POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

Politecnico di Torino

Master's Degree Thesis

Quantitative assessment of radiotherapy-induced microvasculature changes in non-melanoma skin cancer using OCTA

Supervisors

Prof. Kristen Mariko MEIBURGER

Ing. Giulia ROTUNNO

Prof. Mengyang LIU

Lisa KRAINZ, PhD

Gerd HEILEMANN, PhD

Candidate

Francesco GILI

October 2025

Summary

Non-melanoma skin cancer (NMSC) is a term that refers mainly to basal cell carcinomas (BCCs) and squamous cell carcinomas (SCCs). NMSC is the most common form of cancer in Caucasians, with a growing incidence worldwide. Among the possible treatments, radiotherapy is often preferred as a tissue-preserving non-surgical option, which shows effectiveness in terms of cosmetic results and local control. Vessel morphologies vary with different tumour stages and as a result of treatment, so cutaneous blood flow can be used as a biomarker to diagnose NMSC. Optical coherence tomography angiography (OCTA) can visualize both the microstructure and vasculature of skin non-invasively and could thus be used in treatment planning and follow-up. The objective of this thesis is to understand whether feature extraction on OCTA can quantitatively assess lesion-dependent changes in skin microvasculature, as a result of radiotherapy.
The dataset was obtained at the Vienna General Hospital by recruiting 20 patients with BCC, SCC, or actinic keratosis. It's comprised of 146 acquisitions, with 18 of them being healthy volumes, used as a baseline, and the remaining belonging each to a different timepoint: before, right after, 3 months after, and 6 months after radiotherapy. Each time, 4 OCT volumes are acquired and then processed to obtain an OCTA volume through an intensity-based method. These arrays are semi-automatically segmented and further masked to ignore, during feature extraction, any areas that might lack angiographic information. A 2D projection is also computed along the depth axis of the 3D volume. Skeletonizations of both 2D and 3D arrays are computed.

The features are chosen to describe the peculiarities of NMSC vascularization. Therefore, some parameters, such as branch tortuosity and entropy, describe how chaotically the vessels are intertwined with each other. Others, like vascular density, number of trees, and mean radius, observe how packed the blood vessels are and if their dimension changes between different timepoints. The variation of certain parameters with depth is calculated as well. These features allow to further compare lesions to healthy skin and see, for example, whether vascularization is the densest closer or further from the surface.
The statistical analysis performed on the extracted data shows that, overall, 37 features change significantly across timepoints ($p<0.05$), with very strong evidence ($p<0.0001$) for most of them. They describe the vascular architecture as denser, more chaotic and branched, and closer to the surface in lesions than in healthy skin. Moreover, the post-hoc analysis indicates that pairs of timepoints often statistically differ from each other. The foundations for the reproducibility of these results are laid by developing two Python pipelines: one semi-automatic class for the initial computing of the OCTA volume, and an automatic pipeline for feature extraction.
This thesis shows that quantitative parameters extracted from OCTA data allow to distinguish different timepoints in NMSC lesions treated with radiotherapy, with a consistent trend of features converging back to healthy values. Notably, these features are easily interpretable and give direct insight into the evolution of the lesions. Future studies that combine these vascular features with radiomics might be able to provide methods that accurately and non-invasively predict NMSC lesions' response to treatment, thus opening the door to personalized radiotherapy for each patient.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Radiotherapy for non-melanoma skin cancer

Non-melanoma skin cancers (NMSCs) are the most common malignant tumor among fair-skinned people, with a growing incidence worldwide [1]. The term is mainly used to define [2]:

- **Basal-cell carcinomas (BCCs)**: slow-growing, locally invasive epidermal tumors with a metastatic rate of $< 0.1\%$. They are associated with intermittent and childhood sun exposure.
- **Squamous-cell carcinomas (SCCs)**: they arise from dysplastic epidermal keratinocytes, and have a higher metastatic rate of 0.3% to 3.7%. They are associated with chronic UV exposure.

BCCs and SCCs represent 99% of NMSCs [3].
Surgery is the most common practice to treat NMSC. However, whether exclusive or adjuvant, radiotherapy may be preferred for older patients with comorbidities, and tumors involving a sensitive part of the face that, with surgery, may result in an unacceptable functional or cosmetic deficit [3] [4].
Radiotherapy is a treatment modality that delivers ionizing radiation towards tissue to destroy tumor cells, while sparing healthy ones as much as possible (Figure 1.1). The radiation impairs cancer cells' ability to reproduce and kills them by damaging their DNA structure. Fortunately, healthy cells are generally less susceptible to radiotherapy

than cancerous ones, since the latter have a high replication rate and defects in their DNA damage response pathways [4] [5].

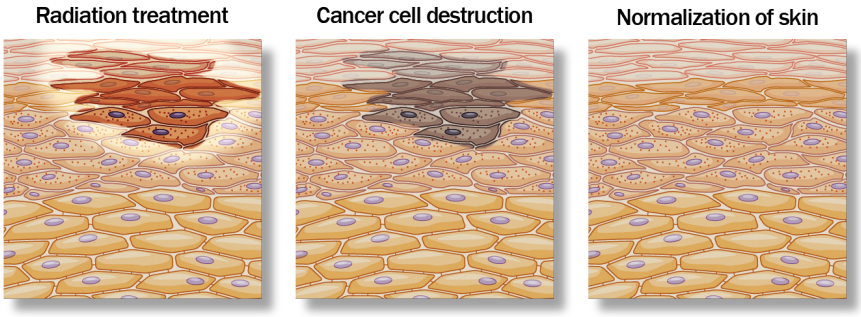


Figure 1.1: Representation of the effect of radiotherapy on cancer cells. Taken from [6].

Nevertheless, the primary aim of radiotherapy is to achieve optimal tumor control probability (TCP) while minimizing the risk of normal tissue complications (NTCPs). This can be achieved by correctly localizing the target and accurately planning the fractionation.
NMSCs are radioresponsive: radiotherapy has shown local control rates exceeding 90-95%, irrespective of the dose or dose per fraction, and optimal cosmetic results [4]. Therefore, radiotherapy is often recommended as a tissue-preserving option, especially for lesions that are localized and in early stages.
Different tumors have different vascular architectures, which also vary at different stages of the same lesion and as a result of treatment. Therefore, the recognition and quantification of distinctive features in these different vascular structures may help in diagnosing and monitoring NMSC, as well as in radiotherapy treatment planning and follow-up.

1.2 Overview on OCT and OCTA

Optical coherence tomography (OCT) is an imaging method based on the optical scattering of internal tissue micro-structures, which are discriminated at their interfaces due to their different refractive indices

[7]. By decorrelating OCT signals that are repeatedly acquired at the same location, light scattered from moving erythrocytes is contrasted with respect to the static background. This functional extension of OCT visualizes microvasculature and is called optical coherence tomography angiography (OCTA). These techniques can be applied in dermatology, and thus present a non-invasive alternative to traditional histology and offer the possibility of carrying out longitudinal studies [8].

1.2.1 Theoretical principles behind OCT

OCT uses the delays of back-reflected light in the near-infrared (NIR) to resolve the depth at which the waves are reflected. NIR light travels too fast to be measured directly, so a reference measurement is done through an interferometer.
Low-coherence interferometry is the underlying principle of all OCT modalities, as the broadband light source it uses has, by definition, low temporal coherence. This means that two waves coming from the same optical source maintain a fixed phase relation only within a short timeframe, corresponding to a low coherence length.
The beam produced by the optical source is split between the reference arm, where it's reflected by the reference mirror, and the sample arm of the system, where it's reflected by the sample itself. When these two reflected beams are recombined, their waves superimpose and an interference signal is detected. Interference is constructive when the two beams reinforce each other, and destructive when they cancel each other out. The interference signal can be obtained only if the path lengths of the two arms are equal within the coherence length of the source, which is in the order of micrometers.
The axial resolution (δz) in air of OCT equals the round-trip coherence length, which is half the standard coherence length [9][10]:

$$\delta z = \frac{2ln(2)}{\pi} * \frac{\lambda_0^2}{\Delta\lambda_{FWHM}} \tag{1.1}$$

where λ_0 is the central wavelength and λ_{FWHM} is the wavelength range of the source, defined as the width at the intensity level equal to half the maximum intensity [9].
The lateral resolution δx depends on the spot size of the laser beam, which, for a gaussian profile, is the radius w_0 where the beam intensity drops at $\frac{1}{e^2}$. It can be thus calculated as a function of the beam diameter at FWHM:

$$\delta x = \sqrt{2ln(2)} * w0 = \sqrt{2ln(2)} * \frac{2\lambda_0}{\pi} \frac{f_{sys}}{n * d} \tag{1.2}$$

where f_{sys} is the focusing length, n the refractive index of the media and d the diameter of the collimated incident beam at the focusing lens [9]. A better lateral resolution can be achieved, but at the cost of a shallower depth of focus.

1.2.2 OCT modalities

Two main categories of OCT systems exist: time-domain (TD) and frequency domain (FD) systems. They differ in the optical source, in the usage of the reference mirror and in the chosen detector. Ultimately, these modalities differ in how they are able to reconstruct the depths of the reflective interfaces from the detected signal.
TD-OCT systems employ a broad-band light source, a scanning reference mirror and a simple photodetector, as shown in Figure 1.2. The position of the mirror is mechanically changed to match the optical path length of the back-reflected beam in the sample arm. The interference signal is present when the displacement of the reference mirror is equal, within the coherence length, to the depth at which the reflection happened in the sample. This allows the depth to be obtained.

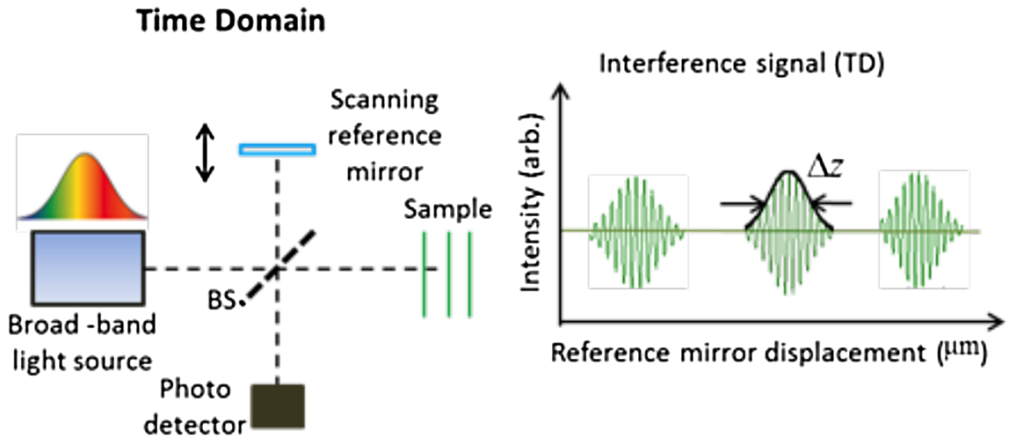


Figure 1.2: On the left, schematic representation of TD-OCT. On the right, graph of the intensity of the interference signal (green) with respect to the displacement of the reference mirror. Δz indicates double the coherence length (equal to the axial resolution). Taken from [10].

FD-OCT systems are further divided into two types: spectral domain (SD) and swept source (SS) OCT, both shown in Figure 1.3.

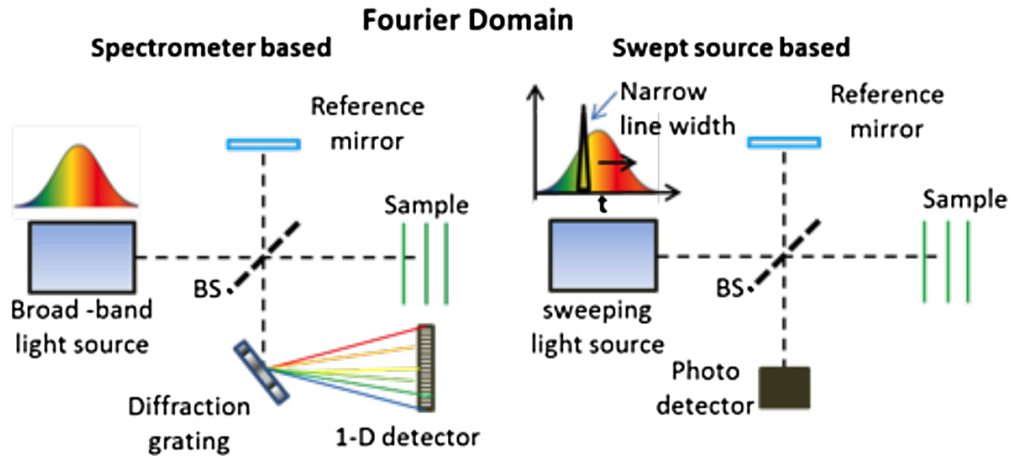


Figure 1.3: Schematic representation of the two FD-OCT systems. Taken from [10].

Different wavelengths penetrate at different depths, with longer ones reaching deeper and not being scattered as much. The interference signal presents a cosinus for each reflective element in the sample, with the modulation periodicity depending on its depth and

the amplitude proportional to the reflectivity of the layer. Therefore, the depths can be reconstructed by computing the fast Fourier transform (FFT) of the signal: there's one peak for each reflection, with the height proportional to the intensity of the reflection and the position corresponding to its depth [11]. Figure 1.4 shows a simple example, where the sample has only one interface, such as a mirror: the symmetric additional peaks are mirror images.

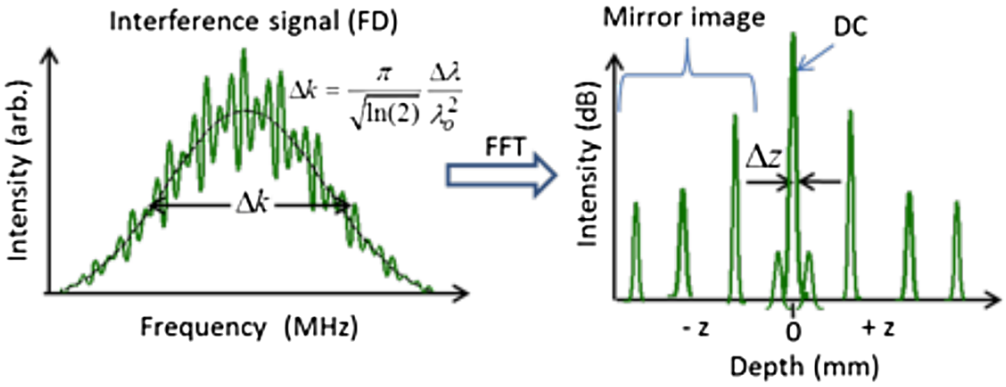


Figure 1.4: Graphs representing the application of the FFT on the interference signal obtained in FD-OCT, where Δk represents the spectral bandwidth and is proportional to the axial resolution (Eq. 1.1). Taken from [10].

SD-OCT employs a broad-band source, a stationary reference mirror and a spectrometer. The spectrometer disperses the beam into its spectral components with a diffraction grating, thus recording the spectrum. Detection is done with a camera, a 1D charged-coupled device (CCD) array. The FFT of the signal is performed on the workstation. On the other hand, SS-OCT employs a sweeping light source, a stationary reference mirror and a simple photodetector. The light source consists of a rapidly-tunable laser that can illuminate at a narrow wavelenght, eliminating the need of a diffraction grating. The signal that is processed by the FFT is a combination of all the signals detected at the chosen wavelengths.

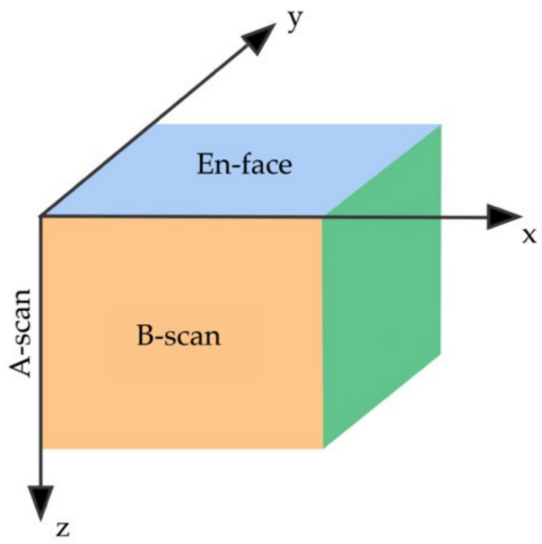


Figure 1.5: Graphical representation of acquired OCT data. Modified from [12].

As shown in 1.5, an OCT volume is composed by the combination of B-scans, which, in turn, are the combination of A-lines, which are depth-resolved reflectivity profiles of the sample [10]. An en-face image is a representation of the volume from above, created by flattening the depth information, and is perpendicular to the A-lines. An OCT volume can also be divided into cross-scans, which are orthonormal sections to the en-face and the B-scans.

	TD-OCT	**SD-OCT**	**SS-OCT**
Light source	Broadband	Broadband	Sweeping
Detector	Photodetector	Spectrometer	Photodetector
Reference mirror	Moving	Stationary	Stationary
Depth reconstruction	Mirror displacement	FFT	FFT

Table 1.1: Quick overview of the main differences between the three OCT modalities: time domain (TD), spectral domain (SD) and swept source (SS).

A general overview of the OCT systems can be seen in Table 1.1. TD-OCT is the system that takes the longest to gather the data, as it requires for the reference mirror to be moved many times to

reconstruct a single A-line. SS-OCT is thus preferred, as it's much faster and doesn't employ as complex of a detector as SD-OCT does. This is crucial, since OCTA is based on the premise that the only thing that moves between the acquisitions are blood cells. If a patient moves, which happens more often than not, having a faster acquisition time means reducing the effect of movement artifacts.

1.2.3 OCTA techniques

As previously said, an OCTA image is computed by decorrelating two or more OCT images taken at the same location in quick succession. The motion contrast can be obtained by an intensity-based, a phase-based or a complex-based technique. The first two evaluate, respectively, differences in the intensity or phase of subsequent OCT volumes. For a number N of B-scans acquired at the same position y, the angiographic volume is computed as follows [10]:

$$A(x, y, z) = \frac{1}{N-1} \sum_{i=0}^{N-1} |\Delta T(x, z)_i| \tag{1.3}$$

with $\Delta T(x, z)$ representing the i-th difference in logarithmically-scaled intensity or phase between A-lines repeatedly acquired at the same position y. Phase-based OCTA tends to have higher contrast than the other. However, the intensity-based technique may be preferred for its simplicity and the lower sensitivity to motion artifacts.

1.3 Purpose of the study

Reconstructing skin microvasculature in 3D is of interest for NMSC diagnosis, with many techniques used for this purpose. These, however, present significant shortcomings. High resolution episcopic microscopy (HREM) provides a very accurate 3D reconstruction, but requires excision. Computed tomography angiography (CTA) is accurate and non-invasive, but uses ionising radiation and requires a contrast agent.

Dermoscopy is a non-invasive, label-free, and clinically established technique and provides accurate information without using harmful radiation. However, its penetration depth is up to 200 μm, which is too shallow for a complete evaluation of the vascular architecture of certain lesions [7][13].
OCTA is fast, non-invasive, label-free and penetrates tissues with electromagnetic wavelengths that don't cause any damage. It can accurately reconstruct vascular micro-structures and resolve individual vessels, reaching depths of up to 1.5 mm. Therefore, it offers a valid alternative to the other available techniques and could be used to aid in NMSC diagnosis and, in longitudinal studies, for radiotherapy treatment planning and follow-up. In fact, OCTA has already found many clinical applications, especially in ophthalmology, where it's used for diabetic retinopathy, macular degeneration, and glaucoma, to name a few [12].
The objective of this thesis is to understand whether feature extraction on OCTA data can quantitatively assess lesion-dependent changes in skin microvasculature, as a result of radiotherapy. In order to do so, lesions are imaged before and after the start of radiotherapy treatment, thus conducting a longitudinal study. A statistical analysis assesses whether these parameters change significantly over the course of the 6-month-long follow-up. It's also crucial for these values to progressively get closer to the healthy baseline, which would arise from radiotherapy's effectiveness. Afterwards, the significant features are commented and contextualized, in order to understand whether they can give descriptive insight into the mutating vascular morphologies. This is particularly important to further solidify OCTA as a technology that can provide clear and interpretable information, that can be useful in decision-making in clinical settings.

Chapter 2

Materials and Methods

2.1 The dataset

The dataset was obtained in an IRB-approved study (Medical University of Vienna, no. 1246/2013) conducted at the Center for Medical Physics and Biomedical Engineering of the Vienna General Hospital. A total of 20 patients, affected by BCC, SCC, or actinic keratosis (AK), were recruited. AK is a precancerous lesion that can develop into SCC. The diagnoses were confirmed with biopsy, and all patients, once surgery was deemed inapplicable, were treated with superficial orthovoltage radiotherapy (Section 2.2).

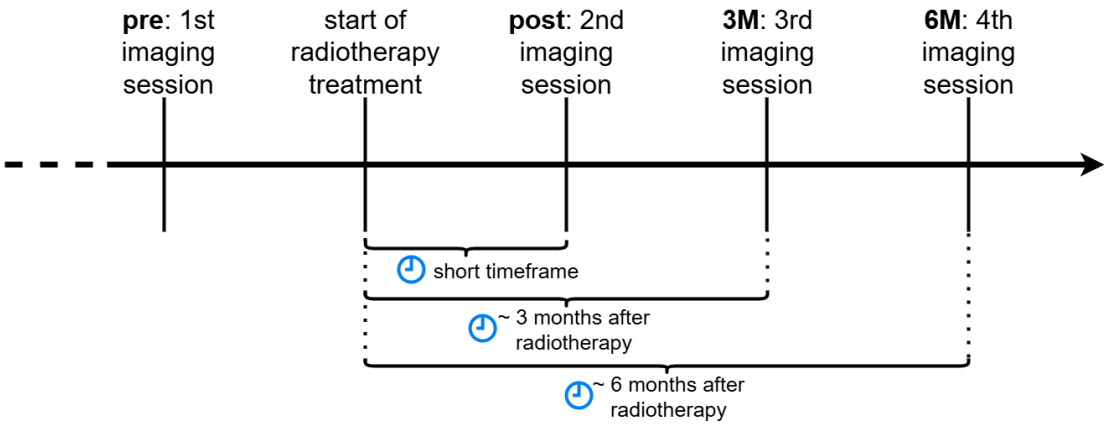


Figure 2.1: Timeline of the imaging sessions. Healthy volumes may be acquired at any session.

