

POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering



Master's Degree Thesis

"GPU-accelerated algorithm for real-time needle tracing in tissue-mimicking silicone phantoms using optical coherence tomography (OCT)"

Supervisors

Prof. Kristen MEIBURGER

Prof. Filippo MOLINARI

Dr. Mengyang LIU

Candidate

Simone REGALDO

Academic Year 2020/2021

Acknowledgements

My two thesis advisors, Prof. Meiburger and Dr. Liu, for giving me the great opportunity to have such a wonderful and professional experience abroad.

The whole Center for Medical Physics and Biomedical Engineering of the Medical University of Vienna. Especially to Prof. Drexler, and the always helpful Lisa Krainz for all the support she gave me and the patience she had during all the imaging session.

Giulia and Max, for all the funny moments we had together in Vienna and for putting up with me when I could only spoke about my algorithm errors or how far behind I was in my project.

My lifetime friends, for all the carefree moments we had during this long long university journey.

My crazy big family for all the support I had from them. Especially my grandparents, the ones who are still with me and the ones who are somewhere in the sky but always on my side.

My mum and dad, for never having made me miss anything in my life.

Last but not the least my Pissi. She is my strength, my courage, my weakness, my bestfriend, my only love. If I finally managed to close this chapter of my life, it is only because of her.

Summary

Since the last two decades, Optical Coherence Tomography (OCT) started to find applications not only in the ophthalmologic field of the medicine. The use of OCT techniques in dermatology allows to obtain high resolution images (3 to 15 μ m) in a completely non-invasive way and using non-ionizing radiations.

Thanks to the recent improvements, image acquisition with OCT technique has reached really high scanning rates. Therefore, OCT image reconstruction algorithm have been developed by using GPU accelerated algorithm in order to acquire, process and visualize OCT volumes in real time. The GPUs architecture provides the developers to process huge amount of data in parallel making real time applications achievable. These applications are extremely useful in image-guided surgery.

The goal of this work, done in collaboration with the Medical University of Vienna, is to demonstrate the feasibility of a real-time application to trace a needle inside a silicone phantom. A plastic black ball is positioned inside the phantom in order to simulate the biopsy area so to calculate how much the needle tip is distant from it.

The image reconstruction algorithm exploit the GPU acceleration to reconstruct the entire volume acquired from one of the Swept Source OCT of the Physics and Biomedical Department of the Medical University of Vienna.

The needle tracing algorithm uses a Fully Convoluted Neural Network (U-net) to segment the needle and Principal Component Analysis (PCA) based method to measure the needle orientation. These techniques are the core of the algorithm, they allow to segment the needle, to measure the depth on the needle tip and its distance from the biopsy-mimicking area.

Table of Contents

List of Figures	VII
1 Optical Coherence Tomography	1
1.1 OCT main principles and terminology	1
1.1.1 Michelson interferometer	2
1.1.2 Low coherence interference	3
1.1.3 Light sources	4
1.1.4 Data terminology	5
1.2 Mathematical treatment	6
1.3 OCT signal properties	8
1.3.1 Probing depth	8
1.3.2 Sensitivity	9
1.3.3 Resolution	10
1.3.4 Penetration depth	12
1.4 OCT techniques	12
1.4.1 Time-domain OCT (TD-OCT)	12
1.4.2 Fourier-domain OCT (FD-OCT)	13
1.5 OCT system	14
2 Image reconstruction	19
2.1 Volume reconstruction algorithm	20
2.2 Volume registration	25
2.3 Timings	27
3 Tissue-mimicking phantom	30
3.1 Materials	30
3.2 Design	31
3.2.1 Biopsy area-mimicking plastic inclusion	33
3.3 Phantom VS reality	34

4	Needle tracing algorithm	36
4.1	Convolutional Neural Network (CNN) VS Fully Convolutional Network (FCN)	37
4.1.1	U-net	41
4.1.2	Inputs and network training	43
4.1.3	Outputs and performance	46
4.2	Algorithm	49
4.3	Simulations	57
4.4	Errors	61
4.5	Timings	62
5	Conclusions	64
5.1	Limitations	64
5.2	Challenges and future work	66
	Bibliography	67

List of Figures

1.1	OCT scheme based on a low time-coherence Michelson interferometer	3
1.2	Representation of a perfectly coherent light	4
1.3	Spatial disposal of the OCT planes	6
1.4	Geometry of sample and probing beam	9
1.5	OCT system imaging setup	15
1.6	Reference arm top view	16
1.7	Reference arm	17
1.8	Sample arm	18
1.9	Power supplies and laser source	18
2.1	Basic OCT reconstruction algorithm	20
2.2	Volume reconstruction algorithm	21
2.3	Light source spectrum profile adopted in the OCT system	22
2.4	Comparison between non-windowed and windowed B-scan	23
2.5	Mirror artifact	24
2.6	OCT volume from the forearm skin	24
2.7	Volume inclination	25
2.8	Volume registration	26
2.9	Total processing timings for the volume reconstruction	28
2.10	Reconstruction algorithm timings	28
2.11	Registration timings	29
3.1	Pictures of the fabricated phantom	33
3.2	Comparison of different colored inclusion	34
3.3	B-scan comparison between phantom and skin	35
4.1	Convolution as alternative for fully connected network	38
4.2	Zero-padding	38
4.3	Activation functions	39
4.4	Max-pooling	40
4.5	Fully-connected layer	40

