern, Switzerland
[d]University of Bern, Center for Space and Habitability, Bern, Switzerland

ABSTRACT

Pancreatic cancer, the third leading cause of cancer-related deaths, often requires complex surgery that relies on intraoperative diagnosis to assess resection margins. A common technique for this is frozen section (FS), in which a biopsy sample is snap-frozen, sectioned, stained, and evaluated by an expert pathologist. However, FS takes 20-30 minutes, and its accuracy is hindered by artifacts generated in the process. In this context, widefield Mueller Matrix Polarimetry (MMP) imaging represents an advantageous and robust alternative.

This study evaluates the usability of MMP to aid in assessing resection margins by segmenting the surface of freshly resected pancreatic tissue in cancer and non-cancer regions.

We imaged 56 fresh pancreatic samples from 19 patients with a custom-built 16-wavelength Mueller matrix imaging polarimeter. The imaging is performed in reflection, with a 16 deg phase angle, and the wavelengths range from 460 to 598nm. Haematoxylin and eosin (H&E) stained slides of these samples were used as ground truth to train a deep-learning binary segmentation model to detect the presence of cancer tissue surface from the MMP imager.

Our pixel-wise segmentation model reached an AUC of 94 ± 5, clearly delineating cancer regions across different patients. Moreover, the end-to-end time for imaging a sample and generating a prediction is approximately 55 seconds, significantly shorter than the 20-30 minutes commonly required for RFS.

These findings indicate that MMP-based segmentation has a strong potential as an intraoperative diagnosis alternative, identifying resection margins with high sensitivity (true positive rate) and specificity (true negative rate) in less than a minute, minimizing waiting time in the operating room, and improving surgery workflow and clinical decision-making.

Keywords: polarimetry, pancreatic cancer, histopathology, machine learning, segmentation, classification

1. INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is one of the deadliest forms of cancer, with a 5-year survival rate below 10%.[1,2] It remains asymptomatic for an extended period and, when diagnosed, often requires complex surgery with intraoperative assessment of resection margins.[3,4] The current standard for intraoperative diagnosis is frozen section (FS),[5] where a snap-frozen biopsy sample is sectioned into a $5\mu m$-thick slice, stained with Hematoxylin and Eosin (H&E), and evaluated by an expert gastrointestinal pathologist. The process takes up to 40 minutes and is prone to technical limitations, such as freezing artifacts, poor section quality, bloated cell morphology, and inconsistent staining.[6,7] The gold standard for diagnosis remains the H&E slides from formalin-fixed paraffin-embedded (FFPE) blocks, but these require at least 2 days of preparation, making them unsuitable for intraoperative use. Compounding these challenges, PDAC is characterized by a dense stromal response, or desmoplastic reaction, where malignant cells infiltrate the surrounding connective tissue.[8] This

Further author information: E-mail: paulo.sampaio@unibe.ch

Applications of Machine Learning 2025, edited by Michael E. Zelinski, Tarek M. Taha, Barath Narayanan Abdul A. S. Awwal, Khan M. Iftekharuddin, Proc. of SPIE Vol. 13606, 1360614
2025 Published by SPIE · 0277-786X · doi: 10.1117/12.3064291

fibrotic environment can closely resemble chronic pancreatitis, making tumor identification difficult even for experienced pathologists, whether analyzing frozen or paraffin-embedded sections.[9]

Mueller matrix polarimetry (MMP) is a label-free optical imaging technique that characterizes how a medium (e.g. tissue) changes the polarization state of light.[10] By measuring the full 4x4 Mueller matrix, MMP captures changes due to birefringence, depolarization and scattering.[11] Its sensitivity to anisotropic organization and density of fibrous structures makes it particularly valuable for biological applications.[12] MMP has shown promising results in studies of cervical,[13] colon,[14] breast, rectal,[15] and skin[16] cancer, as well as in brain tissue segmentation,[17] pancreatic normal and abnormal tissue detection.[18] Based solely on innocuous light intensity measurements and with a rapid acquisition time (less than 1 minute), MMP offers an attractive alternative to accelerate intraoperative diagnostics.

Despite its clinical relevance, few optical studies have investigated the polarimetric features of PDAC tissue in detail. This is partly due to the anatomical complexity of the pancreas, which is located deep in the abdominal cavity and surrounded by other critical organs and blood vessels, making direct in vivo optical imaging highly challenging. Therefore, to characterize the polarimetric signatures of PDAC and its surrounding tissue structures, we focused on fresh ex vivo surgical biopsies, imaged within 45 minutes of resection and without additional sample preparation.

In this work, we imaged 57 fresh ex vivo pancreatic samples using a custom-built, 16-wavelength spectro-polarimeter mounted in a reflection mode with a 16° phase angle, covering wavelengths from 460nm to 590nm. Using annotated FFPE slides as ground truth, we trained a multitask machine learning model to perform both classification and segmentation. The classification and segmentation models achieved ROC-AUC of 97 ± 3 and 94 ± 5, respectivelly, with a total acquisition and prediction time of approximately 1 minute. While the classification provides a fast binary decision analogous to FS analysis, segmentation provides spatial localization of malignant regions, an important enabler for future in vivo margin assessment applications, where cancerous areas could be identifies directly within the surgical field, eliminating the need for biopsy altogether.