Each patient that enrolls in the study shows up for imaging sessions before, right after, 3 months after, and 6 months after the start of radiotherapy treatment (Figure 2.1). This allows to follow the evolution of vasculature morphology and verify if OCTA can be used to quantify these changes.
Initially, the dataset contained 293 volumes, across all timepoints, from which the ones with the highest quality were selected for the segmentation process. The final dataset used in this thesis is thus composed by a total of 146 volumes, detailed in Table 2.1. Notably, not all patients were monitored throughout all timepoints. This is due to two factors:

- patients were not recruited at the same time, so some of them are still in earlier treatment stages;
- some patients withdrew from treatment, so further acquisitions were not possible.

Moreover, since acquisitions of poor quality have not been selected for the final dataset, not all patients have a corresponding healthy volume.

Patients					
Total (#)		**Lesion type**			
20		BCC, SCC or AK			
Only pre-post (#)		**Only pre-3M (#)**		**All timepoints (#)**	
7		6		7	
Acquisitions					
Total (#)	**Healthy (#)**	**Pre (#)**	**Post (#)**	**3M (#)**	**6M (#)**
146	18	43	40	27	18

Table 2.1: Overview on the acquired dataset. Pre: before radiotherapy; post: right after radiotherapy; 3M: 3 months after radiotherapy; 6M: 6 months after radiotherapy.

2.2 The radiotherapy treatment

Orthovoltage radiotherapy is ideal for superficial lesions (< 5 mm), as it uses low-energy X-rays, from 80 kVp to 250 kVp.
For the enrolled patients, treatment was delivered on the skin surface with a Gulmay D3300 orthovoltage therapy unit, while applying a custom lead shielding to protect healthy skin surrounding the lesion. The properties of the unit are showed in Table 2.2.

Gulmay D3300 orthovoltage therapy unit			
Target material	**Focal spot**	**Filtration (Be)**	$\frac{\alpha}{\beta}$
W	8 mm	3 mm	10 Gy

Table 2.2: Main properties of the radiotherapy unit. Be: beryllium; W: tungsten.

The clinical target volume (CTV) included a margin of 5 mm to 20 mm around the visible lesion, as tumor cells may still be present outside the diagnosed area. The fractionation of the dose is individual, e.g. 36 Gy over 6 treatment sessions ($\frac{36\ Gy}{6\ fx}$) or 50 Gy over 20 sessions ($\frac{50\ Gy}{20\ fx}$).

2.3 The laboratory OCT system

A high-resolution OCT system based on a swept-source from Insight Photonic Solutions, USA, was used for the measurements. Its main operating parameters are displayed in Table 2.3.

High-resolution SS-OCT system					
Central wavelength	**Bandwidth**	**Lateral resolution**	**Axial resolution**	**Field of view**	**Depth**
1300 nm	29 nm	31.5 μm	27.3 μm	1 cm^2	1-1.5 mm

Table 2.3: Properties of the laboratory SS-OCT system.

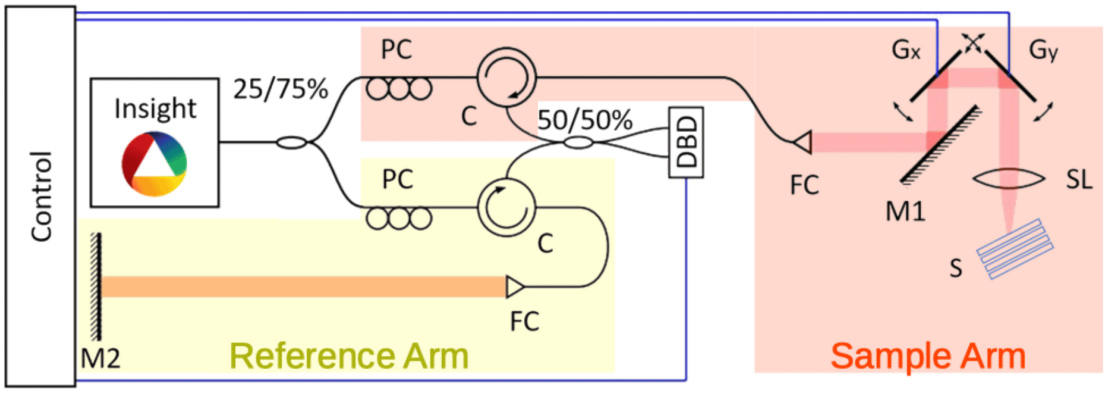


Figure 2.2: OCT setup. Insight: laser source; 25/75% fiber coupler; PC: polarization control; C: circulator; 50/50% fiber coupler; FC: fiber collimator; M1, M2: mirrors, with M2 being the reference mirror; Gx, Gy scanning galvanometers; SL: scan lens; S: sample; DBD: dual-balance detector [14]. Taken from [14].

See Figure 2.2 for a schematic representation of the laboratory system, which functions as follows [14]:

- **Beam splitting**: light from the laser source is split into two beams by a fiber coupler, with a proportion of 25% of the optical power to the reference arm and 75% to sample arm.
- **Sample arm**: the beam that travels through the sample arm is released into free space by fiber collimators (FC), towards the object that needs to be investigated. Two scanning galvanometric mirrors move the beam through the surface of the sample. More precisely, one of them allows to scan every A-line of a single B-scan, and the other one allows to move between different B-scans. Each time, 4 B-scans are acquired at the same location, before moving to the following one.
- **Reference arm**: the beam reaches the reference mirror, which reflects it back.
- **Beam recombination**: the two back-reflected beams are redirected by two circulators towards a 50/50% fiber coupler for recombination.

- **Detection**: lastly, a dual-balance-detector (DBD) records the interference signal and a data acquisition (DAQ) card digitizes it.

2.4 Acquisition protocol

The acquisition protocol is divided into a series of tasks that can be grouped on whether they're performed before, during, or after the acquisition takes place (Figure 2.3).

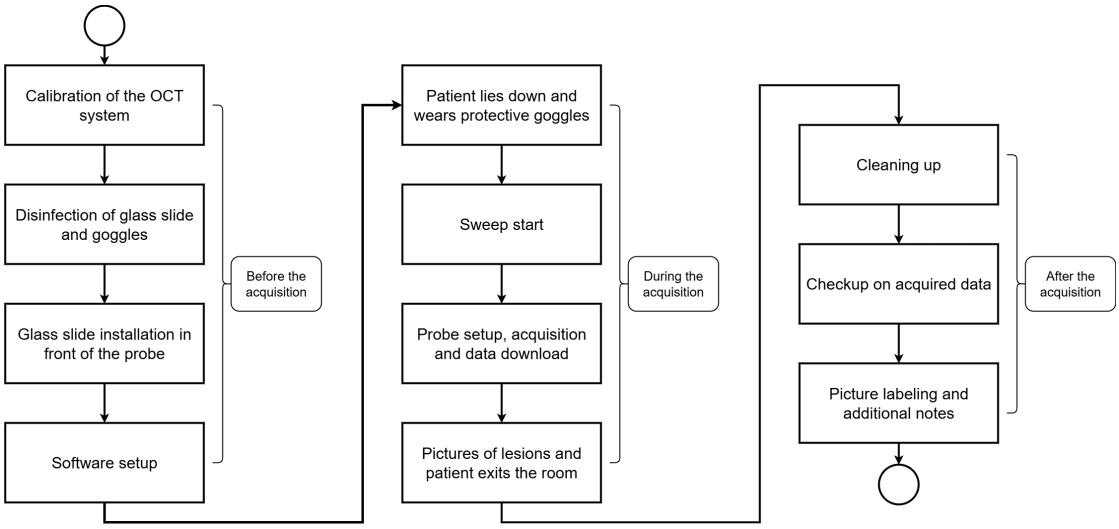


Figure 2.3: Pipeline of the acquisition process.

2.4.1 Before the acquisition

A series of tasks has to be carried out before the patient arrives:

- **Calibration**: in order to identify valid sweep points, the system is calibrated with the dedicated Insight Photonics laser control program. The calibration produces a data valid vector (DVV) file, which contains the indices of the reliable A-line points. The DVV file is crucial during processing, and without it image quality gets heavily compromised. Calibration needs to be done only once a day, unless experiments with different settings are performed between patients.

- **Disinfection**: all objects the patient will be in touch with are disinfected. These are the glass slide, the support for the glass slide, the protective goggles, and two cushions for the legs, for patient comfort. Normally, the examination table the patient lies on is not disinfected, as a suitable waxed paper roll is used as a cover, and replaced for each patient.

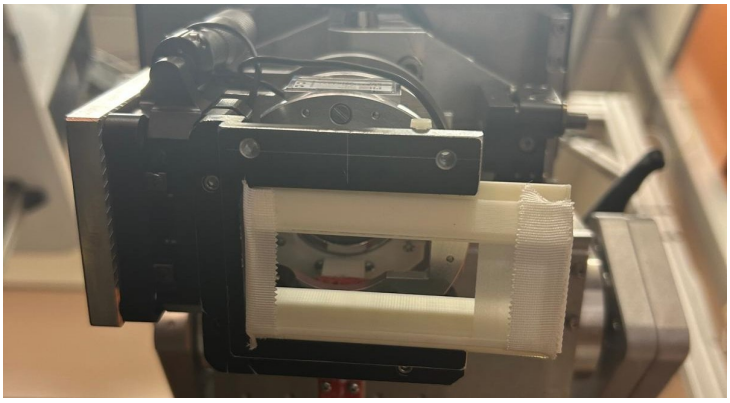


Figure 2.4: Picture showing the glass slide after being attached in front of the scanning lens. The 3D-printed support is not in contact with the lens itself, but is kept at a distance. Glass slide: transparent, but for an opaque band; support: white; medical tape: white, ragged texture.

- **Glass slide insertion**: the disinfected glass slide (Figure 2.4) is attached with medical tape to its 3D-printed support made of a bio-compatible material, which allows it to be firmly placed in front of the scanning lens (SL). The glass slide flattens the skin in a stabilized position. Even if skin surface is never exactly parallel to the en-face of the OCT volume, reduced variability in the sample-SL distance is very important. As the depth of focus is fixed during acquisition, a flattened skin surface leads to a more consistent quality across slices in the acquired volume. It also makes it easier for the operator to focus the laser at the desired depth. Moreover, the glass slide helps to attenuate undesired reflections.
  However, the glass slide can be utilized only if the placement of the lesion allows total adherence with the skin: it's impossible

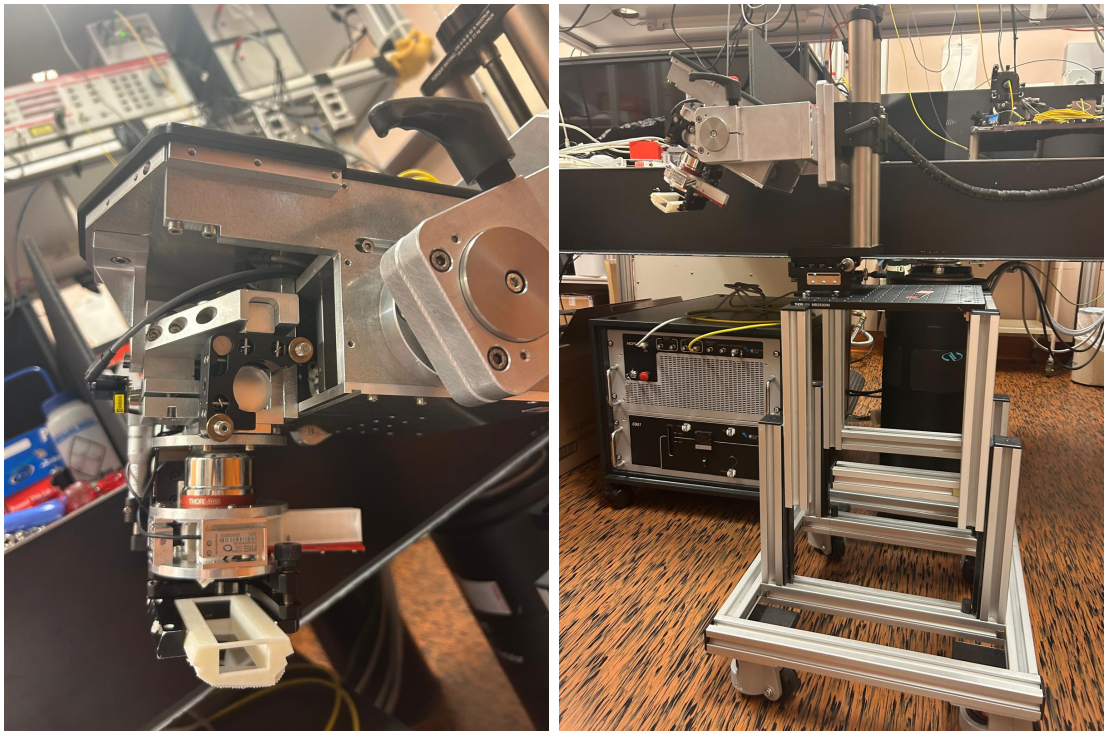
to place it on the nose, and similarly curved structures. In these situations, the acquisition is carried out in any case, at the cost of additional reflections on the surface and diminished resolution at the borders of the volume.

- **Software setup**: the computer station has two screens. On the left, the laser control program has already been started for the calibration process. On the right, a laboratory custom-made LabView software (Figure 2.6) is used to control the acquisition process.

2.4.2 During the acquisition

The patient lies down on the examination table, in a supine position or laterally, depending on the placement of the lesion. Once the patient is lying down, they wear protective goggles to comply with laser safety regulations, even if the laser is not used in proximity of their eyes. Then, the acquisition process can start:

- **Sweep start**: firstly, it's ensured that the patient is correctly wearing the protective goggles. Then, the laser control program is used to start the sweep, which means that, at that moment, the laser beam starts exiting the probe.
- **Probe placement**: as the imaging system is mounted on a wheeled cart (Figure 2.5), it can be moved around the bed to get in the vicinity of the lesion. Moreover, the cart can be lifted in the vertical direction, and a series of levers and knobs on the system itself allow to fine-tune positioning with careful rotations, back-and-forth movements and further up-and-down sliding. Thus, the probe is placed in the desired position, with the glass slide adhering to the skin. If the slide is not used, there's no contact whatsoever with the skin.
- **Water**: air pockets are always present between skin and glass, so distilled water is used to fill them, as it has a refractive index similar to that of tissues. This reduces unwanted reflections.



(a) Side view of the OCT probe. **(b)** The OCT probe mounted on the cart.

Figure 2.5: Pictures of the imaging setup right before acquisition.

- **Fine-tuning**: The LabView software (Figure 2.6) allows to visualize a live preview of a B-scan of the current investigated volume with background subtraction. Moreover, it allows to switch between fixed B-scan positions. This is important for a series of reasons:
  - to understand whether, throughout the volume, there's appropriate skin-glass contact and thus whether more pressure should be applied;
  - to fine-tune the position of the probe with knobs and levers, so that the imaging volume is centered at the desired skin volume. This is particularly important if no glass slide is used;
  - to ensure that image quality is satisfactory: more water may need to be added, the position of the probe may need to be

fine-tuned, or a setting in the interferometer may need to be changed.

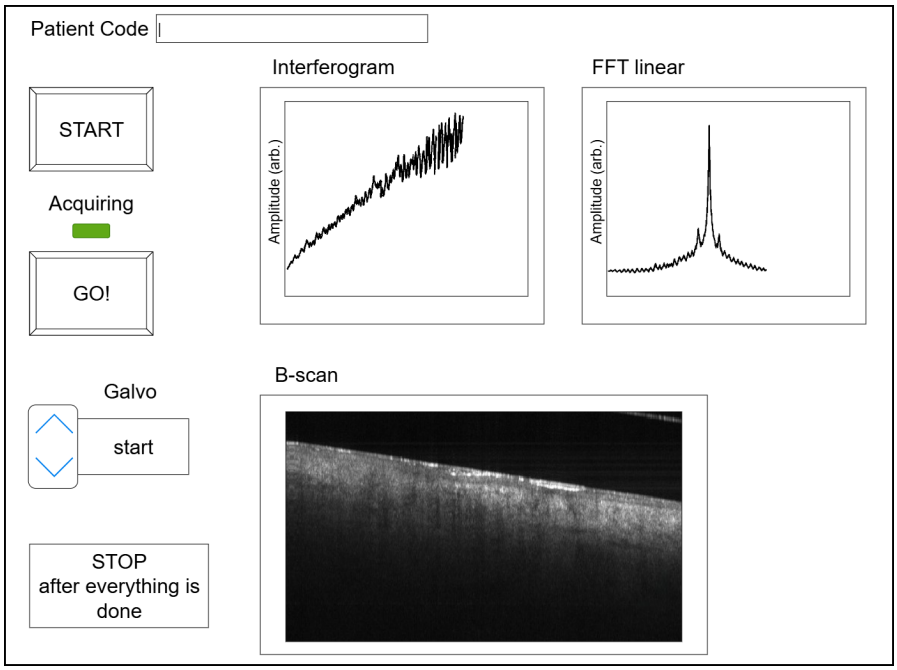


Figure 2.6: Simplified schematic of the interface of the custom LabView software. Only the components cited in the text are present. START: starts showing data in the three plots; Acquiring: turns on while acquiring; GO!: starts the acquisitions; Galvo: allows to shift between different B-scans in the preview; B-scan: live B-scan preview; Interferogram: plot of the interferometric signal; FFT linear: plot of FFT of the interferometric signal; STOP: raw data downloading can be lengthy - if little time passed from the previous acquisition, this button is pressed so that the new downloading starts after the previous, without interrupting it; Patient Code: label of folder in which data will be saved.

- **Acquisition**: once everything has been fine tuned, the acquisition is commenced through the LabView software. At the same time, the data is downloaded into the corresponding folder of the computer, as labeled in LabView. A total of 4 volumes are imaged in the same position in around 20 seconds. Adding the time needed to fully download the data, it takes 1-2 minutes to complete an acquisition.

This process, except for the sweep start, is repeated for every lesion. Notably, some lesions have a larger extension than the 1 cm^2 FOV, so more positions of the same lesion are often taken. Moreover, especially for acquisitions lacking the glass slide, more imagings may be done without moving the probe at all, and the best one is later picked out. After the desired lesions, or healthy skin, are imaged, the sweep is stopped through the laser control program. Then, pictures of each investigated surface are taken with a ruler alongside as reference. Only then, can the patient take off the protective goggles and exit the examination room.
It's important to note that, ideally, two operators are in the examination room: one managing the OCT system and the other operating the softwares.

2.4.3 After the acquisition

A series of tasks has to be carried out after the patient leaves:

- **Cleaning up**: the glass slide is thrown away in the appropriate container. Then, the protective goggles, the 3D-printed glass slide support, and the leg cushions are disinfected. Lastly, the waxed paper on which the patient lied down is ripped and thrown away, and a replacement is placed on the bed for the next patient.
- **Checkup on acquired data**: after ensuring it's all there, it's moved to a specific folder alongside the rest of the dataset. It's fundamental to move the DVV file in the same folder as the OCT data.
- **Picture labeling**: the imaged sections are highlighted and labeled, so that these pictures can be cross-referenced with the folder names in which the relative acquisitions are stored, as seen in Figure 2.8. This is particularly important if different positions of the same lesion were imaged.
- **Additional notes**: `about.txt` file that contains useful information, whether it's important to interpret the acquired data or

patient-specific tips for the next imaging session.

The process is concluded when the current date folder appears as shown in figure 2.7.

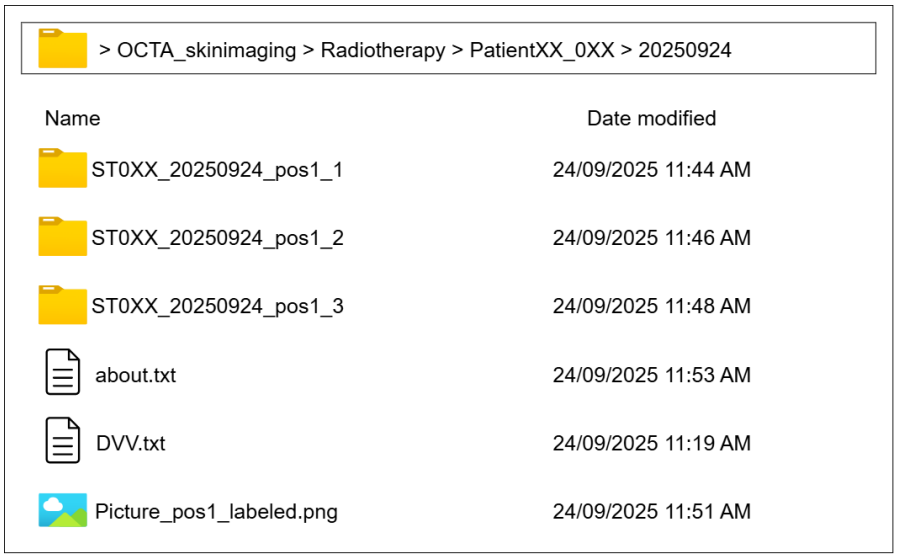


Figure 2.7: Typical structure of an acquisition folder. The main dataset folder contains all patients' folders, which are divided into timepoints. Each timepoint corresponds to an acquisition folder. Each `ST0XX_20250924_posX_X` folder contains the OCT data acquired in that location.