2. METHOD

In this section, we describe the full methodology of our approach. We begin with the design of our custom-built multispectral Mueller Matrix polarimetric (MS-MMP) imaging device, followed by details on the data collection and histopathology workflow used to obtain ground truth labels. We then explain the pre-processing steps applied to the raw data, describe our multitask deep learning architecture for simultaneous classification and segmentation, and outline the training strategy and evaluation metrics used to validate model performance.

2.1 MS-MMP Imager

Our device is a custom-built dual-rotating retarder Mueller matrix polarimeter[10] operating in reflection mode. The system consists of a LED source, a polarization state generator (PSG), a polarization state analyzer (PSA), and a multispectral camera, mounted with a 16° phase angle between PSA and PSG. The PSG produces the desired polarization states by passing the light through a fixed linear polarizer and a rotating quarter-wave plate. After interacting with the tissue sample, positioned on a custom 3D-printed holder ensuring optimal focus, the reflected light passes through the PSA, which is similarly composed of a rotating quarter-wave plate and a fixed linear polarizer.

The detector is a multispectral camera (Photonfocus, Lachen, Switzerland) equipped with 16 passbands ranging from 460 nm to 598 nm, arranged in a 4×4 mosaic with a full width at half maximum (FWHM) of 15 nm. Each spectral band has a spatial resolution of 512×272 pixels, with a pixel size of approximately 140 μm. By acquiring a series of images while rotating the quarter-wave plates of the PSG and PSA at harmonic frequencies, the system computes the full pixel-wise Mueller matrix for each wavelength.[19] The total acquisition time for a 272×512 pixel image across all 16 wavelengths is approximately 45 seconds.

Figure 1 illustrates the overall setup, including the optical layout, internal component arrangement, and final user-facing system. The acquisition software was developed in Python and features a graphical user interface (GUI) with intuitive routines for calibration and data acquisition.

Figure 1. Overview of the MS-MMP system: (left) conceptual schematic of the optical design; (center) open optical box with mounted components; (right) deployed benchtop solution at the Institute for Tissue Medicine and Pathology of the University of Bern, featuring an enclosed optical box, touchscreen interface, and QR-code scanner for sample tracking.

2.2 Data collection

We installed the device at the Institute of Tissue Medicine and Pathology (ITMP) of the University of Bern, which processes tissue specimens from the Department of Visceral Surgery from the Bern University Hospital.

Immediately after gross sectioning, the freshly excised pancreatic tissue samples were placed on pathology cassettes by a pathology resident. Each cassette is identified with a unique QR code, which was scanned and logged by our device. The resident then manually selected the exposure time to optimize image quality by balancing overexposed and underexposed pixels in the tissue region.

This study was approved by the Cantonal Ethics Committee of Bern (IRB no: BE-2024-00553) and conducted in accordance with the Declaration of Helsinki.

2.3 Ground truth generation

After polarimetric imaging, the tissue samples followed the standard histopathology workflow. Each sample was fixed in formalin, embedded in paraffin, sectioned into 2.5μm-thick slices, and stained with Hematoxylin and Eosin (H&E). The resulting slides were scanned at high resolution using a digital slide scanner.

An expert gastrointestinal pathologist manually annotated the digital slides using QuPath,[20] identifying malignant and non-malignant regions of interest.

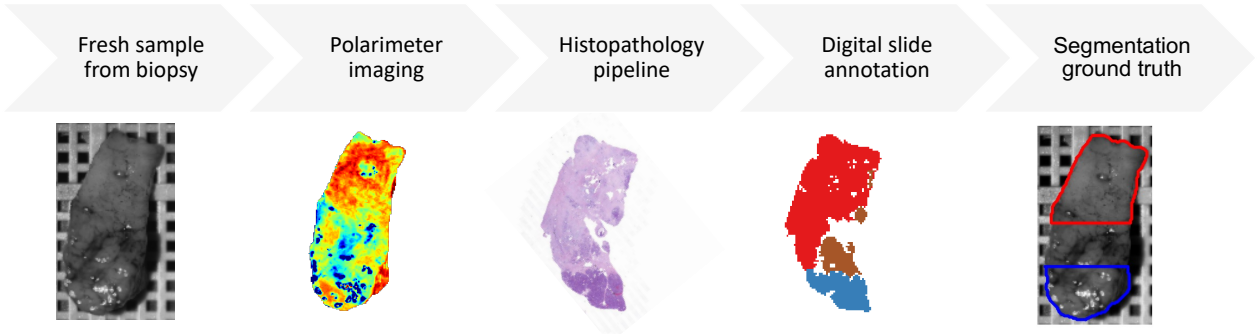

Figure 2. After imaging, the sample follows the histopathology pipeline to produce a digital slide, which is fully annotated by a specialized pathologist. For classification, if the slide contains malignant regions, the sample is labeled as cancer, otherwise, as healthy. For segmentation, annotated regions are visually matched to the polarimetry image to create a pixel-wise ground truth, where red region is labeled cancer while the blue is labeled healthy.