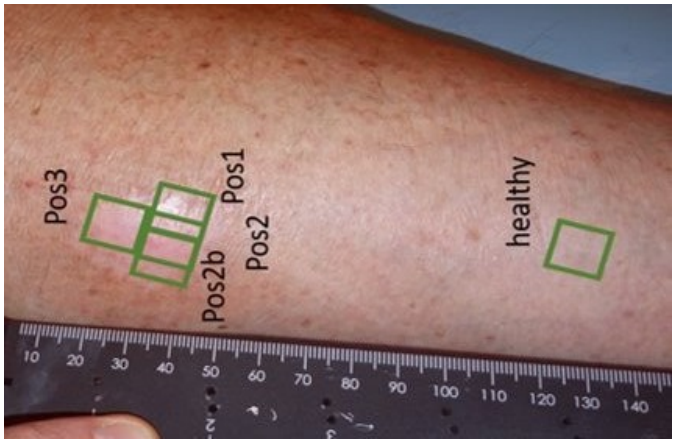


Figure 2.8: Example of labeled picture.

2.5 Processing protocol

The raw OCT data is divided into 32 files named `B_scans0XX`. Every file is a 1D array of size: 2176 * 512 * 64, where 2176 is the length of an A-line, 512 the width of a B-scan, and 64 the number of scans per file. The volumes have a lateral pixel spacing of 0.0196 mm and an axial pixel spacing of 0.0137 mm.
The processing protocol is composed of four sections:

- **Data processing**: each file of raw data is reshaped, processed and combined so that it turns first into 4 separate OCT volumes and then into a single OCTA array.

- **Segmentation**: the OCTA arrays are turned into binary masks, where the vessels are equal to 1 and everything else is equal to 0.

- **Feature extraction**: performed on the segmentation. The parameters are saved into Excel files, one for each timepoint.

- **Statistical analysis**: to determine which features are able to differentiate between different timepoints.

2.5.1 Data processing

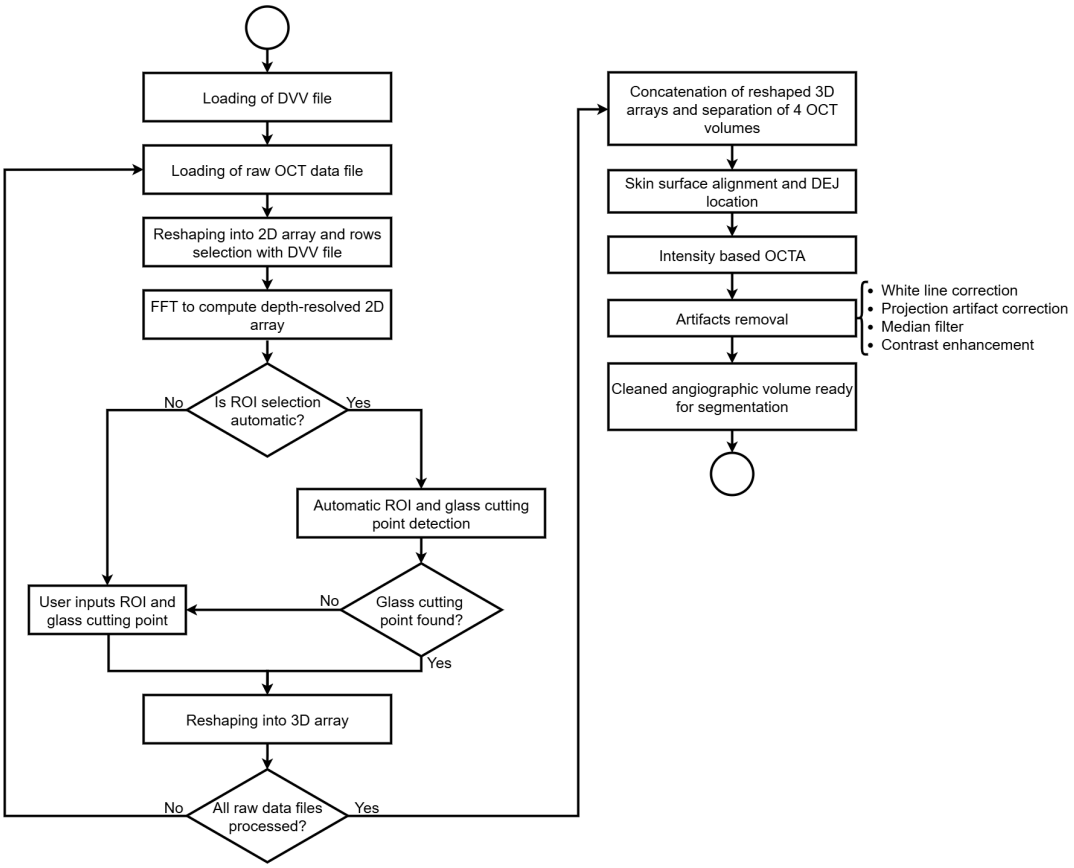


Figure 2.9: Pipeline of the processing of raw OCT data into an OCTA volume.

Previously, data processing was being done in Matlab. However, for scalability and the possibility of running the pipeline with a GPU, everything was adapted and optimized in Python, with a few improvements and additions. The pipeline can be visualized in Figure 2.9. The custom `OCTAProcessing` class has four inputs:

- **`variable_name`**: the initial part of the files that contain the raw OCT data. In our case, it's set as `B_scans`.

- **`automatic_ROI_and_glass`, `ask_confirmation`**: these flags, if set to `True`, allow for, respectively, the automatic determination

the region of interest (ROI), and the possibility of asking the user if they're satisfied with the chosen ROI, automatic or not.

- **`patient_files_path`**: the path to a JSON file, which has to be structured as follows:

```
[
    {
        "pat_number": "ST0XX",
        "pre": ".\\Radiotherapy\\PatientXX_0XX\\20250101",
        "post": ".\\Radiotherapy\\PatientXX_0XX\\20250121",
        "3M": ".\\Radiotherapy\\PatientXX_0XX\\20250421",
        "6M": ".\\Radiotherapy\\PatientXX_0XX\\20250721",
        "9M": ".\\Radiotherapy\\PatientXX_0XX\\20251021"
    }
]
```

  This allows to put in the same file all timepoint paths of all patients, and builds the foundations for an automatic, or semiautomatic, pipeline. There is no specific path for healthy acquisitions, as they are done during one, or more, of the five visits. A 9M path is also present, as some patients may also come 9 months after radiotherapy. However, since very little data is present regarding this timepoint, it is not considered throughout this thesis, and only the other ones are mentioned.

The first function of the pipeline is `ppOCT3D`, which does the initial processing of raw OCT data, schematized in Figure 2.13. This function is structured as follows:

- **Reading the DVV file**: `DVV.txt` is read into a 1D array, where each element is an integer number corresponding to a valid A-line index.

- **`load_intensity_image`**: this function reads every 1D raw data file and reshapes it into a 2D array, so that every column is an A-line. Then, using the DVV file, the valid rows are selected.

- **`process_intensity_image`**: this function receives the output of

the previous one. After subtracting to each row its mean for background removal and after applying zero-padding, it transforms the image into a depth-resolved 2D array, through an inverse FFT.

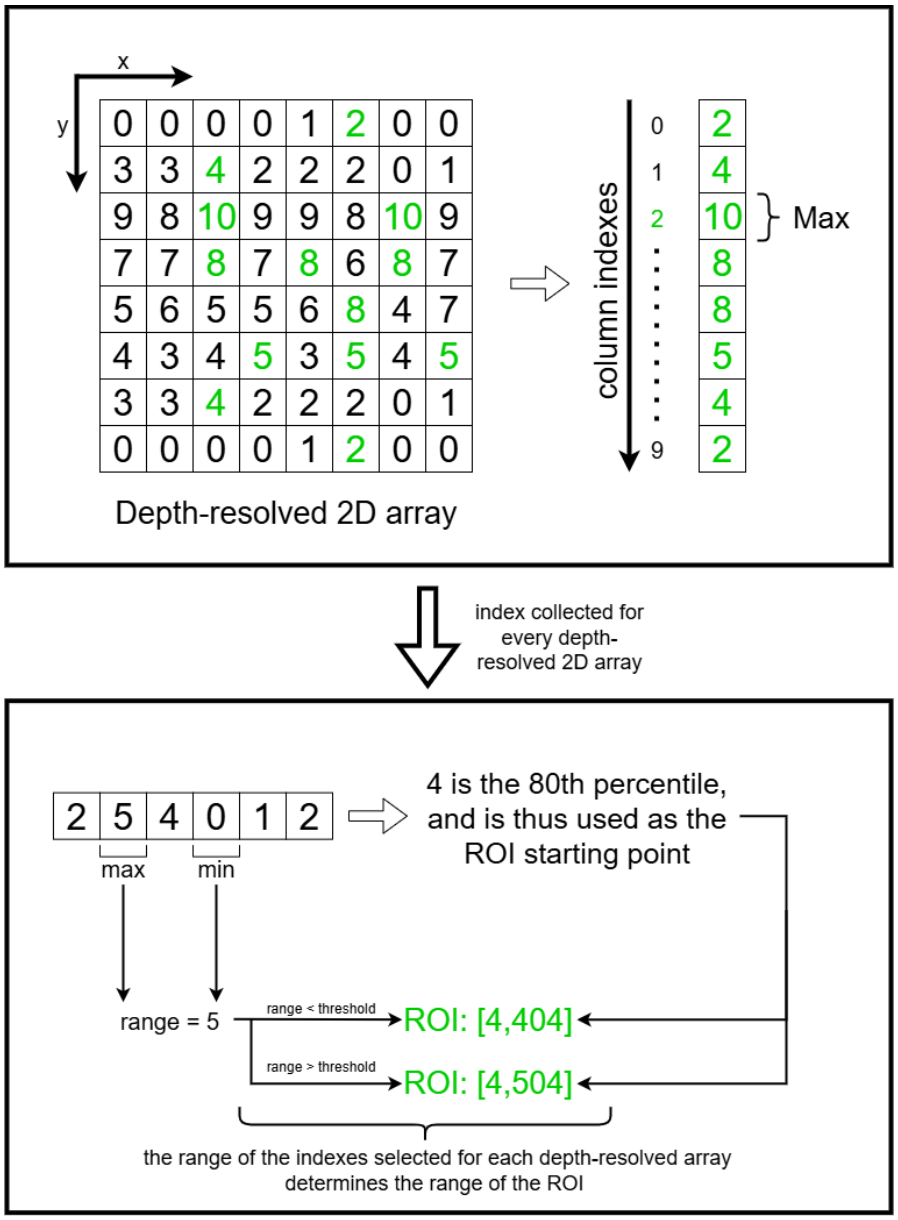
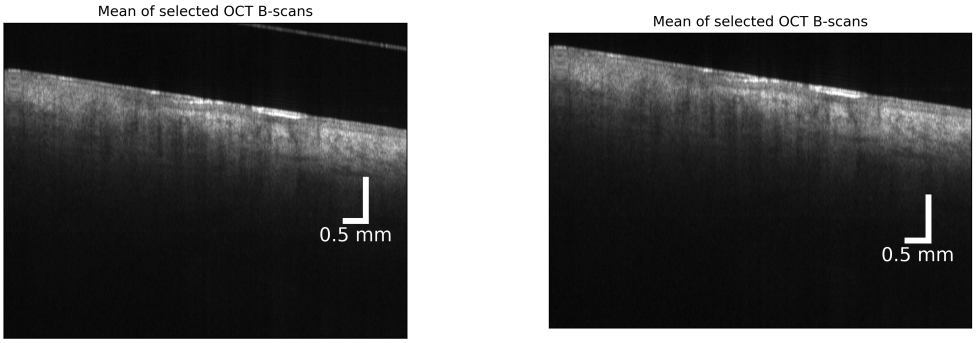


Figure 2.10: Schematic representation of the initial automatic ROI selection.

- **Automatic ROI selection**: firstly, for each output of the previous step, a column is created, where each element is the maximum value inside that row of the 2D array. As this array is also used later, it will be called `max_column_array`. Then, the index of the maximum value of that column is found. Of all these indexes, one for each raw file, the 80th percentile is taken as the initial index of the global ROI. The thickness of the ROI is either 500 or 400 pixels, depending on whether the range of the preliminary start values is more than 200 pixels or not. This takes into consideration the possibility of a very sloped skin surface, which requires a thicker ROI. The process can be visualized in Figure 2.10.
- **Glass line removal**: if a glass slide is used during acquisition, its upper surface is going to be visible as a high-intensity line throughout some B-scans (Figure 2.11).



(a) **(b)**

Figure 2.11: Comparison of the averages of the same 10 B-scans, before (a) and after (b) the automatic selection of the glass cutting point.

The ROI must be adapted and lowered to eliminate it by finding a glass cutting point, otherwise it's impossible to align the skin correctly. For this purpose, the mean of all the `max_column_array` is computed. Then, it's smoothed with a low-pass Butterworth

filter. A peak search, restricted to those with a minimum height, is then conducted on its gradient. The choice of the depth cutting point is based on one of four situations:

- firstly, the variable `manual_glass` is a flag that, if set to `True`, asks the user to manually input the depth cutting point; this flag is initially defined as `False`;

- if only one peak is found, it's assumed that the glass was already removed in the initial ROI determination;

- if exactly two peaks are found, the first is assumed to be due to the glass. Then, the glass cutting point is set as the peak index plus 30 pixels;

- if there are more than two peaks, the choice is considered to be uncertain, and `manual_glass` is set to `True`. This is unless the second peak is equal to or more than 80% of the maximum peak, in which case the first peak is assumed to be due to the glass. Then, the glass cutting point is set as the peak index, if selected, plus 15 pixels;

- if none of this conditions are met, then `manual_glass` is set to `True`.

The process is schematized in Figure 2.12. Overall, let's say that the initial ROI is between rows 5 and 500, if the glass cutting point is found to be 15, then the final ROI is located either between 35 and 500 or between 50 and 500.

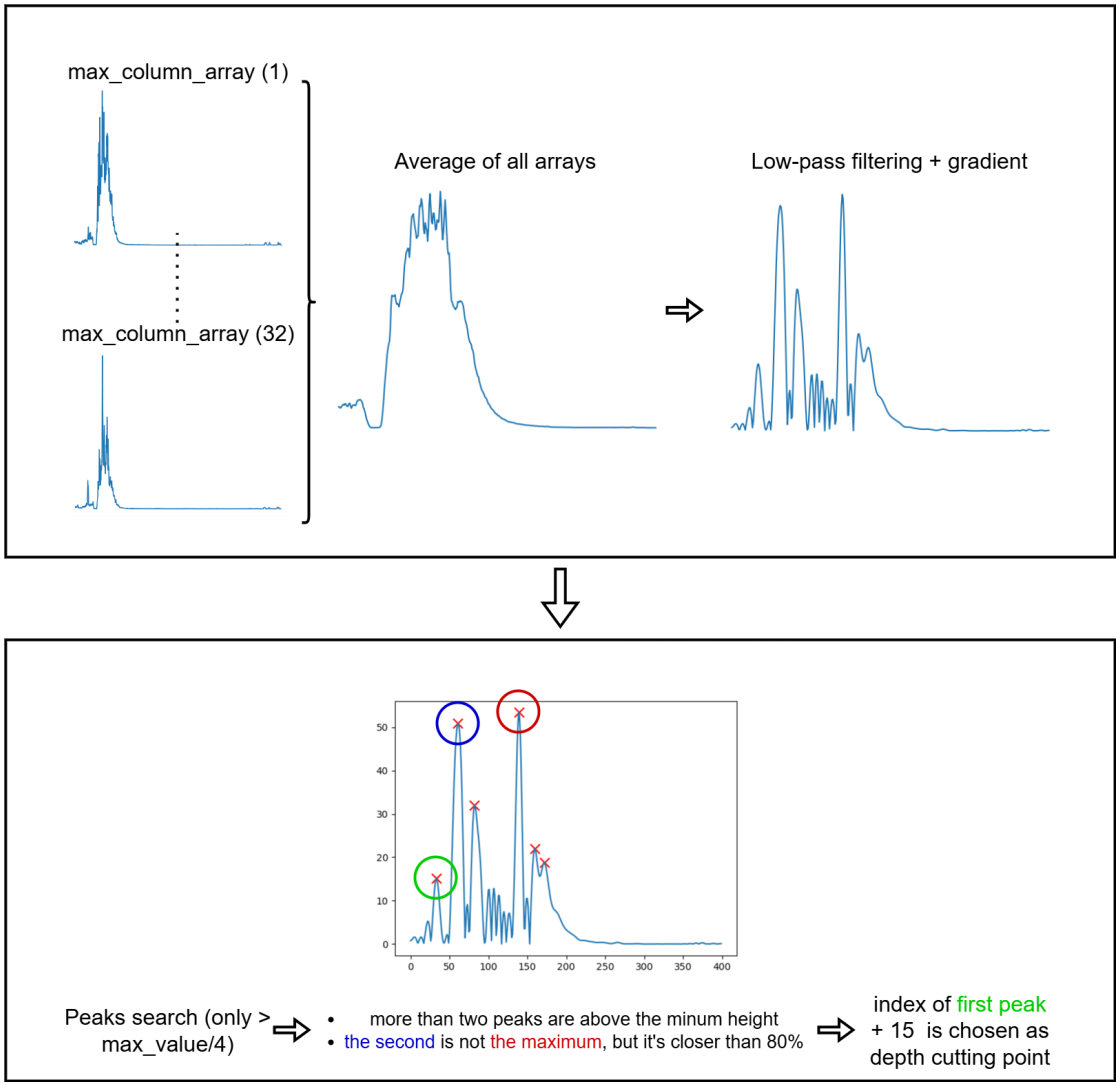


Figure 2.12: Schematic representation of automatic glass cutting point detection.

- **Row selection and reshaping**: in every depth-resolved array, only the rows contained in the final ROI are selected. Then they're reshaped into a 3D array, and each of these volumes is concatenated one after the other. This volume is defined as `all_slices_array`.

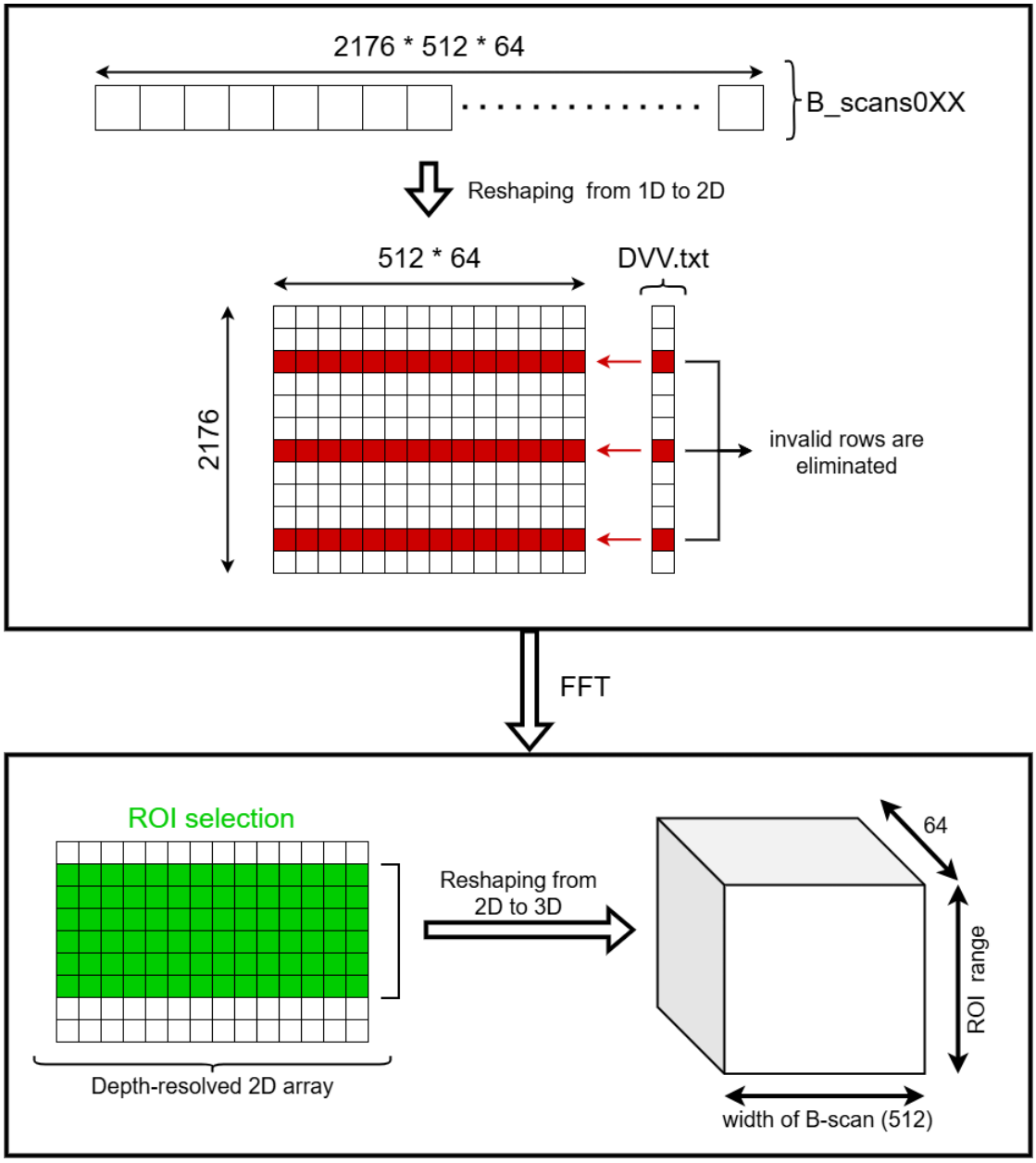


Figure 2.13: Scheme representing how each of the 32 raw data files are processed. These 32 obtained volumes are concatenated in the following steps.

The second main function of the pipeline is `intensity_based_OCTA`, where the OCTA volume is computed:

- **Separation of 4 OCT volumes**: computation done according

to the structure of `all_slices_array`, where one every four A-lines belongs to the same OCT volume, in sequence (Figure 2.14). Then, in each of these 3D arrays the first 30 A-lines are eliminated to account for the mirror artifact. It is caused by the inertia of the galvanometer mirrors, since the passage from the end position of a B-scan to the starting one of the next is not instantaneous.