The pixel-wise digital slide annotations were binarized: pixels containing PDAC or desmoplastic reaction were labeled as cancer, while all other tissue types (including endo-/exocrine pancreas, fat, healthy stroma, fibrosis, and pancreatitis) were labeled as healthy. For the classification ground truth, any sample whose slide contained malignant tissue was labeled as cancer; otherwise, healthy. For the segmentation ground truth, annotated regions were visually matched to the polarimetry image to ensure spatial correspondence. Only regions where the histological annotation aligned with the polarimetric field of view were used, allowing for accurate transfer of pixel-wise labels. This workflow is illustrated in Figure 2.

2.4 Image pre-processing

The input to the model consists of the raw pixel-wise Mueller matrix values acquired at 16 wavelengths. Each wavelength produces a 4×4 Mueller matrix, which values are stacked resulting in a total of 256 input channels. Since the matrix elements are normalized to M_{11}, the M_{11} element is constant across the sample and thus excluded, reducing the input dimensionality to 240 channels.

To isolate the tissue from the background, we applied the Segment Anything Model (SAM)[21] developed by Meta AI to generate a tissue mask from the intensity images. This mask was used to crop the image around the tissue, removing excess background and reducing computational load.

2.5 Machine learning model

We propose a multitask convolutional neural network combining classification and segmentation in a single architecture. The model follows a U-Net[22] inspired encoder-decoder structure, where the encoder extracts hierarchical features through successive convolutional layers, and the decoder reconstructs a pixel-wise segmentation map.

The encoder consists of two downsampling blocks, each composed of two convolutional layers with 3×3 kernels, batch normalization, and ReLU activations, followed by 2×2 max-pooling. This is followed by a bottleneck block with a similar double convolution structure but increased feature dimensionality. For segmentation, the decoder upsamples the bottleneck features using transposed convolutions and skip connections from the encoder, progressively recovering the original image resolution. The segmentation head outputs a two-class pixel-wise prediction (cancer vs. healthy).

In parallel, the bottleneck features are globally averaged using adaptive pooling and passed through a linear classifier to produce a binary class label for the entire image. This enables simultaneous image-level classification and pixel-wise segmentation from a single forward pass.

The network starts with 64 channels, doubling the number at each downsampling step, reaching a 256-channel bottleneck. An overview of the architecture is shown in Figure 3.

2.6 Training strategy

We trained the model using a repeated by-patient cross-validation scheme, consisting of 5 repetitions of a 3-fold split, totaling 15 distinct train/test combinations. Samples from the same patient were always assigned to the same split to prevent data leakage. To address the class imbalance, we used a weighted random sampling strategy during training, assigning inverse class-proportional weights to each sample. This ensured that batches were approximately class-balanced across each epoch, improving robustness to underrepresented cancer samples.

The model was optimized end-to-end for both classification and segmentation tasks simultaneously over 100 epochs. We used the Adam optimizer with a learning rate of 10^{-4} and minimized the sum of the segmentation and classification cross-entropy losses.

During training, input tensors were dynamically zero-padded within each batch to match the height and width of the largest sample in the batch, accommodating variations in tissue size while maintaining consistent tensor shapes for efficient processing.

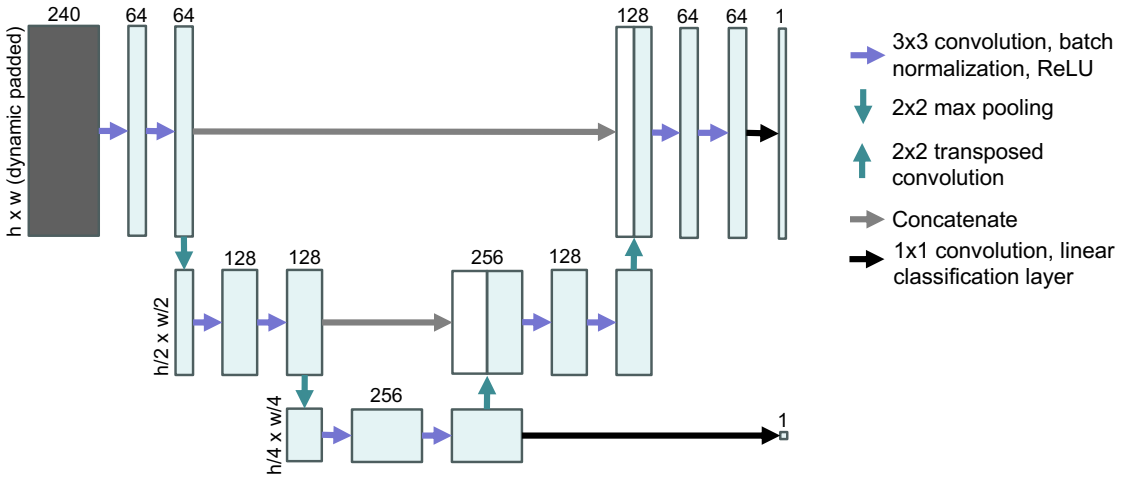

Figure 3. Overview of the multitask architecture. The model consists of two downsampling layers, a bottleneck, and two upsampling layers. The classification head is attached to the bottleneck and outputs a sample-wise label, while the segmentation decoder produces a pixel-wise prediction. The entire model is trained jointly in an end-to-end fashion.

2.7 Evaluation

We evaluated model performance using the area under the receiver operating characteristic curve (ROC-AUC) and the precision-recall (PR) curves for both classification and segmentation tasks. For segmentation, we also performed qualitative analysis of the predicted masks.

To assess the contribution of polarimetric information, we benchmarked our model against a baseline trained solely on multispectral intensity images (without polarization features).

3. RESULTS

In this section, we present a comprehensive evaluation of our system and model. We begin by describing the dataset acquired through our imaging campaign and highlight key characteristics such as class distribution and sample variability. We then conduct an exploratory analysis of biologically relevant polarimetric features to understand their discriminative potential. Next, we report quantitative metrics for classification and segmentation performance, including comparisons with a non-polarimetric baseline and an analysis of prediction stability. Finally, we showcase qualitative examples to assess model behavior on both typical and challenging samples.

3.1 Data collection

Our data collection campaign took place between October 2023 and October 2024. Inclusion criteria were: (1) signed general consent, (2) age above 18 years, and (3) no neoadjuvant treatment. We collected a total of 56 biopsy samples from 19 patients meeting these criteria.

Table 1 summarizes the key dataset characteristics, including the number of patients, samples, pixels, and biopsy area statistics for both healthy and cancerous tissues. Notice that we have five patients for whom we have both cancer and healthy samples. On average, the samples are approximately $1.3mm^2$. The dataset is also mildly imbalanced, with 44 healthy and 12 cancerous samples — a discrepancy reflected in the total number of pixels per class — and this was accounted for during model training by using a weighted random sampling strategy

Figure 4 (left) shows an example of raw Mueller matrix data captured by our imaging device at 598.2 nm. The right panel illustrates some of the derived polarimetric features used for exploratory analysis, which will be discussed in the following section.

Table 1. Overview of collected data with 56 tissue samples from 19 patients.

Label	Patients	Samples	Total pixels	Biopsy area (mm^2)			
				Average	Std. dev.	Min	Max
Healthy	16	44	122,247	124	93	47	420
Cancer	8	12	36,316	152	142	51	512
Total	19	56	158,563	130	104	47	512

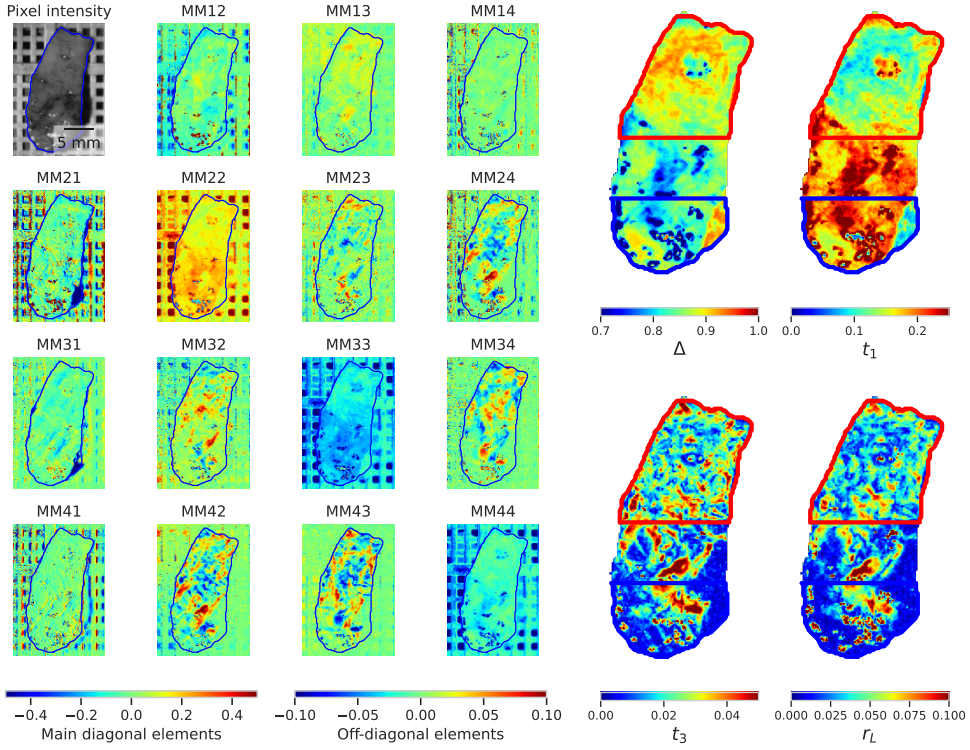

Figure 4. Example images from the dataset. left: raw polarimetric data acquired at 598.2 nm, showing the full 4×4 Mueller matrix. right: derived features used for exploratory analysis, including depolarization, anisotropy degree, linear retardance, and circular-to-linear coupling.