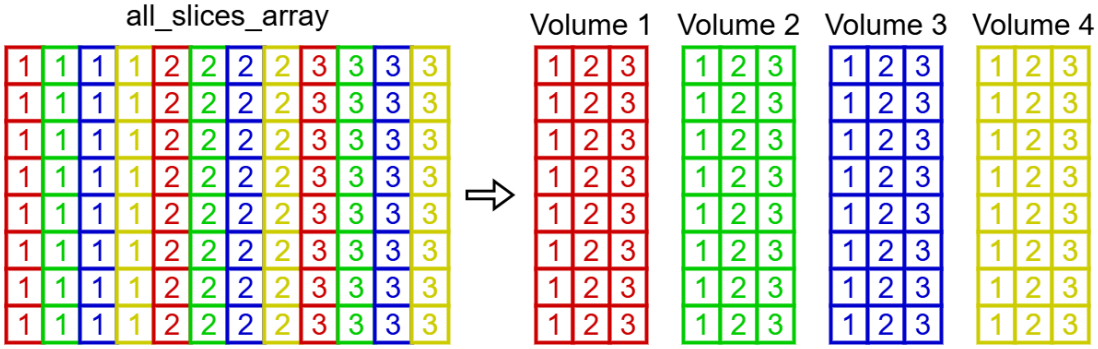
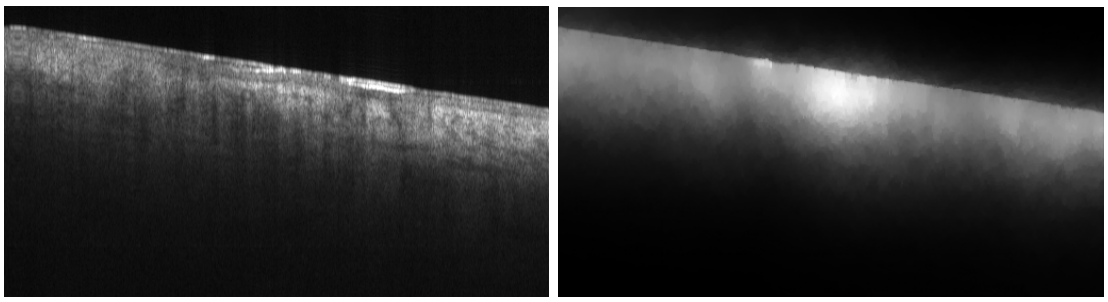


Figure 2.14: Scheme representing the separation of the 4 OCT volumes.

- **Skin alignment**: as visible in Figure 2.11, the skin is always sloped with respect to the en-face of the volume. Therefore, it has to be aligned before computing the OCTA, otherwise any analysis concerning depth will be negatively affected. Since the glass slide has been removed, the skin surface is the first, most prominent, and continuous intensity profile. For this purpose, the algorithm for skin detection developed by Li et al. is adapted into Python [15]. Firstly, a weighted least square (WLS) filter is applied as an edge-preserving smoothing method on each B-scan [16]. Skin detection is reduced to a problem of finding the shortest path between continuous points, while favouring lower-weight pixels, which present a strong gradient. A result of the WLS filter can be visualized in Figure 2.15.
  This algorithm is thus based on the premise that, again, skin is the first, most prominent and continuous intensity profile in the volume. Crucially, the WLS filter is used exclusively for

skin detection, while all the next passages, starting from the alignment, utilize the original OCT volumes.
Once the position of the surface is registered, the A-lines in each volume are shifted, so that the detected skin surface points are placed at the first layer of the 3D array. As shown in Figure 2.16, skin surface is followed almost pixel per pixel.



(a) **(b)**

Figure 2.15: Comparison of a set of 10 averaged B-scans before (a) and after (b) the application of the WLS filter.

- **Intensity-based OCTA**: the four aligned volumes are logarithmically scaled [10]. The OCTA volume is computed as the average of the three differences between sequentially acquired volumes (Eq. 2.1).

$$vol_{OCTA} = \frac{\Delta_{OCT_2-OCT_1} + \Delta_{OCT_3-OCT_2} + \Delta_{OCT_4-OCT_3}}{3} \quad (2.1)$$

- **Dermal-epidermal junction (DEJ) detection**: ideally, the DEJ should be found in every volume. Its location could be used to detect changes in epidermis thickness and as a zero-point for depth-related vascular calculations. This doesn't always happen, either due to suboptimal acquisitions, or the tumor breaking the DEJ or obscuring it [13] [17]. Therefore, its detection is possible only when the DEJ is clearly visible, or when it is not affected by pathological structural changes.
Nevertheless, the DEJ detection will still be explained. The implementation is again derived from the algorithm by Li et al.

[15], with some modifications to account for noise. The method is based on the DEJ presenting itself as a black stripe along the volume:

- one of the aligned OCT volumes is selected, then, the pipeline proceeds with the computation of the moving average of nine cross-scans, centered around one cross-scan after the other;
- the `detect_DEJ` function takes as input an averaged cross-scan and computes the gradient on the moving average of 5 A-lines centered around one line after the other. The gradient is low-pass filtered, so that peak detection is less sensitive to noise;
- peak detection is performed with the constraint of peaks being a minimum of 10 pixels apart. The DEJ is detected as the location of the only identified peak. If more are detected, the previous A-lines DEJ location is assumed.

Figure 2.16 and Figure 2.17 show the results of skin surface and DEJ detection.

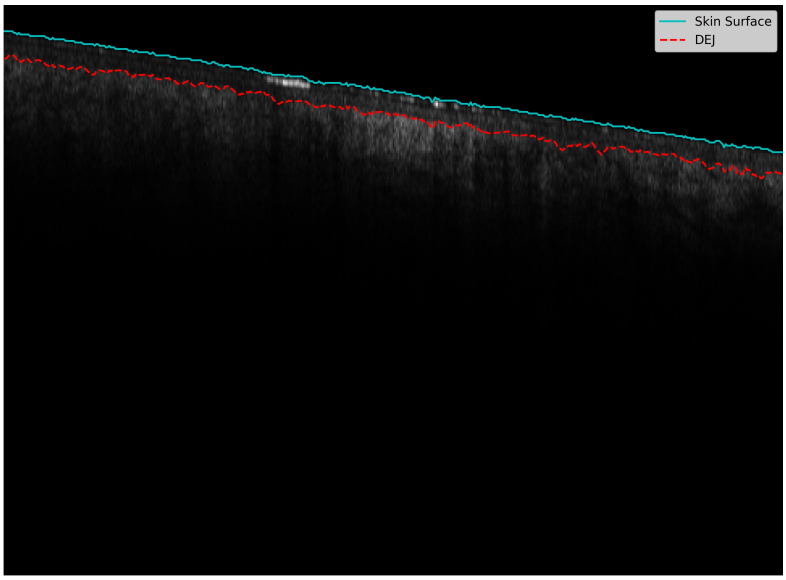


Figure 2.16: Skin and DEJ overlay on the moving average over a central B-scan.

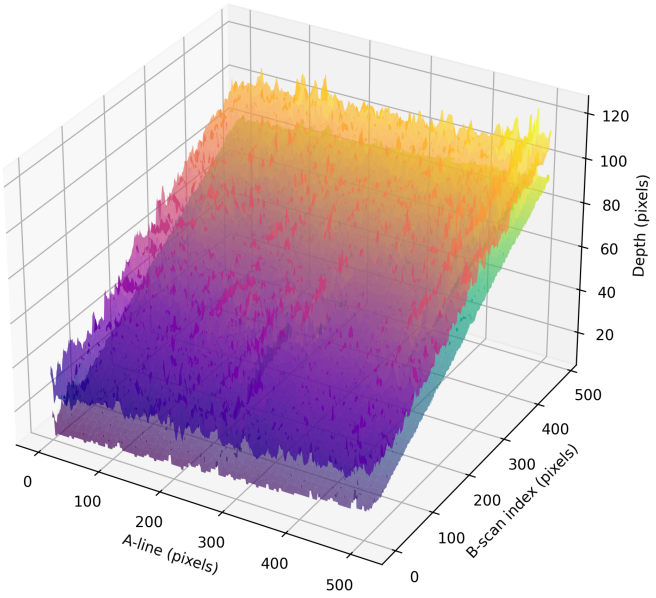


Figure 2.17: Representation in 3D of the detected skin surface and DEJ, that show how significant the slope is, and thus the importance of skin alignment.

The OCTA volume, which is the output of `intensity_based_OCTA`, is defined as `morph_angio`. This array is not ready yet for visualization and segmentation, but goes through the `artifacts_removal` function:

- **White line artifacts**: sudden shifts in intensity in B-scans, due to movement between subsequent acquisitions, appear as white vertical lines in the OCTA en-face. They are corrected by normalizing every B-scan by its mean [18].
- **Projection artifacts**: projections of superficial vessels appear on lower layers, as if they are casting a shadow [18]. They are corrected through a step-down exponential filtering method [19]:

$$P(i,j,k) = M(i,j,k) * \exp(-\frac{1}{\gamma} \sum_{r=0}^{i-1} * P(r,j,k)) \qquad (2.2)$$

  where P is the array after the correction of the artifact, M is the original volume, γ is the exponential decay constant, and (i,j,k) are depth and lateral coordinates.

- **Overall quality improvement**: before applying contrast enhancement, so that the vessel contrast more, a median filter is used to account for remaining noise.

In order to verify the success of the processing, the pipeline also outputs a series of en-face images, showed in Figures 2.18, 2.19, and 2.20.

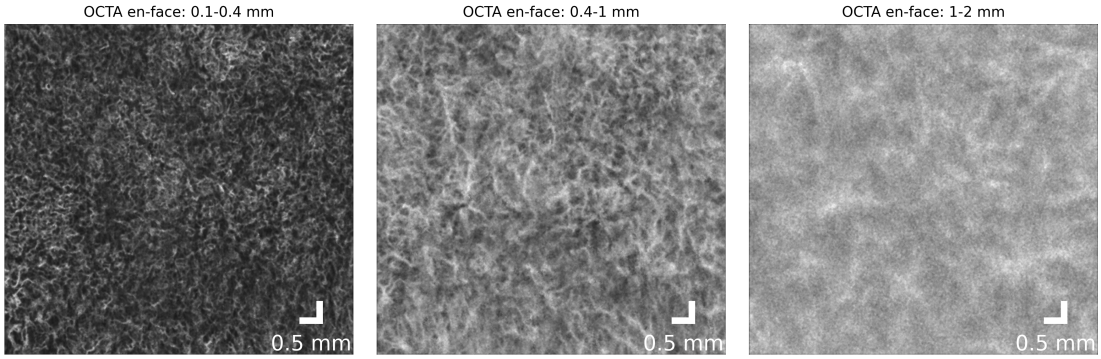


Figure 2.18: En-face representations of selected OCTA volume sections at different depths, obtained by computing the median intensity projection.

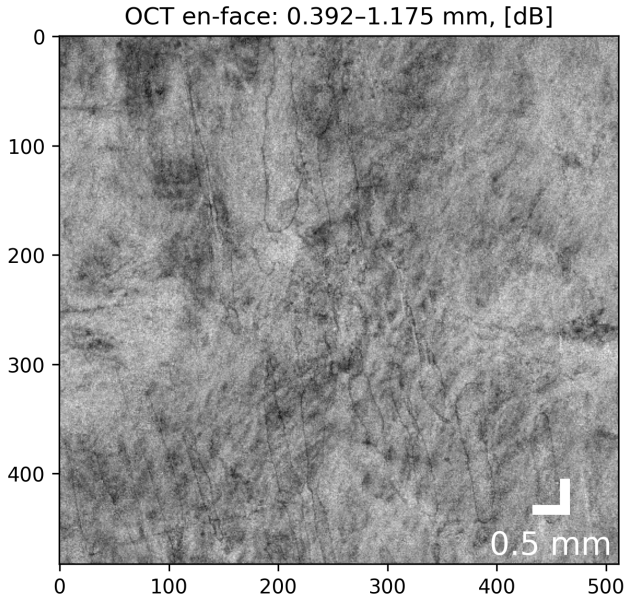


Figure 2.19: En-face representation of selected OCT volume section, obtained by computing a maximum intensity projection (MIP).

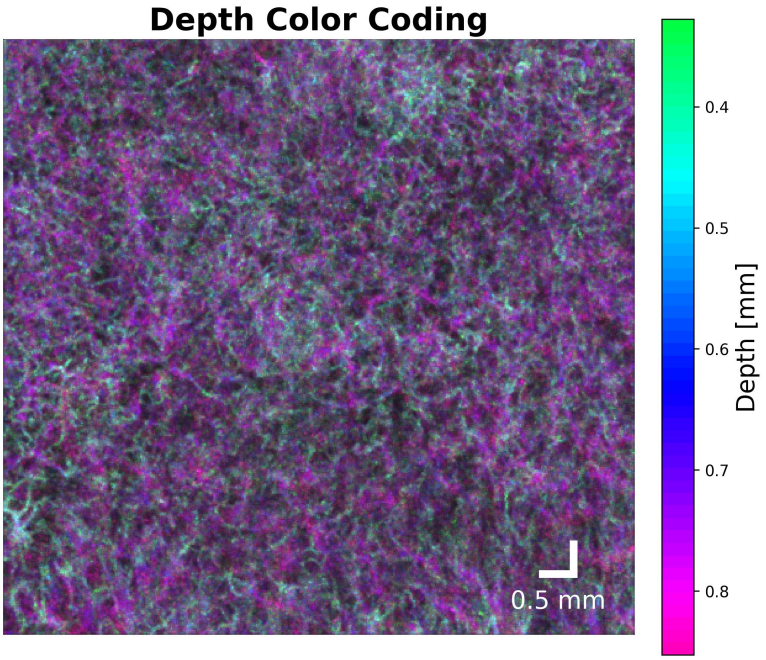


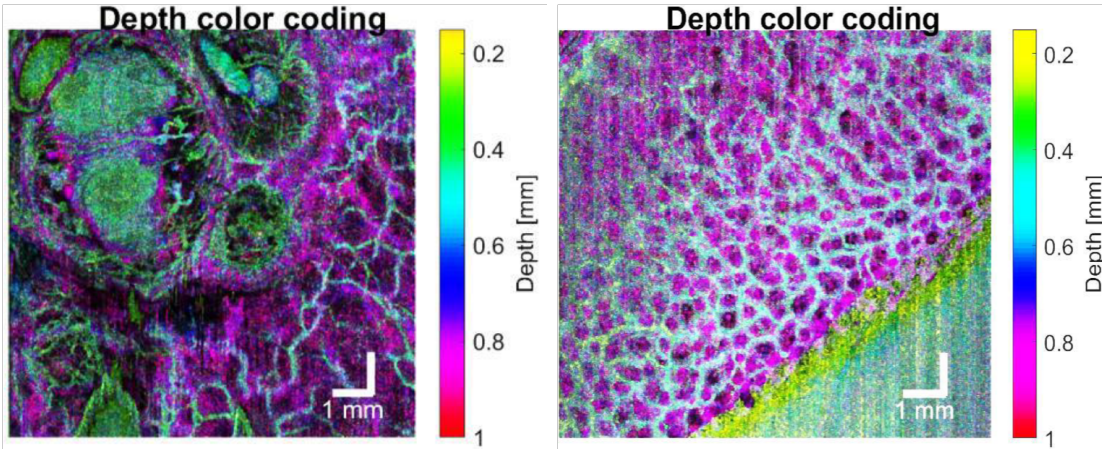
Figure 2.20: Depth color-coded MIP from a section of the OCTA volume.

These images are particuarly important, as they allow an initial and quick qualitative assessment of the lesion, without having to wait for segmentation and feature extraction.
The final array is thus ready for segmentation.

2.5.2 Segmentations and NaN masks

The outputs of the data processing pipeline are semi-automatically segmented with the AMIRA software. These segmentations are 3D binary masks, where 1 corresponds to vessels and 0 corresponds to the background.
Certain areas in OCT volumes don't contain any information. This happens especially with acquisitions in thin and heavily sloped areas such as the nose and the ear, which often leads to parts of the volume being unusable (2.21b). Other volumes contain blisters or small bubbles, that come out as angiographic signals in the OCTA volume, even if they're not blood vessels (2.21a).

Therefore, with AMIRA, it was decided to manually create dedicated binary masks, where the voxels that should be ignored are equal to 1, and the rest are equal to 0. Since these voxels are then symmetrically set as NaN in the segmentations, these masks are called NaN masks. All of them were checked alongside a clinician, in order to confirm that it was appropriate to judge certain areas as useless.



(a) **(b)**

Figure 2.21: Depth color-coded MIPs that show two different kinds of unwanted areas. The left one (a) shows bubbles and blisters, that appear as green. The other one (b), on the bottom right, shows an area with heavy artifacts due to the slope of the ear.

2.5.3 Feature extraction

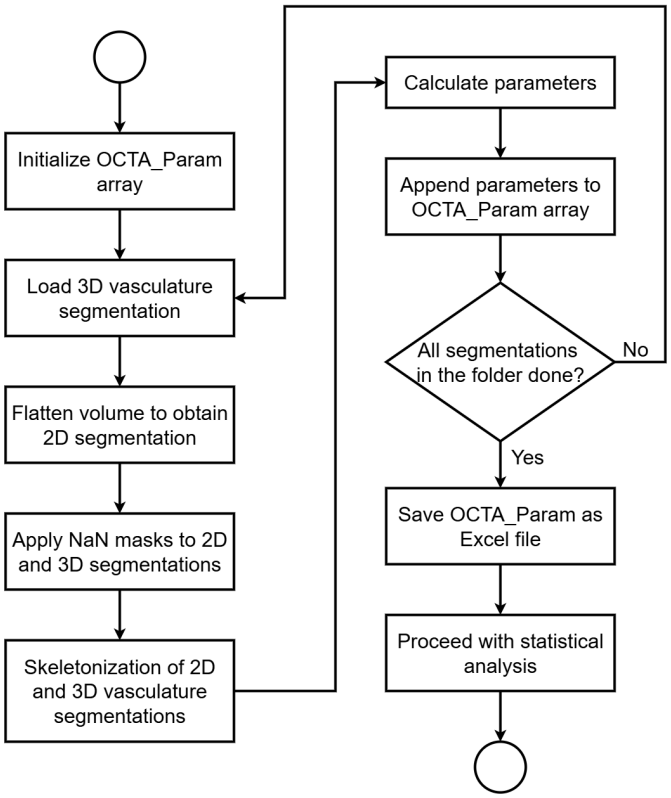


Figure 2.22: Pipeline of the feature extraction process, which gets repeated for every acquisition folder.

Feature extraction is performed on the segmentations, sometimes referred to as binary masks or just masks. The Python pipeline, summarized in Figure 2.22) is wholly automatic and uses a series of custom-made classes: `OCTAFeatures`, `globalParameters`, `ROIFeatures`, `applyNaNmasks` and `branchParameters`.
`OCTAFeatures` loads the segmentations and, by calling functions from the other classes, computes the values and saves them in a dedicated excel file, one for each timepoint, which then gets downloaded into the computer. It has three inputs:

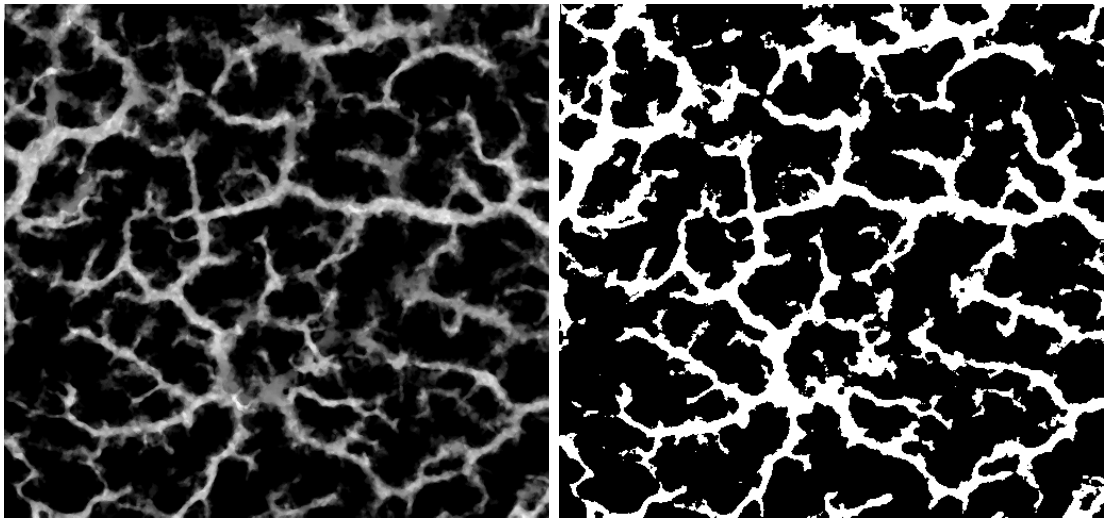
- **`patient_files_path`**: the path to a JSON file, which is divided into the timepoints and, for each one, contains the paths to the

binary segmentations and the NaN masks.

- **`variable_name`, `NaN_masks_name`**: which are, respectively, the file extensions of the segmentations and the Nan masks.

Features are extracted both from the 3D vasculature arrays and their 2D projections. These flattened arrays, that are en-face visualizations, give an additional insight into the vasculature architecture and allow the operator to have a quicker overview than the 3D arrays do. These are the necessary steps to obtain them:

- **Grayscale en-face**: sum of the vasculature mask along the depth and normalization by dividing it with the length of the depth axis, which results in a grayscale image (Figure 2.23a).



(a) **(b)**

Figure 2.23: Comparison between the en-face of a 3D vasculature segmentation (a) and its binarization with the Otsu method (b).

- **Linear thresholding**: binarization with threshold found with Otsu method (Figure 2.23b).

The `applyNaNmasks` class is then initialized and the NaN masks, if created for the currently processed volume, are applied to both the 3D and 2D binarizations:

- **3D segmentations**: here, where the NaN mask is equal to 1, the pixels are set as NaN.

- **2D segmentations**: the 3D NaN mask is flattened in the same way as the vasculature binary volume is. Then, where the 2D NaN mask is equal to 1, the 2D segmentation is set to NaN.