3.2 Exploratory analysis

As an initial assessment, we extracted optical features commonly associated with tissue microstructure from the raw Mueller matrix data. Although our model is trained directly on the raw Mueller matrix values, analyzing these standard features can provide insights into the separability potential of polarimetric data. For this analysis, we computed the following features:

- **Depolarization**, computed as $1 - \frac{|\mathrm{tr}(M_\Delta)-1|}{3}$, where M_Δ is obtained from the polar decomposition $M = M_\Delta M_R M_D$[23]
- **Anisotropy degree** (t_1), calculated as $t_1 = \frac{\sqrt{(m_{22}-m_{33})^2+(m_{23}+m_{32})^2}}{2}$.[24]
- **Linear retardation** (t_3), computed as $t_3 = \frac{\sqrt{m_{42}^2+m_{43}^2}}{2}$.[24]
- **Conversion from circular to linear polarization** (r_L), defined as $r_L = \sqrt{m_{24}^2 + m_{34}^2}$.[25]

These features have been previously reported as markers of collagen fiber density and orientation.[26] As illustrated in Figure 4 (right), clear visual contrasts between healthy and cancerous regions are apparent, particularly in t_1 and r_L. Healthy tissue exhibits a higher anisotropy degree, suggesting a more organized and

directionally aligned fibrous structure—consistent with the presence of well-ordered collagen in normal stroma. In contrast, cancerous regions show elevated r_L values, indicating increased scattering and a more disordered tissue microstructure, potentially reflecting the chaotic architecture of desmoplastic stroma.

These observations reinforce the biological relevance of polarimetric information and support their utility in downstream classification and segmentation of pancreatic ductal adenocarcinoma (PDAC).

3.3 Quantitative results

Our multitask model using polarimetric data (MS-MMP) achieved a classification ROC-AUC of 97 ± 3 and a segmentation ROC-AUC of 94 ± 5, averaged across all cross-validation folds. The corresponding average precision (AP) scores were 94 ± 5 for classification and 84 ± 10 for segmentation. Considering cancer as the positive class, the model reached a sensitivity of approximately 90% at a false positive rate of 10%. When trained without polarimetric information (MS) (using only multispectral intensity), performance dropped consistently across all metrics, with reductions of at least 30 percentage points. These results are summarized in Figure 5, which presents the ROC and precision-recall curves for both tasks along with variability across folds with and without polarimetry.

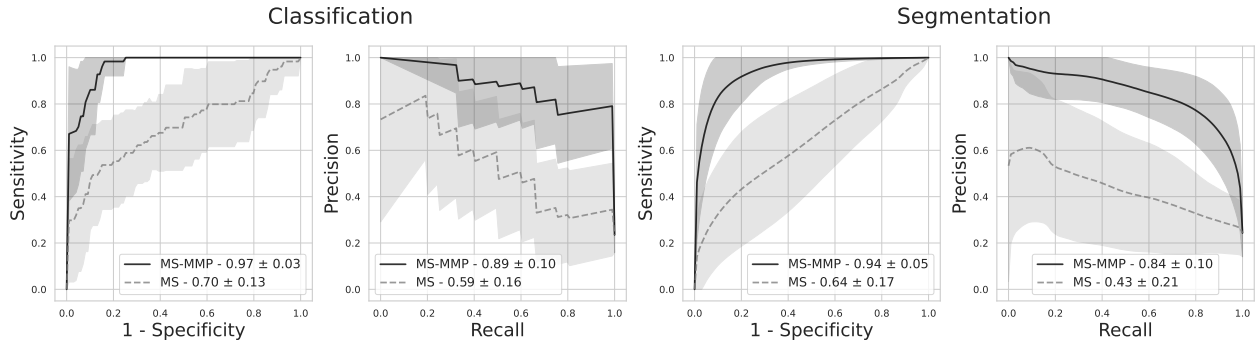

Figure 5. Quantitative evaluation of the multitask model across all cross-validation folds, with (MS-MMP) and without polarimetry (MS). Left: ROC and Precision-Recall (PR) curves for classification. Right: ROC and PR curves for segmentation. Shaded areas represent standard deviation across folds.

Assessing prediction stability across folds, we observed lower standard deviation for samples with very high or very low average prediction scores, indicating greater model confidence on these cases. In contrast, predictions with intermediate scores showed higher variability across folds, suggesting greater uncertainty in borderline cases. This behavior is illustrated in Figure 6, which shows the relationship between prediction confidence and cross-fold variability for each sample for both classification and segmentation. Two healthy samples stand out in both tasks with consistently high predicted scores, resulting in false positives in some of the folds where they appear. One of these samples contains a blood vessel, an uncommon structure in our dataset, which may have influenced the model's prediction. The second sample corresponds to a case of pancreatitis, a condition that is notoriously challenging to distinguish from cancer, even for experienced pathologists.

3.4 Qualitative results

For the qualitative evaluation, we examined the predicted segmentation masks and classification scores in comparison to the annotated digital slides. Figure 7 shows results for six samples from the same patient, each presented as a pair: on the left, the annotated digital slide with the sample-level ground truth label; on the right, the corresponding segmentation mask with the predicted classification score.

Across all cases, cancerous samples received high predicted scores and displayed coherent segmentation masks that align well with the annotated regions. Similarly, healthy samples were assigned low cancer scores and exhibited homogeneous background predictions. This consistent performance within the same patient suggests the model's robustness and its ability to generalize beyond intra-patient correlations.

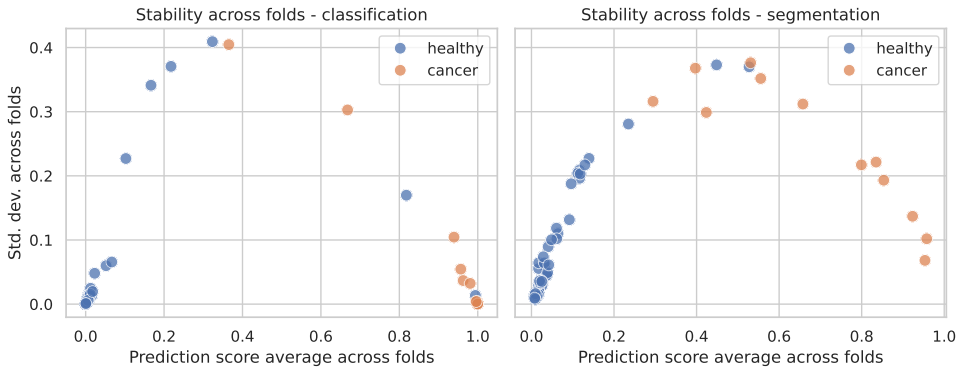

Figure 6. Prediction stability across folds, with classification results on the left and segmentation on the right. Each point represents a sample, with the average prediction score on the x-axis and the standard deviation across folds on the y-axis. The model behaves similarly in both tasks: lower variance is observed for highly confident predictions (at the extremes), while higher variance reflects uncertainty in borderline cases. In both settings, two healthy samples are consistently predicted with higher scores, contributing to false positives.

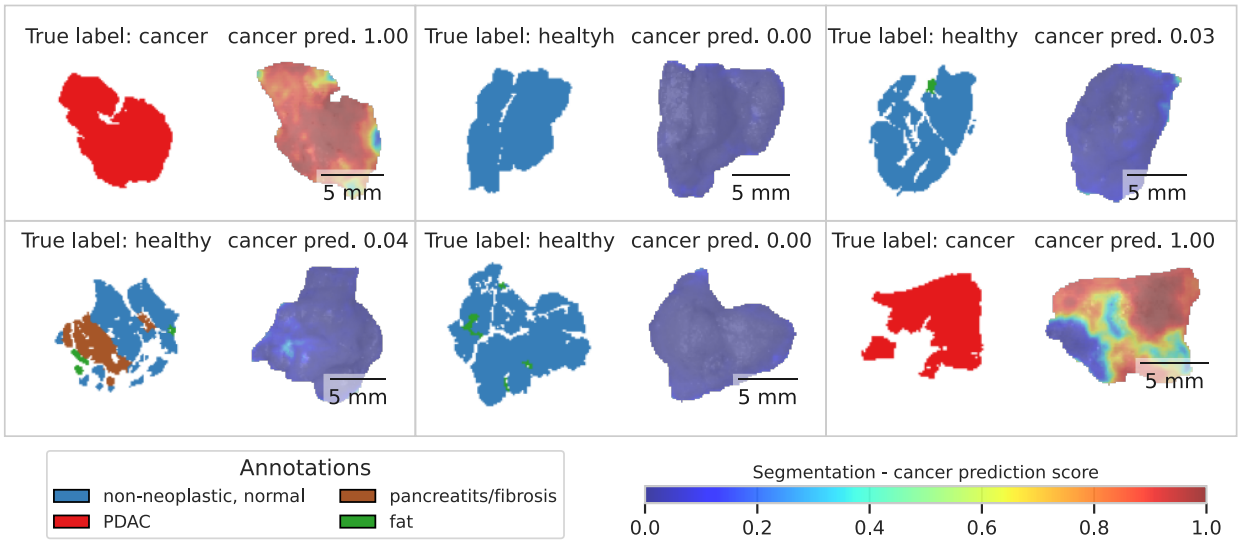

Figure 7. Qualitative results for six samples from the same patient. Each sample is shown as a pair: (left) annotated digital slide with ground truth label, and (right) model output with predicted segmentation mask and classification score. Cancer samples are accurately identified with high scores and localized masks, while healthy samples exhibit low predicted scores and uniform background segmentation, indicating strong intra-patient generalization.

Another insightful experiment involves analyzing segmentation masks for heterogeneous samples—i.e., those containing both healthy and cancerous regions. Figure 8 displays, on the leftmost column, the annotated digital slides for two such samples, followed by the predicted segmentation masks across all folds in which each sample appeared in the test set.

In the first row, the sample was consistently classified as cancer, with the segmentation masks correctly highlighting the malignant region in the upper portion, in agreement with the annotated slide. However, some false positive activations appear in the lower part of the sample.

In the second row, the classification scores vary across folds. This variability is likely due to the small cancerous region present, which may be lost or diluted during the downsampling and global average pooling operations of the model. This highlights a known challenge in detecting subtle or spatially limited lesions with

encoder-decoder architectures.