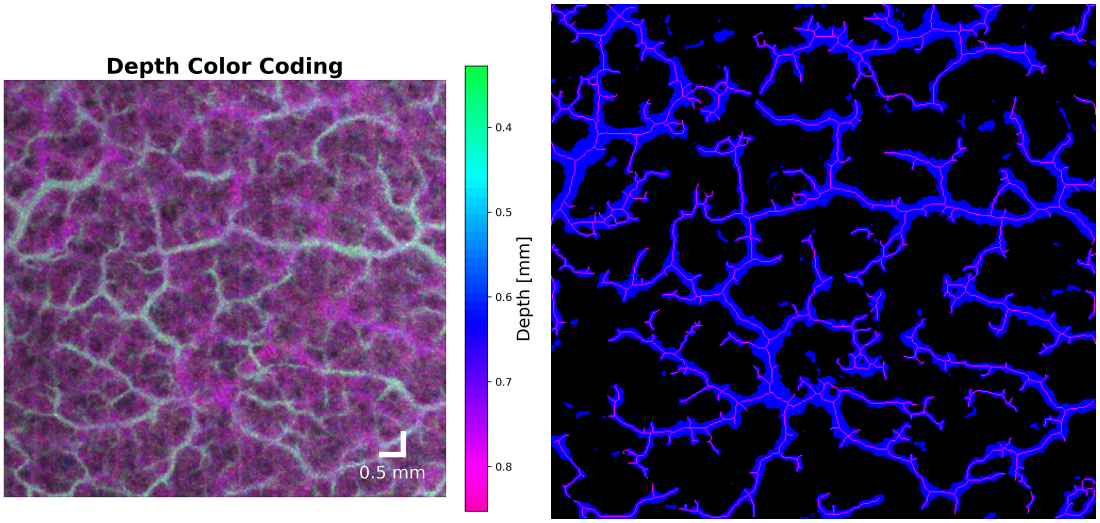
Setting the useless areas to NaN allows for the next calculations to ignore them. Before proceeding with feature extraction, both the segmentations follow some post-processing steps:

- **Morphological closing**: for hole filling, it removes small dark spots in favour of the blood vessels. As it's a dilation followed by an erosion, the dark holes that are closed are smaller than the structuring element, and those that are bigger retain their original size [20]. The structuring elements are a 2-by-2 square and a 2-by-2-by-2 cube, for the 2D and 3D segmentations respectively.

- **Skeletonization**: two additional arrays are created in the pipeline by reducing blood vessels to 1-pixel wide representations, centered around their median axis [21]. This allows to capture the essential characteristics of blood vessels.

- **Small objects removal**: blood vessels smaller than 21 pixels are removed in both skeletons.

An example of the output of this section is shown in Figure 2.24.



(a) **(b)**

Figure 2.24: On the right (b), overlap of a skeletonization (magenta) on its corresponding segmentation mask (blue), after all post-processing steps. On the left (a), the depth color-coded en-face of the OCTA volume the overlap was calculated from.

Then, feature extraction stars with the branch-related parameters, by calling a dedicated function from the `branchParameters` class. They are calculated exclusively for the 2D arrays, as calculating most of them for the 3D segmentations is too computationally heavy, due to the high number of branches. The subsequent features are calculated:

- **Number of endpoints**: they are at the extremities of the vascular trees. Endpoints have just 1 pixel in their 8-connected neighborhood.
- **Number of branchpoints**: points at which branches separate from each other. Branchpoints have 3 pixels or more in their 8-connected neighborhood. As they tend to aggregate, they're not counted individually, but one group is counted as a single branchpoint.
- **Number of branches and branches length**: the pixels contained between two endpoints, two branchpoints, or one endpoint

and one branchpoint are considered to be branches. The length is calculated by adding 2 to the number of these pixels. Branches without two extremities, such as loops, are eliminated.

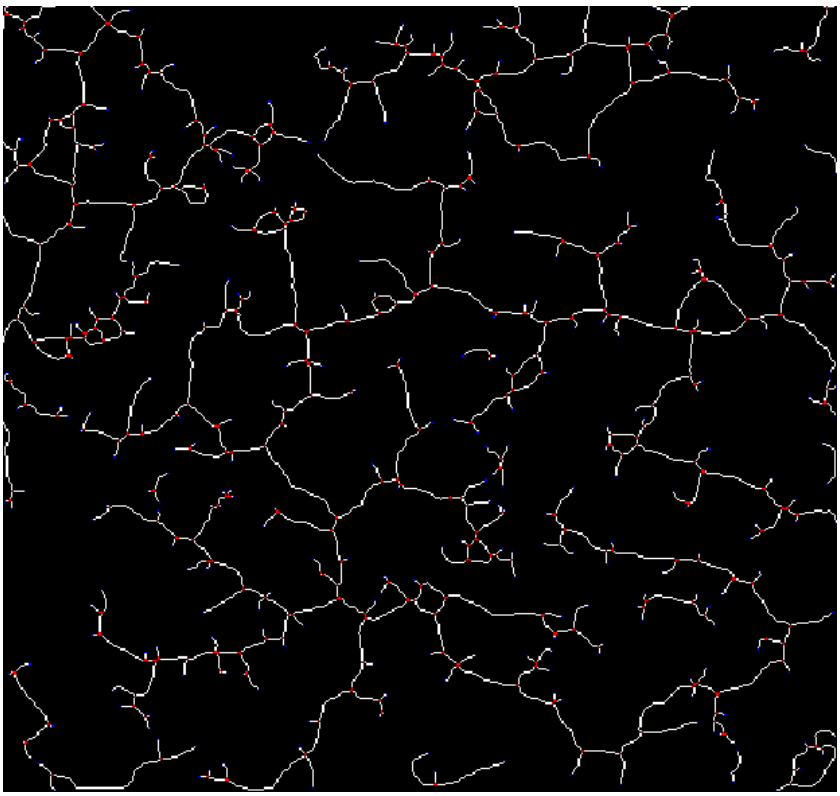


Figure 2.25: Picture showing branchpoints (red) and endpoints (blue) highlighted on a skeleton.

- **Tortuosity measures**: for each branch, along with its extremities, three parameters are calculated [22]. These are:
  - **Distance metric (DM)**: ratio between the number of pixels in the branch and the euclidean distance between its extremities (Eq. 2.3). The higher the value, the more tortuous the branch is.

$$DM = \frac{branch\ length}{euclidean\ distance\ between\ extremities} \tag{2.3}$$

  - **Inflection count metric (ICM)**: the DM of the branch

multiplied by the number of inflections along its path, which is calculated as the number of times the vessel diverges from its lowest-cost path (Eq. 2.4). Again, a high value corresponds to a highly tortuous vessel.

$$ICM = DM * n\_inflections \tag{2.4}$$

– **Sum of angles metric (SOAM)**: a sum of the angles in the vessel normalized by the branch length (Eq. 2.5). Two types of angles are calculated: in-plane angles (IP), between the tangent vectors of two adjacent pixels, and torsional angles (TP), between the normal vectors of two subsequent voxels, which are null in 2D arrays.

$$SOAM = \frac{1}{branch\ length} \sum_{j=1}^{n-2} \sqrt{IP_j^2 + TP_j^2} \tag{2.5}$$

The subsequent features are then calculated, for both 2D and 3D segmentations and their skeletonizations, unless specified otherwise, by calling `globalParameters`:

- **Radius**: distance between the skeleton and the border of the corresponding segmentation. It's calculated by computing the distance transform within the vessel mask to the closest skeleton pixel. Then, the distance values corresponding to the borders of the vessels are taken as radii values.
- **Vascular density (VD)**: ratio of the white pixels to the whole size of the array, minus the number of NaN pixels.
- **Avascular area (AA) and avascular volume (AV)**: defined as portions of tissue that are further from the nearest blood vessel than a predetermined value (Figure 2.26). This value is assumed to be 150 μm, the maximum distance that necessary compounds can travel before being metabolized by cells. Anything beyond that can be assumed to be an avascular volume, or area, as it's not receiving enough oxygen due to the lack of close enough vessels

[23]. Given the lateral and axial pixel spacings, the threshold distances in pixels are:

$$\begin{aligned} max\_lat\_distance &= \frac{max\_distance}{lat\_pixel\_spacing} \approx 8\ pixels \\ max\_ax\_distance &= \frac{max\_distance}{ax\_pixel\_spacing} \approx 11\ pixels \end{aligned} \tag{2.6}$$

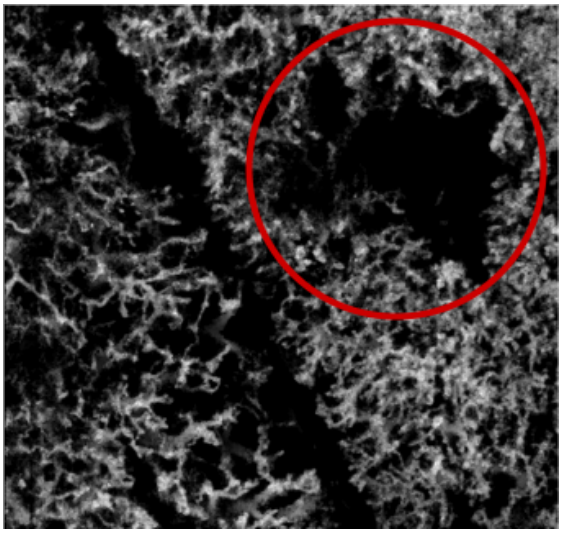


Figure 2.26: Grayscale en-face of the 3D segmentation. Inside the red circle there's one of the avascular areas targeted by the pipeline.

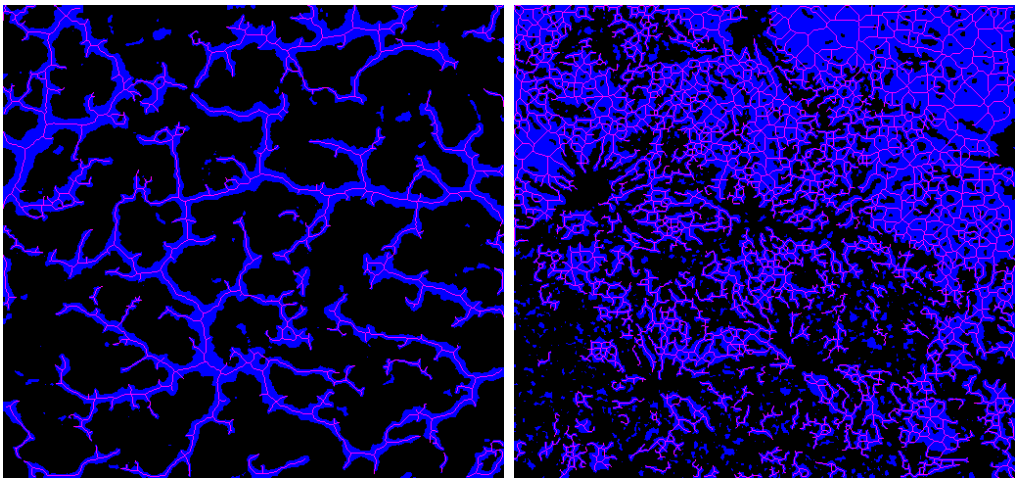
To keep the code simpler and due to the two values differing by just 3 pixels, the maximum distance is always assumed to be equal to the lateral one, even in 3D.
A distance transform is computed on the complementary arrays of the two segmentations. An area, in 2D, or a volume, in 3D, of the array is considered to be avascular if it's more than 8 pixels distant from the nearest vessel. The number of avascular pixels, or voxels, is thus calculated, and the percentage with the number of total valid pixels is calculated as well.

- **Shannon entropy**: it quantifies the amount of uncertainty, or information, associated with the description of the variable's

states. In information theory, it's calculated as such:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \tag{2.7}$$

where X is a discrete random variable and $p(x_i)$ is the probability of the variable assuming value x_i.
In our context, X represents the possible pixel values, 0 and 1, and $p(x_i)$ is the probability distribution for one of these, calculated as the total sum of that value divided by the total number of valid pixels. The higher the entropy value of an array, the more it can be described as chaotic, with the minimum being 0 and the maximum being 1 (Figure 2.27).



(a) **(b)**

Figure 2.27: Comparison between two acquisitions with different entropy values. The one on the right (b) is more chaotic, and the calculation reflects it: the segmentation (blue) and its skeleton (magenta) have entropy values of respectively 0,998 and 0,458. The left one, on the other hand, has entropy values of 0,819 and 0,205, respectively for mask and skeletonization.

- **Fractal dimension (FD)**: it's an index used to describe the complexity in a pattern and how detail changes with the scale at which it is measured. The higher its value, the more complex and irregular a pattern is. The function used in `globalParameters`

class uses the box-counting method, which measures the change in detail with increasingly smaller boxes [24].

- **Number of trees (NT)**: the number of connected objects in a vascular binary array is considered to be the number of vascular trees.
- **Distance-based metrics**: they are calculated only in 3D. There are two of these:
  - the axial position of the center of mass of the segmentation and its skeleton;
  - the axial distance of a chosen voxel of the segmentation from the top layer. The voxel is that with the highest distance from the background.

It's important to understand how some of these features vary within the same array. The `ROIFeatures` class divides each of the four arrays (2D and 3D segmentations and their skeletonizations) into fixed-dimension ROIs (Table 2.4) and calculates for each of them VD and Entropy. The same is done for AA and AV, while of course just taking the two segmentations into account, as these two parameters are not computed for the skeletonizations. By computing the max value between all ROIs, it's possible to assess these features where they're most intense.

Dimension	ROI size (pixels)	Step (pixels)
2D	64 x 60	32 x 30
3D	10 x 64 x 60	5 x 32 x 30

Table 2.4: Characteristics of the ROIs into which the arrays are divided.

Moreover, ROI values are used to create heatmaps of these features, that allow for an assessment of how values are distributed in these arrays. They are created for each acquisition and some examples are shown in Figure 2.28.

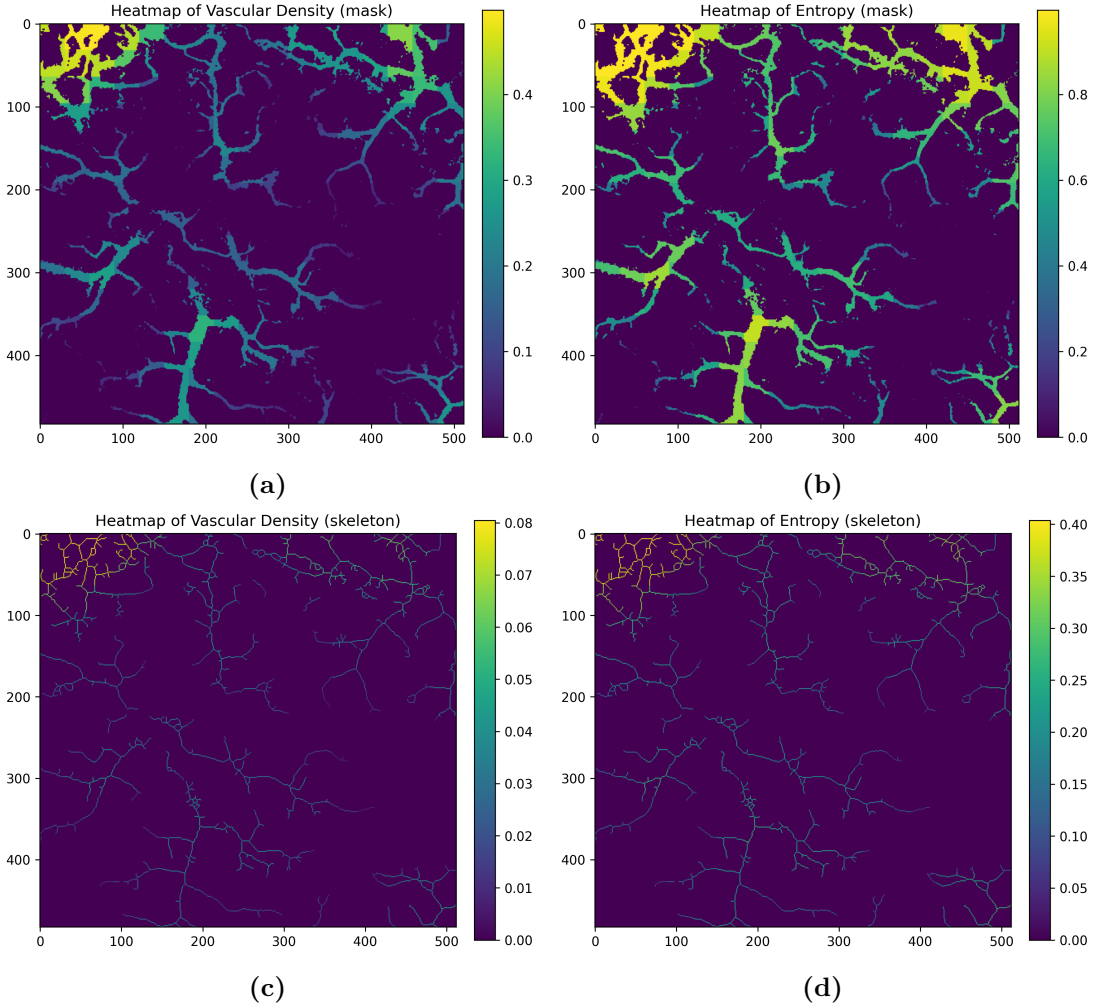


Figure 2.28: Heatmaps of vascular density and entropy across the 2D segmentation and the 2D skeletonization of the same acquisition.

The `ROIFeatures` class also allows to calculate the variation of VD and entropy along with depth. The 3D mask and its skeleton are divided into slices that are 10 pixels thick, with a 5 pixel step. Then, in order to characterize the distributions of these values, their skewness is calculated. However, as the segmentations may have different depths, a cutting point of 13 slices, so 70 pixels, is applied to the distributions before skewness computations.
A total of 53 features are calculated for each acquisition. For a summary, see Figure 2.29.

2D vascularization segmentation	2D vascularization skeletonization
• Vascular density • Avascular area • % of avascular area • Entropy • Fractal dimension • Max ROI vascular density • Max ROI avascular area • Max ROI % of avascular area • Max ROI entropy • Mean radius (with skeleton)	• Number of endpoints • Number of branchpoints • Number of branches • Mean and max branch length • Mean DM • Mean ICM • Mean SOAM • Branch parameters • Vascular density • Entropy • Fractal dimension • Max ROI vascular density • Max ROI entropy • Number of trees
3D vascularization segmentation	**3D vascularization skeletonization**
• Vascular density • Avascular volume • % of avascular volume • Entropy • Fractal dimension • Distance-based metrics • Max ROI vascular density • Max ROI avascular volume • Max ROI % of avascular volume • Max ROI entropy • Skewness of vascular density depth distribution • Skewness of entropy depth distribution • Mean radius (with skeleton)	• Vascular density • Entropy • Fractal dimension • Distance-based metrics • Max ROI vascular density • Max ROI entropy • Skewness of vascular density depth distribution • Skewness of entropy depth distribution • Number of trees

Figure 2.29: All calculated features are listed here and divided by the arrays from which they're calculated. Mean radii are calculated both from the masks and their skeletonizations: they're under the former for convenience.

Overall, the outcome of the feature extraction process are 5 Excel files, 4 relative to the timepoints and one for the healthy volumes. These files are structured so that each column corresponds to a variable, and each row corresponds to a different volume.

2.5.4 Statistical analysis

The purpose of the statistical analysis, shown in 2.30, is to determine whether features change significantly across timepoints. Additionally, boxplot diagrams help contextualize these results, hopefully showing, with time, a return of values to the healthy baseline. Thus, this

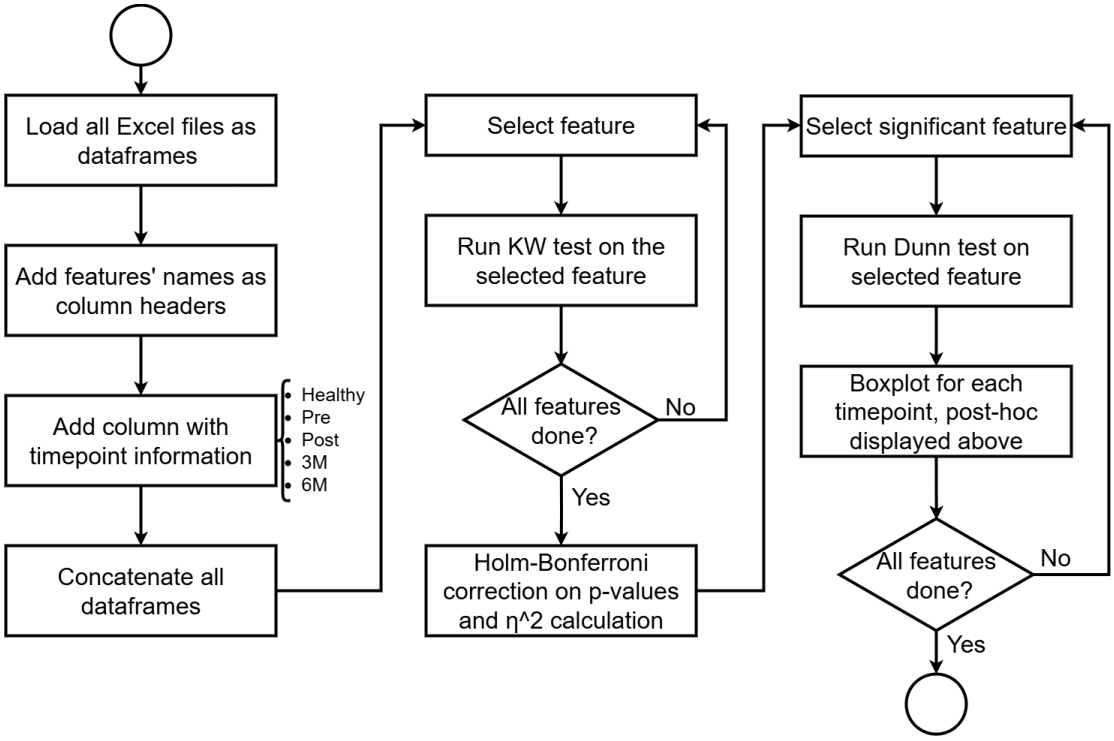


Figure 2.30: Pipeline of the statistical analysis.

analysis is fundamental in understanding which features can characterize vascular changes after radiotherapy treatment and differentiate between the timepoints. The analysis proceeds as such:

- **Kruskal-Wallis (KW)**: it's a nonparametric test used to assess whether two or more independently sampled groups come from the same distribution. The single variable of interest doesn't need to meet normality assumptions, but the groups should come from populations that share the same distribution shape.
  To test the null hypothesis, which states that the medians of all groups are equal, all samples are combined and their values are ranked. The ranks are summed separately for each group and, along with the sample sizes, they are used to calculate the H statistic. It reflects the variance in ranks between groups, so the higher its value the more at least one group differs from the others [25].