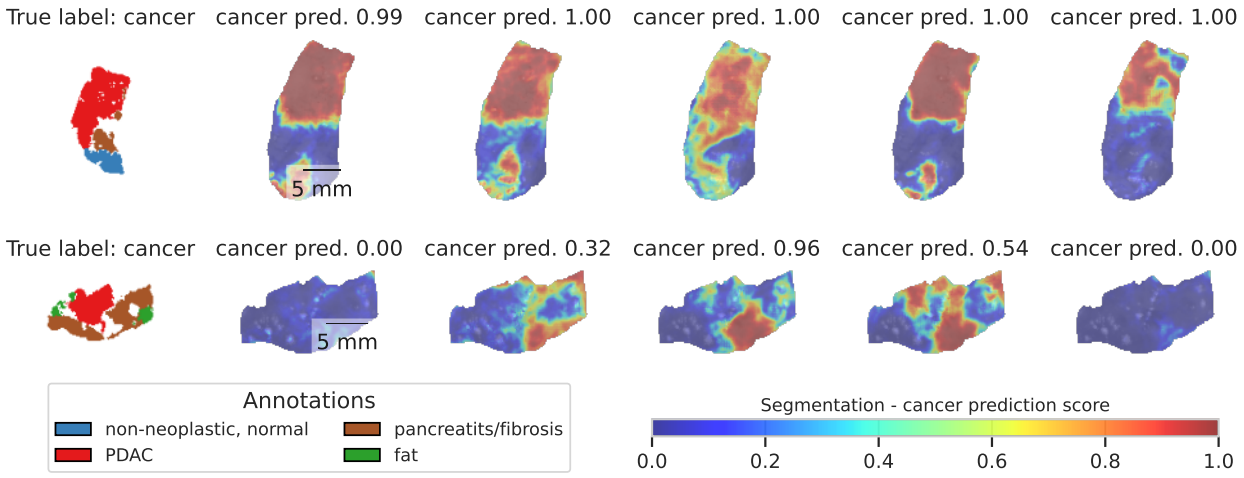

Figure 8. Qualitative analysis of heterogeneous samples. Left: annotated digital slides for two representative cases. Right: predicted segmentation masks across all folds where each sample appeared in the test set. Top row: consistent detection of a large cancerous region, with some false positives. Bottom row: a sample with a small cancer focus, leading to inconsistent predictions across folds due to downsampling and averaging effects in the model.

4. DISCUSSION

In this study, we combined raw polarimetric data acquired from a custom-built multispectral Mueller Matrix polarimeter with a multitask deep learning model to simultaneously output a sample-wise cancer probability (analogous to frozen section diagnosis) and a pixel-wise segmentation mask, providing valuable spatial context. The complete workflow, from image acquisition to prediction, takes approximately one minute, a significant reduction compared to the 30–40 minutes typically required for frozen section analysis.

Our model achieved high ROC-AUC and average precision scores for both classification and segmentation tasks, with consistent results across intra-patient samples. This suggests that the network effectively captures tissue-specific signatures directly from the polarimetric measurements.

The exploratory analysis of handcrafted polarimetric features supports this interpretation. Metrics such as anisotropy degree (t_1) and circular-to-linear coupling (r_L) exhibited patterns consistent with established histopathological differences between normal stroma and desmoplastic tissue, reinforcing the biological relevance of polarimetric contrast.

Nonetheless, some limitations must be acknowledged. First, the dataset is relatively small, and certain benign structures,such as blood vessels or precursor lesions, are underrepresented, which may limit generalizability. Second, the model's predictions were less stable for samples with very small malignant regions, likely due to spatial resolution loss introduced by downsampling and global average pooling. Incorporating attention mechanisms may help improve sensitivity in such cases; however, these methods typically benefit from larger datasets, underscoring the importance of continued data collection.

While the segmentation output is a critical enabler for future intraoperative direct imaging applications, several challenges remain before clinical translation. These include managing the complex surface geometry of in vivo tissue, accounting for motion artifacts from patient respiration, and accommodating variable focal distances and incident angles during acquisition. These technical considerations warrant future investigation.

Despite these challenges, our results mark a promising step toward fast, spatially resolved intraoperative cancer detection, with the potential to streamline surgical workflows and improve decision-making at the point of care.

ACKNOWLEDGMENTS

This work was supported in part by the University of Bern, FreeNovation grant, and the Fond'actio contre le cancer grant.

REFERENCES

[1] Siegel, R. L., Miller, K. D., and Jemal, A., "Cancer statistics, 2020," *CA: a cancer journal for clinicians* **70**, 7–30 (5 2020).

[2] Rawla, P., Sunkara, T., and Gaduputi, V., "Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors," *World Journal of Oncology* **10**(1), 10 (2019).

[3] Chavez, J. A., Chen, W., Freitag, C. E., and Frankel, W. L., "Pancreatic frozen section guides operative management with few deferrals and errors," *Archives of Pathology & Laboratory Medicine* **146**, 84–91 (Mar. 2021).

[4] Vazzano, J., Chen, W., and Frankel, W. L., "Intraoperative frozen section evaluation of pancreatic specimens and related liver lesions," *Archives of Pathology & Laboratory Medicine* **149**, e63–e71 (May 2024).

[5] Ganjali, H., "Frozen section: An overview," *Scholar's Research Library* **3**(11) (2012).