The `kruskal` function used in the pipeline returns both the H statistic and the p-value, using the assumption that H has a chi square distribution [26]. All p-values are tested against α=0.05. If they are lower than α, the null hypotesis is rejected, which signals significant differences between the timepoints.

- **Holm-Bonferroni (HB) correction**: the Bonferroni method is used to reduce the optimism of the KW test, as the previously calculated p-values are multiplied by the number of total tests performed.
  However, the Bonferroni correction by itself is considered to be too conservative, so Holm's sequential version is used here [27]. All tests are ranked by their raw p-value, from lowest to highest. Then, each corrected p-value is calculated as such:

$$p_{i,\ corrected} = p_i * (N_t - i + 1) \tag{2.8}$$

  where N_t is the number of tests, or features, and p_i is the i-th ranked p-value. Starting from p_1, its correction is tested against α. All other values are then checked in ascending rank order, and at the first $p_{i,\ corrected}$ higher than 0.05, all larger raw p-values are automatically considered non-significant.

- **Eta squared (η^2)**: given the total variance in all values assumed by a feature, it measures how much of that is associated with those values being calculated in specific timepoints. Here, the η^2 is obtained as such [28]:

$$\eta_i^2 = \frac{H_i - n_{timepoints} + 1}{n_{values,\ i} - n_{timepoints}} \tag{2.9}$$

  where H_i is the H statistic for that feature and $n_{values,\ i}$ is the total number of values of that feature across all timepoints.
  The higher the η^2 value, the larger the effect of timepoints on the value that a feature can assume.

- **Dunn test**: even if the KW test leads to significant results, it cannot tell which groups are different from the others. Therefore,

the Dunn test is essential as a post-hoc analysis to make pair-wise comparisons.

η^2 **(effect size - equal or higher than)**		
Small	**Medium**	**Large**
0.0099	0.0588	0.1379
p-value (significance - lower than)		
Verified	**Strong**	**Very strong**
0.05	0.001	0.0001

Table 2.5: Table with noteworthy values regarding η^2 and the p-value [28].

Overall, an ideal feature is one that presents a large H-value and a corrected p-value lower than 0.05. This means that the feature changes significantly across timepoints. An η^2 higher than 0.1379 enforces the observation that the specific timepoint at which the parameter is calculated largely affects its value. The Dunn test, as a post-hoc analysis, shows for this feature which timepoints differ significantly from each other. Lastly, boxplots diagram allow for a visual assessment of the distribution of its values across timepoints, showing a gradual return towards the healthy baseline.

Chapter 3

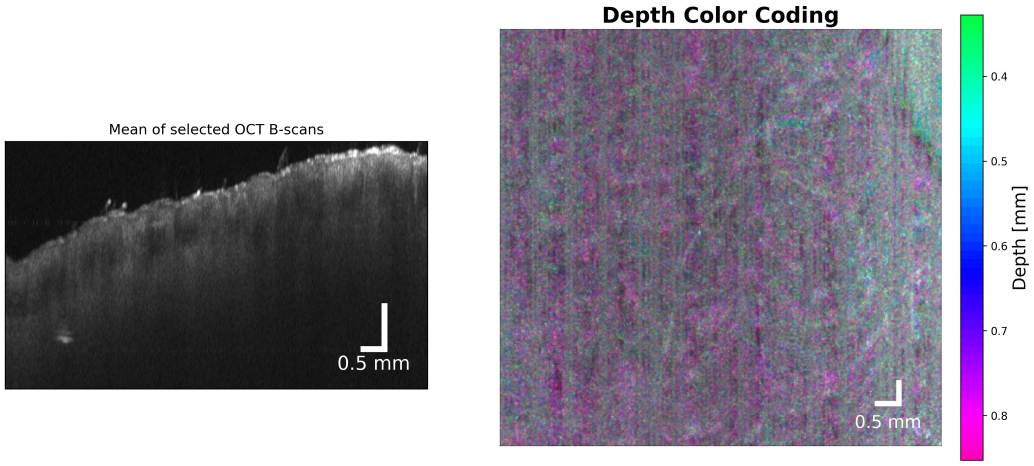
Results

3.1 Qualitative analysis

The images produced by the processing pipeline are important to make a first assessment of the lesion and to contextualize the results of the statistical analysis.

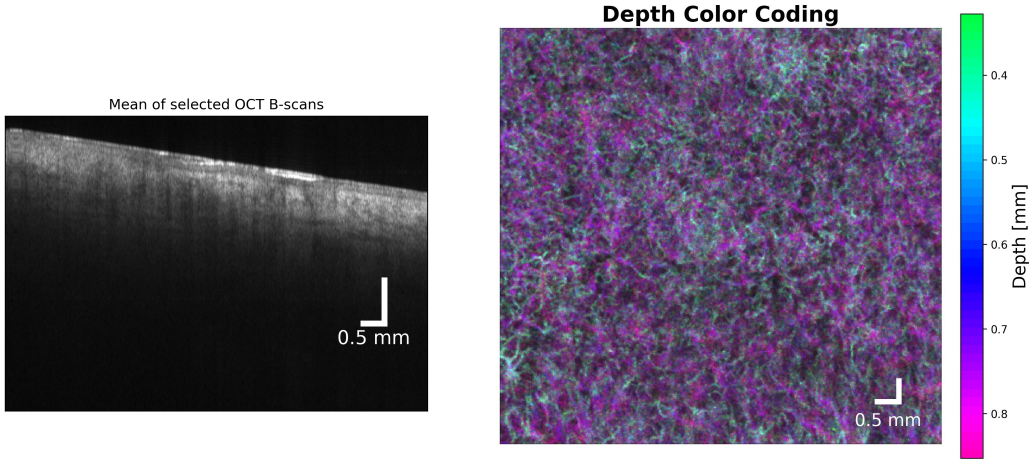
3.1.1 Determining the quality of the acquisition

As it was said in Section 2.1, the overall amount of acquired volumes is much larger than the final selection. Some volumes are, in fact, low-quality, which can be quickly determined by looking at the B-scans, en-face and depth color-coded plots.
Figure 3.1 shows an acquisition taken from the nose. The insufficient quality can be traced back to the lack of the glass slide. This leads to more unwanted reflections, which compromise the overall quality, and to a highly irregular surface, which affects skin alignment and hinders the correct focusing of the laser. Therefore, vessels that are out of focus appear to be blurry. Moreover, there are heavy movement artifacts in the OCTA en-face. While these could be mainly caused by the patient, they could have been partially mitigated by keeping the skin firmer with the slide.



(a) **(b)**

Figure 3.1: Images extracted from an acquisition taken at the nose: average of 10 B-scans, which shows a highly irregular skin surface (a); depth color-coded MIP from a section of the OCTA volume, where heavy movement artifacts can be seen (b).



(a) **(b)**

Figure 3.2: Images extracted from an acquisition taken at the leg: average of 10 B-scans, which shows a sloped planar surface (a); depth color-coded MIP from a section of the OCTA volume (b).

Figure 3.2 shows images with higher quality. The skin here is planar, and allows for the use of the glass slide. The flattened surface

leads to a higher-quality OCTA processing, with no visible movement artifacts and vessels being focused on more precisely.
Of course, some acquisitions are bound to have slightly suboptimal quality, but the ones like that shown in Figure 3.1 are never considered for segmentation. With the fast raw data processing pipeline implemented in the main data acquisition workstation, it could be possible to take a quick look at the image outputs to determine whether an additional acquisition is necessary.

3.1.2 Visual timepoint comparison

The premise that vasculature features may change across timepoints, and the choice of the feature themselves, is based on qualitative observations. Visual monitoring is still crucial, both as a first assessment, and as a way to contextualize features' values.
Figure 3.4 shows a comparison of the same lesion across all timepoints. A healthy portion of skin, from the same patient, can be used as reference (Figure 3.3).

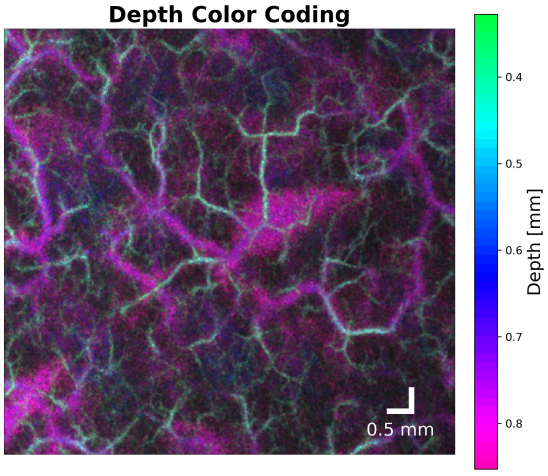


Figure 3.3: Depth color-coded MIP from a section of an OCTA volume of healthy skin.

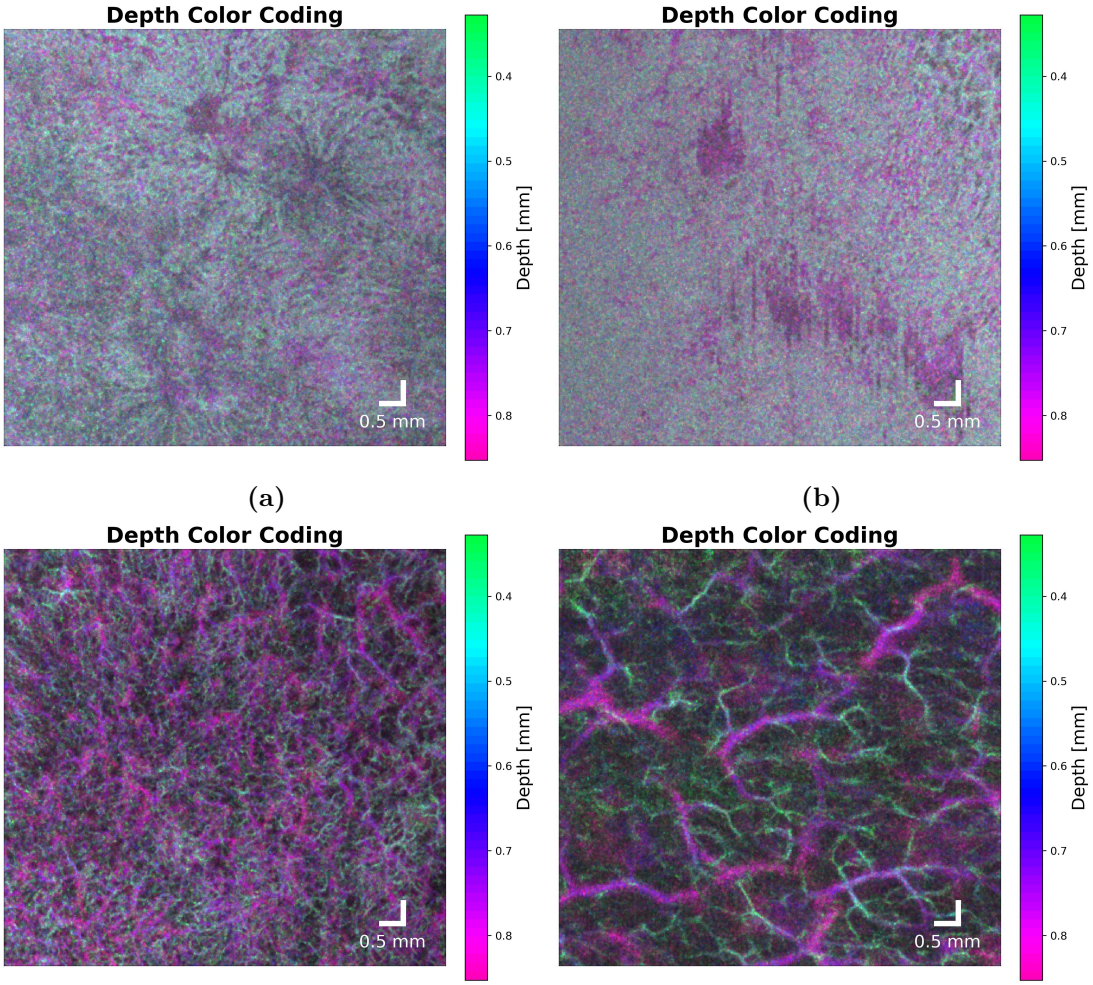


Figure 3.4: Four figures showing the progression of a lesion throughout the timepoints: pre (a), post (b), 3M (c), and 6M (d). The first two seem blurry, due to the high density of the arrays.

As it can be seen, the pre and post radiotherapy lesions have very peculiar layouts, so packed and dense that it's very difficult to see any vessels in the deeper layers. The healthy skin, on the other hand, presents a much more organized vasculature, which is visibly far less dense. Moreover, deeper vessels are clearly visible. The 3M and 6M vasculatures signal a return to structures that are progressively more similar to the healthy ones. Vessel layouts are rare and more

distinguishable than in the previous timepoints.

3.2 Quantitative analysis

Metric	KW H-stat	η^2	Effect size	Raw p-values	Corrected p-values	Reject Hypotesis
Number of Trees (3D)	29,58	0,1881	large	5,97E-06	1,85E-04	True
Global AA (2D)	42,62	0,2840	large	1,24E-08	5,58E-07	True
Global AA % (2D)	38,19	0,2514	large	1,02E-07	4,20E-06	True
Global AV (3D)	19,11	0,1111	medium	7,47E-04	1,49E-02	True
Global AV % (3D)	21,98	0,1322	medium	2,03E-04	4,86E-03	True
Number of Endpoints (2D)	21,18	0,1263	medium	2,91E-04	6,70E-03	True
Number of Branchpoints (2D)	39,30	0,2595	large	6,05E-08	2,60E-06	True
Number of Branches (2D)	37,59	0,2470	large	1,36E-07	5,32E-06	True
Mean Radius (2D)	19,10	0,1111	medium	7,50E-04	1,49E-02	True
Mean Radius (3D)	37,99	0,2499	large	1,13E-07	4,51E-06	True
Mean ICM (tortuosity, 2D)	42,44	0,2826	large	1,35E-08	5,95E-07	True
Mean SOAM (tortuosity, 2D)	42,82	0,2855	large	1,13E-08	5,18E-07	True
Max ROI VD (m, 2D)	46,99	0,3161	large	1,53E-09	7,67E-08	True
Max ROI Entropy (m, 2D)	34,57	0,2248	large	5,69E-07	2,10E-05	True
Max ROI VD (s, 2D)	51,80	0,3515	large	1,52E-10	7,90E-09	True
Max ROI Entropy (s, 2D)	51,80	0,3515	large	1,52E-10	7,90E-09	True
Max ROI VD (m, 3D)	29,02	0,1840	large	7,74E-06	2,01E-04	True
Max ROI Entropy (m, 3D)	34,71	0,2258	large	5,33E-07	2,02E-05	True
Max ROI VD (s, 3D)	29,39	0,1867	large	6,51E-06	1,95E-04	True
Max ROI Entropy (s, 3D)	29,39	0,1867	large	6,51E-06	1,95E-04	True
Global VD (m, 2D)	44,78	0,2999	large	4,41E-09	2,16E-07	True
Global VD (s, 2D)	43,52	0,2906	large	8,08E-09	3,88E-07	True
Global Entropy (m, 2D)	38,29	0,2521	large	9,78E-08	4,11E-06	True
Global Entropy (s, 2D)	43,52	0,2906	large	8,08E-09	3,88E-07	True
Global VD (m, 3D)	29,13	0,1848	large	7,37E-06	1,99E-04	True
Global VD (s, 3D)	33,82	0,2192	large	8,13E-07	2,93E-05	True
Global FD (m, 3D)	31,20	0,2000	large	2,78E-06	8,90E-05	True
Global FD (s, 3D)	32,30	0,2081	large	1,66E-06	5,66E-05	True
Global Entropy (m, 3D)	29,24	0,1856	large	6,99E-06	1,96E-04	True
Global Entropy (s, 3D)	33,82	0,2192	large	8,13E-07	2,93E-05	True
Mean Branch Length (2D)	31,74	0,2040	large	2,16E-06	7,13E-05	True
Distance (m, 3D)	17,58	0,0998	medium	1,49E-03	2,69E-02	True
Skewness VD - cut (m, 3D)	17,36	0,0983	medium	1,64E-03	2,79E-02	True
Skewness Entropy - cut (m, 3D)	20,06	0,1181	medium	4,86E-04	1,07E-02	True
Global FD (m, 2D)	19,89	0,1168	medium	5,26E-04	1,10E-02	True
Global FD (s, 2D)	24,34	0,1496	large	6,82E-05	1,70E-03	True

Figure 3.5: Here all the significant features are displayed. VD: vascular density; AA: avascular area; AV: avascular volume; ICM: inflection count metric; SOAM: sum of angles metric; FD: fractal dimension; m: mask; s: skeleton.

The null hypothesis is rejected for 36 of the features, which thus have statistically significant differences across timepoints. Out of these, 6 present a strong statistical significance and 21 a very strong one. Additionally, only 8 features present a medium size effect, with the rest having a large one; this means that the variance seen in features is largely due to the timepoint in which they are calculated. These results are shown extensively in Figure 3.5.
The rejected features are displayed in Figure 3.6.

Metric	KW H-stat	η^2	Effect size	Raw p-values	Corrected p-values	Reject Hypotesis
Mean DM (tortuosity, 2D)	5,39	0,0102	small	2,50E-01	1,00E+00	False
Max ROI AA (2D)	12,65	0,0636	medium	1,31E-02	1,70E-01	False
Max ROI AA % (2D)	13,91	0,0728	medium	7,60E-03	1,06E-01	False
Max ROI AV (3D)	6,03	0,0149	small	1,97E-01	1,00E+00	False
Max ROI AV % (3D)	8,90	0,0360	small	6,38E-02	5,74E-01	False
Max Branch Length (2D)	11,94	0,0584	small	1,78E-02	2,14E-01	False
Center of Mass (m, 3D)	11,10	0,0522	small	2,55E-02	2,80E-01	False
Center of Mass (s, 3D)	10,29	0,0463	small	3,58E-02	3,58E-01	False
Distance (s, 3D)	6,26	0,0166	small	1,81E-01	1,00E+00	False
Skewness VD - cut (s, 3D)	14,62	0,0781	medium	5,55E-03	8,32E-02	False
Skewness Entropy - cut (s, 3D)	15,41	0,0839	medium	3,92E-03	6,27E-02	False
Skewness VD (m, 3D)	5,05	0,0077	negligible	2,82E-01	1,00E+00	False
Skewness VD (s, 3D)	3,60	-0,0029	negligible	4,63E-01	1,00E+00	False
Skewness Entropy (m, 3D)	6,48	0,0182	small	1,66E-01	1,00E+00	False
Skewness Entropy (s, 3D)	3,79	-0,0015	negligible	4,34E-01	1,00E+00	False
Number of Trees (2D)	8,25	0,0313	small	8,27E-02	6,62E-01	False

Figure 3.6: Here all the non-significant features are displayed. DM: distance metric; VD: vascular density; m: mask; s: skeleton.

The results of the post-hoc analysis are displayed above the boxplots as asterisks, where one indicates the presence of a statistically significant difference, two indicate a strong one, and three indicate a very strong one. This is a quick index of the boxplot diagrams below:

- **Branch parameters**: Figure 3.7 and Figure 3.8. These features are calculated exclusively on branches, thanks to the skeletonization.

- **Global parameters**: Figure 3.9, Figure 3.10, Figure 3.11, Figure 3.12, Figure 3.13 and Figure 3.14.

- **ROI parameters**: Figure 3.15 and Figure 3.16. These features are calculated on the ROIs in which an array is divided. Then, the max of each of the ROI's values is computed.

- **Distance parameter**: Figure 3.17, the distance of the innermost point of the 3D segmentation from the surface.

- **Depth-related parameters**: Figure 3.18, where each feature is the skewness of the distribution of a specific parameter along with depth.

When looking at these boxplots, one has to pay attention to a set of key characteristics:

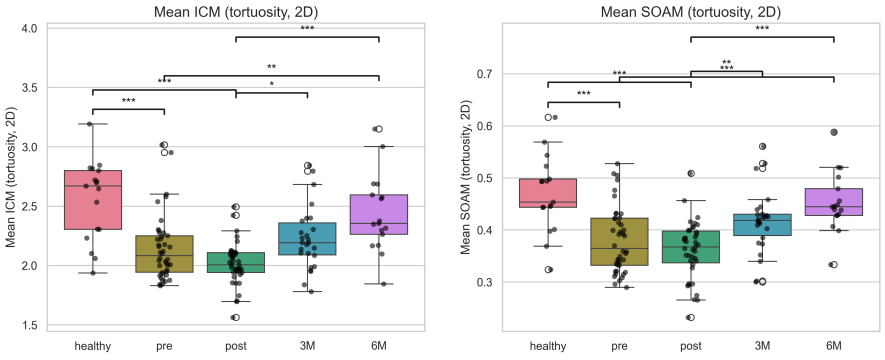
- **Relative positions**: if all the boxplots are aligned, then the feature is not able to discriminate timepoints.

- **Trend**: not only do they need to be shifted in height, but there needs to be a trend that shows features' values initially distant from the healthy baseline and then slowly converging back towards it.

- **Interquartile range (IQR)**: sometimes, the pre- and post-radiotherapy groups have a larger or smaller IQR than the healthy baseline.

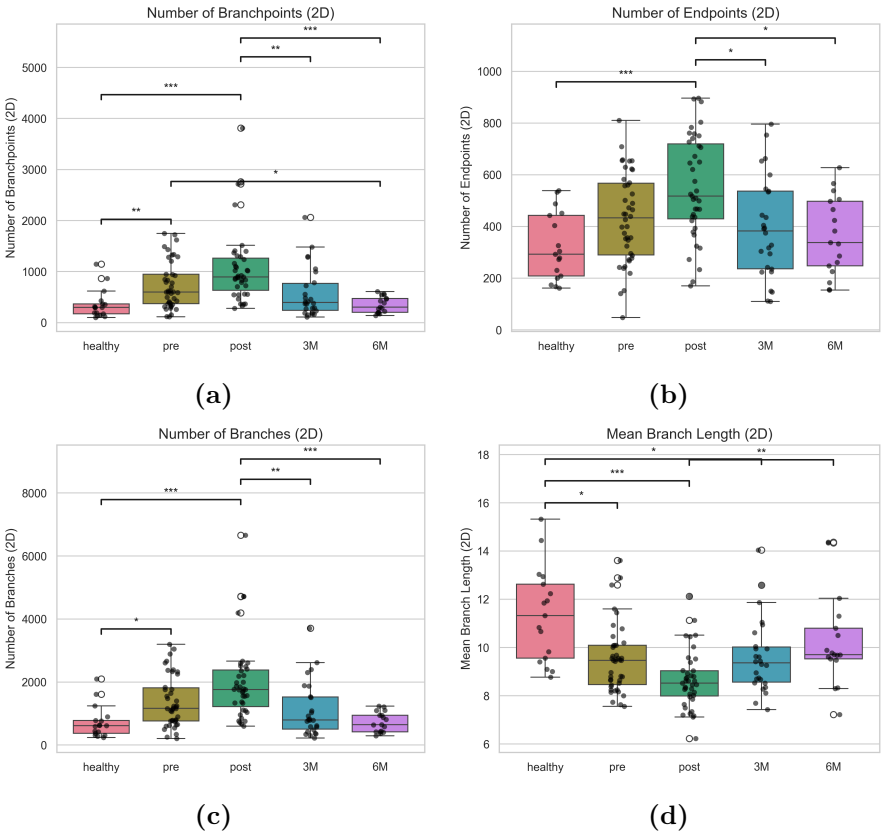
- **Dunn's test results**: crucial to assess the entity of the difference between timepoints. Sometimes, such as in Figures 3.16a and 3.16c, due to heavy outliers, the boxplots are barely visible. In these cases, the post-hoc results are fundamental for the interpretation of results.

All figures show that the first two timepoints drift consistently far from the healthy baseline, with the last two slowly shifting back towards it.



(a) **(b)**

Figure 3.7: Boxplots of two tortuosity measures across timepoints, with Dunn's test results above.



(a) **(b)**

(c) **(d)**

Figure 3.8: Boxplots of branch-related parameters across timepoints, with Dunn's test results above.

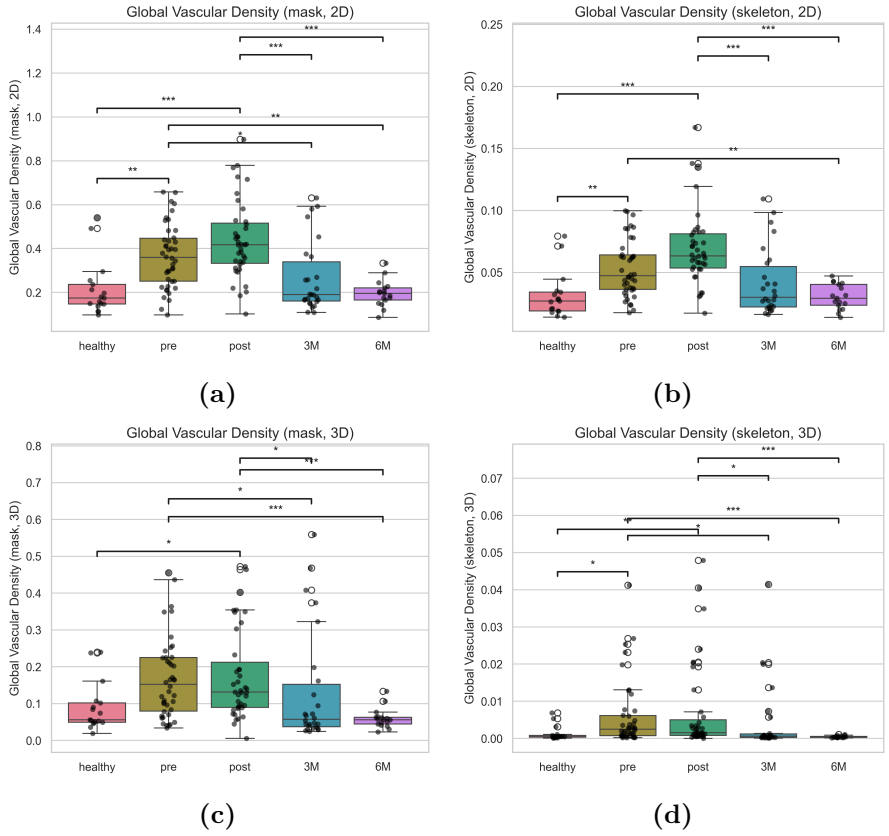


Figure 3.9: Boxplots across timepoints of the vascular density of 4 different arrays, with Dunn's test results above.

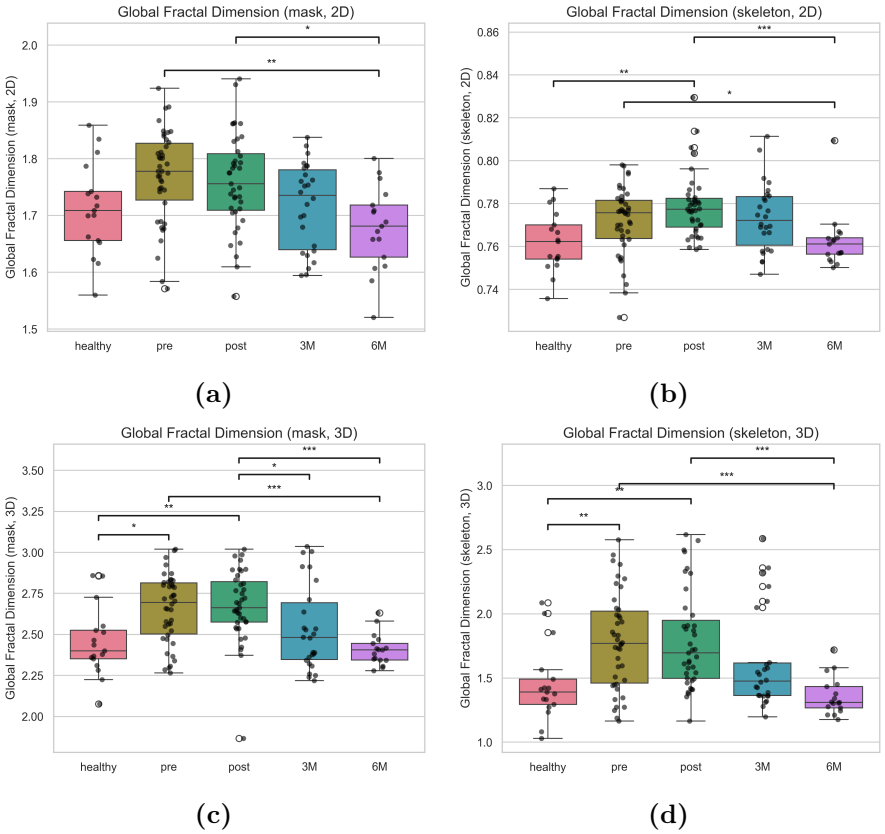


Figure 3.10: Boxplots across timepoints of the fractal dimension of 4 different arrays, with Dunn's test results above.

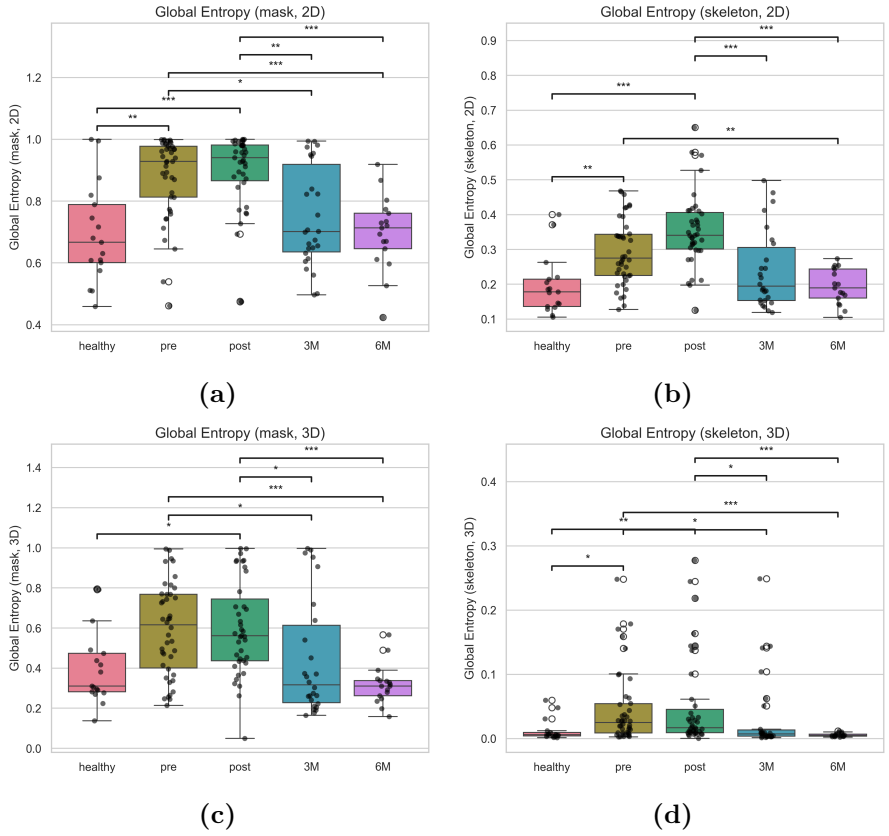


Figure 3.11: Boxplots across timepoints of the entropy of 4 different arrays, with Dunn's test results above.

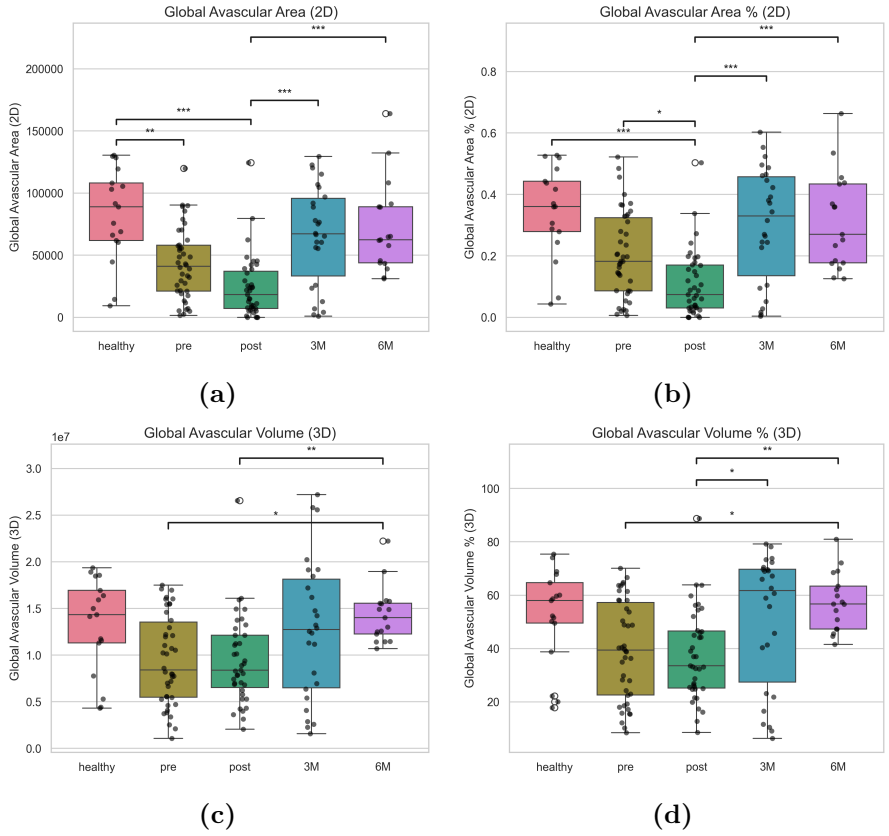


Figure 3.12: Boxplots across timepoints of 4 different avascular parameters, with Dunn's test results above.

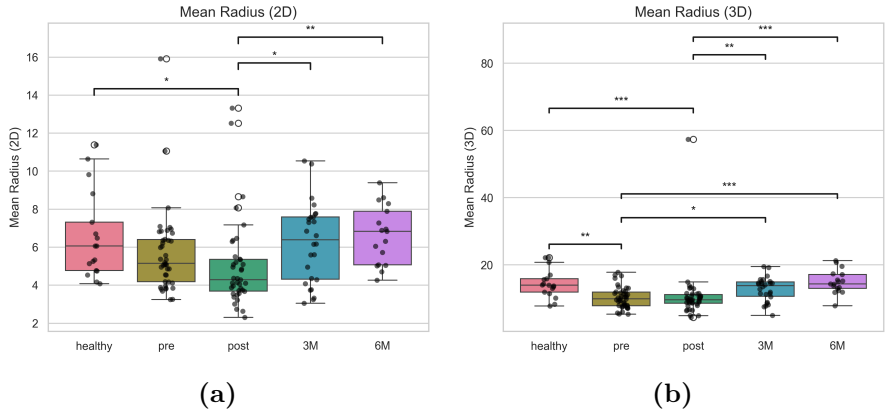


Figure 3.13: Boxplots of mean radius values across timepoints, in 2D and 3D.

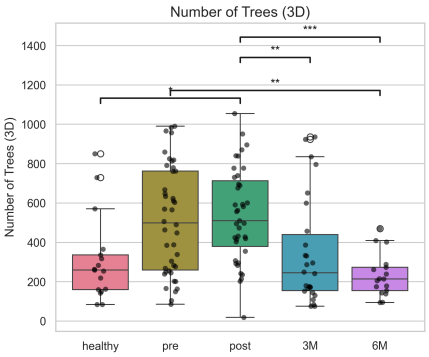


Figure 3.14: Boxplots of the number of trees across timepoints.

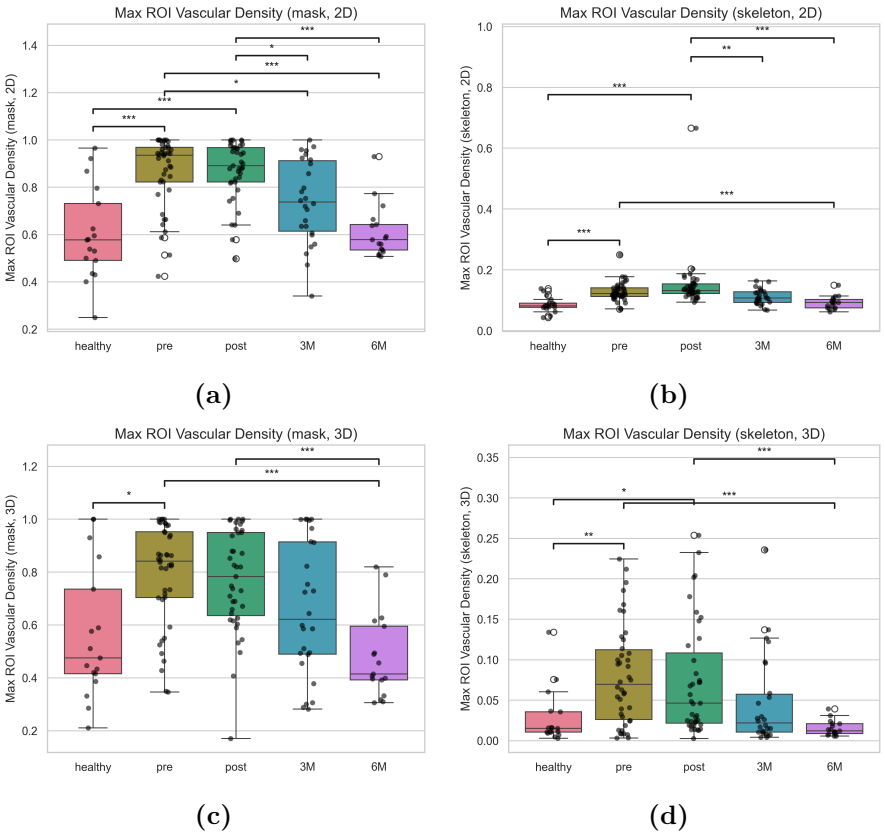


Figure 3.15: Boxplots across timepoints of the maximum vascular density values of ROIs into which 4 different arrays were divided, with Dunn's test results above.

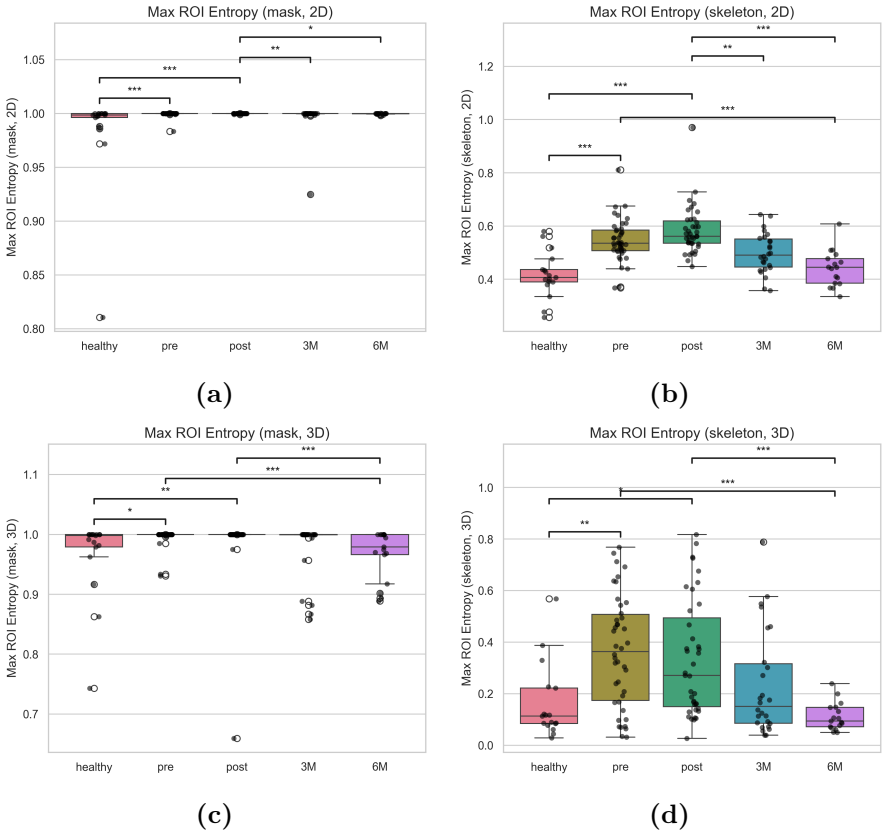


Figure 3.16: Boxplots across timepoints of the maximum entropy values of ROIs into which 4 different arrays were divided, with Dunn's test results above.

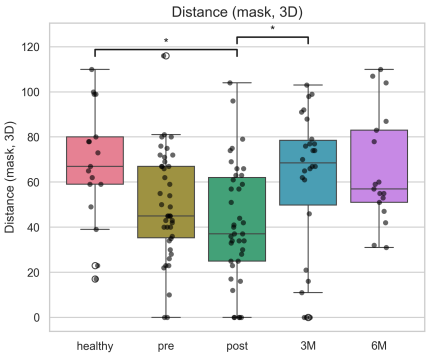
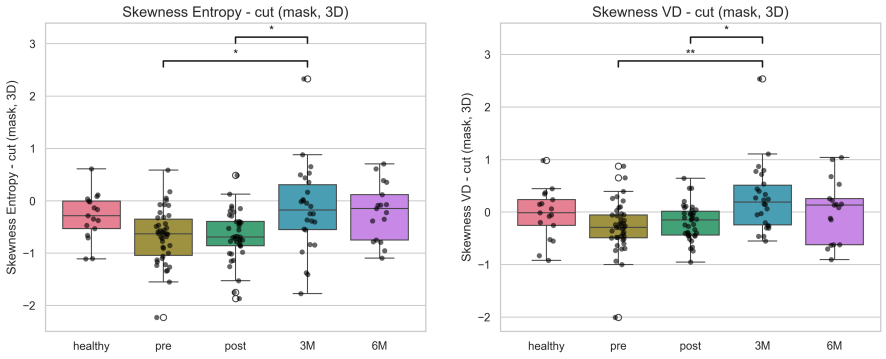


Figure 3.17: Boxplots across timepoints of the distance from the surface of the innermost point of the vasculature segmentation mask.



(a) **(b)**

Figure 3.18: Boxplots across timepoints of the skewness of vascular density and entropy along with depth, with Dunn's test results above.

3.2.1 Results without the NaN masks

If the NaN masks are not used, there are 35 significant features, of which 19 present a p-value under 0.0001, and 29 have a large size effect. The only feature that gains its significance with the NaN masks is the mean radius calculated for the 2D arrays. Therefore, it may seem that they might not be affecting results in a relevant way. The post-hoc analysis proves otherwise.
In fact, by looking at the comparison shown in Figure 3.19, it's clear that the NaN masks allow for greater distinction between the

timepoints, that overlap much less in the presented examples. This is further proved by the results of Dunn's test, that show more asterisks in the right column of the figure.