[6] Nelson, D. W., Blanchard, T. H., Causey, M. W., Homann, J. F., and Brown, T. A., "Examining the accuracy and clinical usefulness of intraoperative frozen section analysis in the management of pancreatic lesions," *American journal of surgery* **205**, 613–617 (5 2013).

[7] Campanale, R. P., Frey, C. F., Farias, L. R., Twomey, P. L., Guernsey, J. M., Keehn, R., and Higgins, G., "Reliability and Sensitivity of Frozen-Section Pancreatic Biopsy," *Archives of Surgery* **120**, 283–288 (3 1985).

[8] Feig, C., Gopinathan, A., Neesse, A., Chan, D. S., Cook, N., and Tuveson, D. A., "The pancreas cancer microenvironment," *Clinical Cancer Research* **18**, 4266–4276 (Aug. 2012).

[9] Klöppel, G. and Adsay, N. V., "Chronic pancreatitis and the differential diagnosis versus pancreatic cancer," *Archives of Pathology & Laboratory Medicine* **133**, 382–387 (Mar. 2009).

[10] Goldstein, D. H., "Mueller matrix dual-rotating retarder polarimeter," *Applied Optics* **31**, 6676 (Nov. 1992).

[11] Ghosh, N. and Vitkin, I. A., "Tissue polarimetry: concepts, challenges, applications, and outlook," *Journal of Biomedical Optics* **16**, 110801 (Nov. 2011).

[12] He, C., He, H., Chang, J., Chen, B., Ma, H., and Booth, M. J., "Polarisation optics for biomedical and clinical applications: a review," *Light: Science & Applications* **10** (Sept. 2021).

[13] Pierangelo, A., Nazac, A., Benali, A., Validire, P., Cohen, H., Novikova, T., Ibrahim, B. H., Manhas, S., Fallet, C., Antonelli, M.-R., and Martino, A.-D., "Polarimetric imaging of uterine cervix: a case study," *Optics Express* **21**, 14120 (6 2013).

[14] Pierangelo, A., Benali, A., Antonelli, M.-R., Novikova, T., Validire, P., Gayet, B., and De Martino, A., "Ex-vivo characterization of human colon cancer by Mueller polarimetric imaging," *Optics Express* **19**, 1582 (1 2011).

[15] Pierangelo, A., Manhas, S., Benali, A., Fallet, C., Totobenazara, J.-L., Antonelli, M.-R., Novikova, T., Gayet, B., De Martino, A., and Validire, P., "Multispectral Mueller polarimetric imaging detecting residual cancer and cancer regression after neoadjuvant treatment for colorectal carcinomas," *Journal of Biomedical Optics* **18**, 46014 (4 2013).

[16] Jacques, S. L., Ramella-Roman, J. C., and Lee, K., "Imaging skin pathology with polarized light.," *Journal of biomedical optics* **7**, 329–340 (7 2002).

[17] Felger, L., Rodríguez-Núñez, O., Gros, R., Maragkou, T., McKinley, R., Moriconi, S., Murek, M., Zubak, I., Novikova, T., Pierangelo, A., and Schucht, P., "Robustness of the wide-field imaging mueller polarimetry for brain tissue differentiation and white matter fiber tract identification in a surgery-like environment: an ex vivo study," *Biomedical Optics Express* **14**, 2400 (May 2023).

[18] Sampaio, P., Lopez-Antuña, M., Storni, F., Wicht, J., Sökeland, G., Wartenberg, M., Márquez-Neila, P., Candinas, D., Demory, B. O., Perren, A., and Sznitman, R., "Müller matrix polarimetry for pancreatic tissue characterization," *Scientific reports* **13** (5 2023).

[19] Azzam, R. M. A., "Photopolarimetric measurement of the Mueller matrix by Fourier analysis of a single detected signal," *Optics Letters* **2**, 148–150 (June 1978). Publisher: Optica Publishing Group.

[20] Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M., and Hamilton, P. W., "QuPath: Open source software for digital pathology image analysis," *Scientific reports* **7** (5 2017).

[21] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollar, P., and Girshick, R., "Segment Anything," in [*IEEE/CVF International Conference on Computer Vision (ICCV)*], 4015–4026 (2023).

[22] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," (2015).

[23] Lu, S.-Y. and Chipman, R. A., "Interpretation of Mueller matrices based on polar decomposition," *Journal of the Optical Society of America A* **13**, 1106 (May 1996).

[24] He, H., Zeng, N., Du, E., Guo, Y., Li, D., Liao, R., and Ma, H., "A possible quantitative mueller matrix transformation technique for anisotropic scattering media/eine mögliche quantitative müller-matrix-transformations-technik für anisotrope streuende medien," *Photonics & Lasers in Medicine* **2** (Jan. 2013).

[25] Li, P., Lv, D., He, H., and Ma, H., "Separating azimuthal orientation dependence in polarization measurements of anisotropic media," *Opt. Express* **26**, 3791–3800 (Feb 2018).

[26] Dong, Y., Liu, S., Shen, Y., He, H., and Ma, H., "Probing variations of fibrous structures during the development of breast ductal carcinoma tissues via mueller matrix imaging," *Biomed. Opt. Express* **11**, 4960–4975 (Sep 2020).