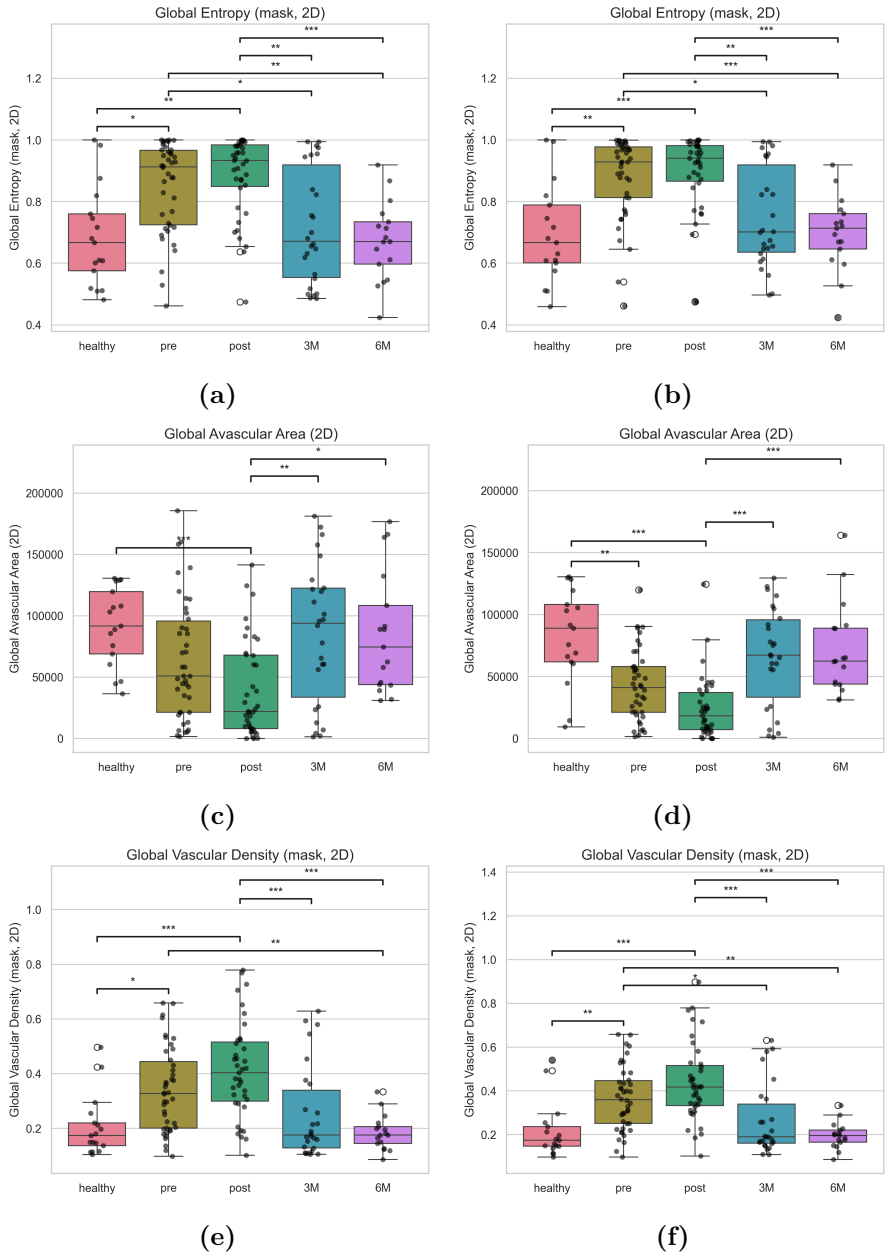


Figure 3.19: Comparison between the boxplots of some features, calculated before applying the NaN masks (a,c,e) and after (b,d,f).

Chapter 4

Discussion

4.1 Conclusions

The results spark a series of observations:

- **p-values and η^2**: many parameters assume significantly different values, based on the timepoint they are calculated in. It's observed that the variance in these values is largely due to the timepoint group to which they belong. These two assessments, that depend respectively on the p-value and η^2, support each other. Therefore, it's not far-fetched to say that quantitative measures extracted from OCTA data change consistently depending on the time elapsed from the start of radiotherapy.

- **Dunn's test**: the previous observation, while precious, is not enough to prove the usefulness of these features, as there's no assessment regarding which timepoints change from the others, and which trend they present and how strong this trend is. Let's take Figure 3.10b: its corrected p-value is higher than its 3D counterpart (Figure 3.10d), and, while the boxplots slowly converge back to the healthy baseline, the 3M is still quite similar to the pre- and post-radiotherapy timepoints. However, both the healthy and 6M groups are significantly different from the the post group. The trend of Figure 3.10d is much stronger: boxplots in 3M, 6M, and healthy are aligned, significantly different from

the pre- and post-radiotherapy ones, and they also have similar IQRs.
Overall, optimal features behave as such:

- significant difference between pre-radiotherapy group and healthy baseline: an untreated lesion is expected to have a dissimilar vascular architecture with respect to healthy skin [7] [12].
- values calculated immediately after radiotherapy are expected to be close in distribution to the previous timepoint. Not enough time is elapsed for the observer to expect significant changes.
- the 3M and 6M groups gradually shift towards the healthy boxplot and have similar IQRs. Not only this, but these groups may not be significantly different from the latter, as radiotherapy is being effective. Between them and the pre- and post-radiotherapy groups a p-value under 0.05 should be observed.

All of the features calculated for this thesis follow, more or less, this ideal behaviour, with a convergence of values towards the healthy baseline sometimes already from the 3M timepoint.

It's equally crucial to showcase which changes, specifically, are described by the features. These are the differences in vasculature that are observed between healthy skin and lesions:

- **Higher density**: the vascular architecture in lesions is denser (Figures 3.9 and 3.15), with the branches being more numerous, shorter and of smaller girth (Figure 3.8c, Figure 3.8d and Figure 3.13).
- **More complexity**: blood vessels in the en-faces are less coiled and tortuous (Figure 3.7) but, overall, lesions have a more complex and chaotic layout (Figure 3.10, Figures 3.11 and 3.16). Moreover, there's a higher number of vascular beds, that branch more and towards more directions (Figure 3.14, Figures 3.8a and 3.8b).

- **Less avascularities**: avascular regions in lesions are less extended (Figures 3.12.
- **Different distribution with depth**: in lesions, the vascular mass appears to be closer to the surface and with a higher density towards the epidermis (Figure 3.17 and Figure 3.18).

It's fundamental to calculate features for the whole volumes, as their values capture general characteristics at a glance. In addition, the calculation of these parameters also on the 2D flattened vascularization proved to be useful for a series of reasons:

- **Branch parameters**: due to the lower number of elements, computation is much faster. This is, however, just a minor aspect that could be fixed in the future by optimizing the pipeline.
- **Reduction of noise and artifacts**: as these are flattened arrays, noise has a lower effect than in the 3D volumes, allowing for more significant results in 2D than in 3D.
- **Unique characteristics**: en-face visualizations allow to capture vascular patterns that might otherwise be overlooked and are usually used in dermoscopy for diagnosis.

Lastly, it's important to remind that the NaN masks allowed for greater statistical significance between the timepoints, showing the importance of ignoring those useless sections that are bound to be present inside these volumes.

4.1.1 Clinical interpretation of the results

Neo-angiogenesis is an established characteristic of skin tumors. Blood vessels need to grow more extensively in order for NMSC to keep developing, and the scope of this phenomenon has been linked to tumor relapse [17] [23]. Therefore, vascular density is an essential biomarker to describe vascular regrowth, alongside metrics such as the number of trees, branches, endpoints and branchpoints.

However, other parameters that quantify changes in transport efficiency from vasculature to tissue are important as well. The avascularity features, AA and AV, show a reduced number of voids in lesions, as nutrients need to reach more areas to feed tumor growth [23].
In NMSC, some vessels appear larger than in healthy skin and others are very tiny and fragile [17]. Here, the average radius has been detected to be lower in lesions, reflecting the larger influence of these new, smaller vessels that have grown to reach new cells. This observation is supported by the fact that branches, on average, are also detected to be shorter than in healthy tissue.
Moreover, this growth happens in a disorderly way, favoring fast growth over efficiency, so that lesions have higher information entropy and fractal dimension values. These parameters thus describe the vascular architecture in NMSC as more complex.
Consequences that seemingly contrast with these previous observations come from the tortuosity values, that are significantly lower in lesions. These results are actually due to lesion-specific characteristics that stand out in en-face visualizations and are traditionally observed in dermoscopy. In this technique, the vascular architectures of NMSC lesions can present globular vessels, in the form of dots and blobs, and arborized vessels [17] [29].
Lastly, the distance metric confirms what had already been observed, which is that vessels tend to appear more superficially in NMSC lesions than in healthy skin [30].
Overall, it can be said that the features calculated in this thesis allow for an accurate assessment of the lesion-dependent vasculature peculiarities that are often described in literature.

4.2 Limitations and future developments

A major limitation in the acquisition protocol stands with the shape of the probe and of the glass slide. If OCTA were to become of common use in dermatology, a more mobile probe would be necessary,

perhaps with an arm that allows to twist it to reach more difficult positions. This would be very useful especially with elder patients, that often feel unconformable in moving into positions that allow the probe to be placed correctly. Moreover, the glass slide is extremely important, but the rectangular ones, used in microscopy, cannot be applied in curved zones, like the nose. The development of more appropriate slides will be useful in the future, if this technology is to be used in clinical settings. These future prospects, being heavily tied with budget, are more difficult to reach in short-term. However, the more OCTA solidifies as a useful technology in clinical practice, the more such proposals become more likely, in the future.
Concerning the vasculature segmentation themselves, while AMIRA allows to create very accurate masks, doing so semi-automatically takes a lot of time from the operator, and creates a bottleneck. Therefore, implementing an automatic segmentation method in this pipeline would make for a seamless transition from processing to feature extraction [31].
The last future upgrades to the pipeline regard the feature extraction process:

- **Code optimization**: the code may be optimized to allow for the calculation of branch features for the 3D volumes too. Tortuosity measures on 3D vascularizations would yield complementary results to those obtained from the en-faces.
- **DEJ inclusion**: as previously said, sometimes it's not possible to detect the DEJ. This can be due to lesion-dependent structural changes: for example, SCC and AK can present thick scales that obscure it. However, it could be worth it to explore the possibility of exploiting this as a possible feature, as, sometimes, not being able to find the DEJ is the direct cause of a pathological state.
- **Additional depth-related considerations**: even if three skewness measures showed statistically significant differences, these parameters' distributions with depth should be investigated more, to help with the interpretation of the results.

Research should also be done towards the identification of the depth reached by each lesion, and its evolution throughout the timepoints. This measure could be achieved with gradient-dependent assessments on OCT A-lines.

- **Dermoscopy-based features**: features that are often used in dermoscopy could potentially be implemented in the 2D section of the pipeline [29].

- **Radiomics**: it has been demonstrated that OCTA radiomics can be potentially combined with vascular parameters such as the ones calculated in this thesis. First-order and texture features have been shown to differentiate between pre- and post-radiotherapy OCTA acquisitions, and more studies are being conducted in this direction [32].

The addition of these features to the existing pipeline might help in distinguishing specific lesions, namely BCC and SCC, that can present different structures at the en-faces and different distributions with depth [29] [30]. This, of course, depends heavily on the amount of available data, as much of the specificity presented in this thesis is due to the large number of analyzed volumes. If BCCs, SCCs and AKs had been separated, this specificity would have been lost. This is why this thesis characterized changes with respect to a healthy baseline, which has been proven here to have quantifiable differences from NMSCs.
Studies that develop methods and classifiers that integrate vascular features along with radiomics, will open the door to patient-specific radiotherapy treatment planning and follow-up, representing a promising direction towards precision medicine.

4.3 Final conclusions

Before drawing the last considerations, some important aspects of the two Python pipelines ought to be highlighted:

- **`OCTAProcessing` class**: it does the whole computation in 5-10 minutes, and wholly automatically, if the user wishes. It produces the OCTA volume, cleaned from artifacts, and a useful series of figures: B-scan visualizations, the rendering of the detected skin surface, en-face images of both the OCT and the angiographic volume. The quicker the clinicians get these images, the more efficiently they may use their time, as further assessments could be made on the lesion before the patient exits the examination room.
- **`OCTAFeatures` class**: it provides automatically a large amount of features in a suitable format for any statistical analysis. Its modularity allows for any modifications, additions, and combinations one may want to make.

These pipelines come from the translation into Python of previously existing Matlab pipelines, that have been optimized and expanded. They are very useful, as they build the foundations for the reproducibility of these results, which is crucial for the diffusion of the technique.
This thesis demonstrated that it's possible to quantitatively assess radiotherapy-induced microvasculature changes in non-melanoma skin cancer using OCTA. This thesis showed that these OCTA features are not only a device for future prediction algorithms, but give interpretable and clear information about the evolution of the lesion. This is particularly valuable to establish the use of OCTA for decision-making in clinical settings.

Bibliography

[1] Vishal Madan, John T Lear, and Rolf-Markus Szeimies. «Non-melanoma skin cancer». In: *The Lancet* 375 (2010), pp. 673–685. DOI: 10.1016/S0140-6736(09)61196-X (cit. on p. 1).

[2] Venura Samarasinghe and Vishal Madan. «Nonmelanoma Skin Cancer». In: *Journal of Cutaneous and Aesthetic Surgery* 5 (2012), pp. 3–10. DOI: 10.4103/0974-2077.94323 (cit. on p. 1).

[3] Sofian Benkhaled, Dirk Van Gestel, Carolina Gomes da Silveira Cauduro, Samuel Palumbo, Veronique del Marmol, and Antoine Desmet. «The State of the Art of Radiotherapy for Non-melanoma Skin Cancer: A Review of the Literature». In: *Frontiers in Medicine* 9 (2022). DOI: 10.3389/fmed.2022.913269 (cit. on p. 1).

[4] Michael J Veness, D Delishaj, EA Barnes, A Bezugly, and Agata Rembielak. «Current Role of Radiotherapy in Non-melanoma Skin Cancer». In: *Clinical Oncology* 31 (2019), pp. 749–758. DOI: 10.1016/j.clon.2019.08.004 (cit. on pp. 1, 2).

[5] Haoran Chen, Zhongyu Han, Qian Luo, Yi Wang, Qiju Li, Lisui Zhou, and Houdong Zuo. «Radiotherapy modulates tumor cell fate decisions: a review». In: *Radiation Oncology* 17 (2022). DOI: 10.1186/s13014-022-02171-7 (cit. on p. 2).

[6] Saint John's Cancer Institute. *Radiation Therapy for Melanoma*. https://www.saintjohnscancer.org/melanoma/treatment/radiation/. Accessed: 2025-10-11. 2025 (cit. on p. 2).

[7] Mengyang Liu and Wolfgang Drexler. «Optical coherence tomography angiography and photoacoustic imaging in dermatology». In: *Photochemical & Photobiological Sciences* 18 (2019), pp. 945–962. DOI: 10.1039/C8PP00471D (cit. on pp. 3, 9, 67).

[8] Rainer Leitgeb et al. «Enhanced medical diagnosis for dOCTors: a perspective of optical coherence tomography». In: *Journal of Biomedical Optics* 26 (2021), pp. 100601–100601. DOI: 10.1117/1.JBO.26.10.100601 (cit. on p. 3).

[9] Silke Aumann, Sabine Donner, Jörg Fischer, and Frank Müller. «High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics [Internet]». In: Springer, 2019. Chap. 3 , Optical Coherence Tomography (OCT): Principle and Technical Realization. DOI: 10.1007/978-3-030-16638-0_3 (cit. on pp. 3, 4).

[10] Wolfgang Drexler, Mengyang Liu, Abhishek Kumar, Tschackad Kamali, Angelika Unterhuber, and Rainer A. Leitgeb. «Optical coherence tomography today: speed, contrast, and multimodality». In: *Journal of Biomedical Optics* 19 (2014). DOI: 10.1117/1.JBO.19.7.071412 (cit. on pp. 3, 5–8, 30).

[11] Adrian Podoleanu. «Optical coherence tomography». In: *British Journal of Radiology* 78 (2005), pp. 976–988. DOI: 10.1259/bjr/55735832 (cit. on p. 6).

[12] Kristen M. Meiburger, Massimo Salvi, Giulia Rotunno, Wolfgang Drexler, and Mengyang Liu. «Automatic Segmentation and Classification Methods Using Optical Coherence Tomography Angiography (OCTA): A Review and Handbook». In: *Applied Sciences* 11 (2021). DOI: 10.3390/app11209734 (cit. on pp. 7, 9, 67).

[13] Josep Malvehy and Giovanni Pellacani. «Dermoscopy, Confocal Microscopy and other Non-invasive Tools for the Diagnosis of Non-Melanoma Skin Cancers and Other Skin Conditions». In: *Acta Dermato-Venereologica* 97 (2017), pp. 22–30. DOI: 10.2340/00015555-2720 (cit. on pp. 9, 30).

[14] Leopold Veselka, Lisa Krainz, Leonidas Mindrinos, Wolfgang Drexler, and Peter Elbau. «A Quantitative Model for Optical Coherence Tomography». In: *Sensors* 21 (2021). DOI: 10.3390/s21206864 (cit. on p. 13).

[15] Annan Li, Jun Cheng, Ai Ping Yow, Carolin Wall, Damon Wing Kee Wong, Hong Liang Tey, and Jiang Liu. «Epidermal segmentation in high-definition optical coherence tomography». In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015. DOI: 10.1109/EMBC.2015.7319034 (cit. on pp. 29, 31).

[16] @goldbema. *WeightedLeastSquaresFilter*. https://github.com/goldbema/WeightedLeastSquaresFilter. Accessed: 2025-09-27 (cit. on p. 29).

[17] Sandra Schuh, Jon Holmes, Martina Ulrich, Lotte Themstrup, Gregor B. E. Jemec, Nathalie De Carvalho, Giovanni Pellacani, and Julia Welzel. «Imaging Blood Vessel Morphology in Skin: Dynamic Optical Coherence Tomography as a Novel Potential Diagnostic Tool in Dermatology». In: *Dermatology and therapy* 7 (2017), pp. 187–202. DOI: 10.1007/s13555-017-0175-4 (cit. on pp. 30, 68, 69).

[18] Richard F. Spaide, James G. Fujimoto, and Nadia K. Waheed. «IMAGE ARTIFACTS IN OPTICAL COHERENCE TOMOGRAPHY ANGIOGRAPHY». In: *Retina* 35 (2015), pp. 2163–2180. DOI: 10.1097/IAE.0000000000000765 (cit. on p. 32).

[19] Woo June Choi, Bjorn Paulson, Sungwook Yu, Ruikang K. Wang, and Jun Ki Kim. «Mean-Subtraction Method for De-Shadowing of Tail Artifacts in Cerebral OCTA Images: A Proof of Concept». In: *Materials* 13 (2020). DOI: 10.3390/ma13092024 (cit. on p. 32).

[20] scikit-image developers. *Morphological Filtering*. https://scikit-image.org/docs/0.25.x/auto_examples/applications/plot_morphology.html. Accessed: 2025-09-29 (cit. on p. 38).

[21] scikit-image developers. *Skeletonize*. https://scikit-image.org/docs/0.25.x/auto_examples/edges/plot_skeleton.html. Accessed: 2025-09-29 (cit. on p. 38).

[22] Kristen M. Meiburger et al. «Automatic skin lesion area determination of basal cell carcinoma using optical coherence tomography angiography and a skeletonization approach: Preliminary results». In: *Journal of biophotonics* 12 (2019). DOI: 10.1002/jbio.201900131 (cit. on p. 40).

[23] Nader Allam, W. Jeffrey Zabel, Valentin Demidov, Blake Jones, Costel Flueraru, Edward Taylor, and I. Alex Vitkin. «Longitudinal in-vivo quantification of tumour microvascular heterogeneity by optical coherence angiography in pre-clinical radiation therapy». In: *Scientific Reports* 12 (2022). DOI: 10.1038/s41598-022-09625-8 (cit. on pp. 42, 68, 69).

[24] ChatzigeorgiouGroup. *FractalDimension: A python function to calculate fractal dimension in 3D*. https://github.com/ChatzigeorgiouGroup/FractalDimension. Accessed: 2025-09-30 (cit. on p. 44).

[25] Patrick E. McKight and Julius Najab. «Kruskal-Wallis Test». In: *The corsini encyclopedia of psychology* (2010). DOI: 10.1002/9780470479216.corpsy0491 (cit. on p. 47).

[26] SciPy Developers. *scipy.stats.kruskal*. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html. Accessed: 2025-09-29 (cit. on p. 48).

[27] Hervé Abdi. «Holm's Sequential Bonferroni Procedure». In: *Encyclopedia of research design* 1 (2010), pp. 1–8. DOI: 10.4135/9781412961288.n178 (cit. on p. 48).

[28] John T.E. Richardson. «Eta squared and partial eta squared as measures of effect size in educational research». In: *Educational Research Review* 6 (2011), pp. 135–147. DOI: 10.1016/j.edurev.2010.12.001 (cit. on pp. 48, 49).

[29] Yaei Togawa. «Review of vasculature visualized on dermoscopy». In: *The Journal of Dermatology* 44 (2017), pp. 525–532. DOI: 10.1111/1346-8138.13686 (cit. on pp. 69, 71).

[30] V. Sigsgaard, L. Themstrup, P. Theut Riis, J. Olsen, and G. B. Jemec. «In vivo measurements of blood vessels' distribution innon- melanoma skin cancer by dynamic optical coherencetomography — a new quantitative measure?» In: *Skin Research and Technology* 24 (2018), pp. 123–128. DOI: 10.1111/srt.12399 (cit. on pp. 69, 71).

[31] Giulia Rotunno et al. «DERMA-OCTA: A Comprehensive Dataset and Preprocessing Pipeline for Dermatological OCTA Vessel Segmentation». In: *Scientific Data* 12 (2025). DOI: 10.1038/s41597-025-05763-6 (cit. on p. 70).

[32] Gerd Heilemann et al. «Texture analysis of optical coherence tomography angiography for detecting microstructural changes in skin cancer lesions post kV-based radiotherapy». In: *Radiotherapy and Oncology* 206 (2025). DOI: 10.1016/S0167-8140(25)00462-1 (cit. on p. 71).