4.6	CNN to FCN transformation	41
4.7	Custom U-net architecture	42
4.8	Intersection over Union equation	44
4.9	Accuracy and IoU score comparison over the nine training sets with different loss functions	45
4.10	Training and validation accuracy trend over the epochs	46
4.11	Training and validation loss trend over the epochs	47
4.12	Network performance	47
4.13	U-net output example	48
4.14	Needle visualization in OCT images	49
4.15	Tracing algorithm concept	50
4.16	First scanning step	51
4.17	Visual comparison between different side needle insertion	52
4.18	Second scanning step	54
4.19	Correction of the distance and xz-angle calculation	55
4.20	Example of the final algorithm output	56
4.21	Phantom locking system	57
4.22	First simulation	58
4.23	Second simulation	59
4.24	Third simulation	60
4.25	Needle tip position error measurement	61
4.26	Tracing needle tip error on x, y, and z-direction	62
4.27	Tracing algorithm performed over the y-direction	63
4.28	Tracing algorithm performed over the x-direction	63

Chapter 1

Optical Coherence Tomography

The Optical Coherence Tomography (OCT) is an optical imaging technique used in the medical field that has achieved quite a success, especially in ophthalmology. Over the last two decades, OCT techniques were employed, as an excellent diagnostic tool, in different medical fields than ophthalmology, such as dermatology. In this chapter, advantages and limitations of this imaging technique will be analyzed.

1.1 OCT main principles and terminology

Why this technique was named OCT? Let's consider each term of the acronym.

- *Optical*
OCT is an optical imaging technique: it uses a laser source in order to image skin volumes. Light is an electromagnetic wave, and a light beam can be characterized by parameters as amplitude and wavelength.
- *Coherence*
Coherence is a property of the light. Two light waves have perfect coherence when they are characterized by identical frequencies, waveforms, and constant phase difference. Thus, only monochromatic light beams can produce perfectly coherent light.
- *Tomography*
OCT techniques is a tomographic technique. Indeed, it allows to image 2D scans of the sample and recreate 3D tissue volume by raster-scanning.

A fourth essential term needs to be understood to fully comprehend how this technique works. The phenomenon of *interference* can be observed when an interaction between two or more electromagnetic waves occurs. The interference results in the sum of the waves involved. The interference can be *constructive* or *destructive*. Constructive interference is observed when the waves phase difference is equal to 0° (waves are "*in-phase*"). On the other hand, destructive interference is observed when the phase difference is equal to 180° (waves are "*out-of-phase*").

The reflectivity of a tissue can be measured by the OCT as a function of depth using ballistic and non-ballistic photons. Initially, OCT techniques were based on Low time-Coherence Interferometry (LCI)[1]. Starting from the late 1990s, also spectral interferometry based OCT techniques were used[2].

The OCT concept is based on the working of the Michelson interferometer.

1.1.1 Michelson interferometer

Fig.1.1 shows the Michelson interferometer and its components: the low time-coherence light source, the reference mirror, the beam splitter, the sample and the photodetector.

The light beam is projected from the light source and then splitted into the reference beam and the sample beam by the beam splitter.

While, the sample beam hits the sample in exam, the reference beam hits the reference mirror placed at a known distance. Both the beams are reflected. After they traveled the same path in the opposite direction, they recombine at the beam splitter and the photodetector record the interference signal. Basically, the echo time delay of the two light beams is compared. The maximum interference signal is recorded when the two reflected beams interfere on the beam splitter simultaneously. The tomogram of the sample is reconstructed by the interferometer through integration of several data points collected in depth and laterally. For that reason there are basically two scan procedures in OCT. The depth-scan is performed by the reference mirror. The lateral-scan is performed by moving the sample or by scanning the probe beam which illuminates the sample.

As stated in Sec.1.1, perfect coherence can be obtained when a monochromatic light beam is used. In this case, the reference and the sample beams splitted from this type of light source are "*in-phase*" and a constructive interference can be observed if the pathlengths of the two beams are equal. If this condition occurs, it is possible to measure the exact depth of a certain reflective surface. Since the Michelson interferometer works with a low time-coherence light source, there is no possibility to locate this surface.

In the next section is presented a brief introduction to the low coherence

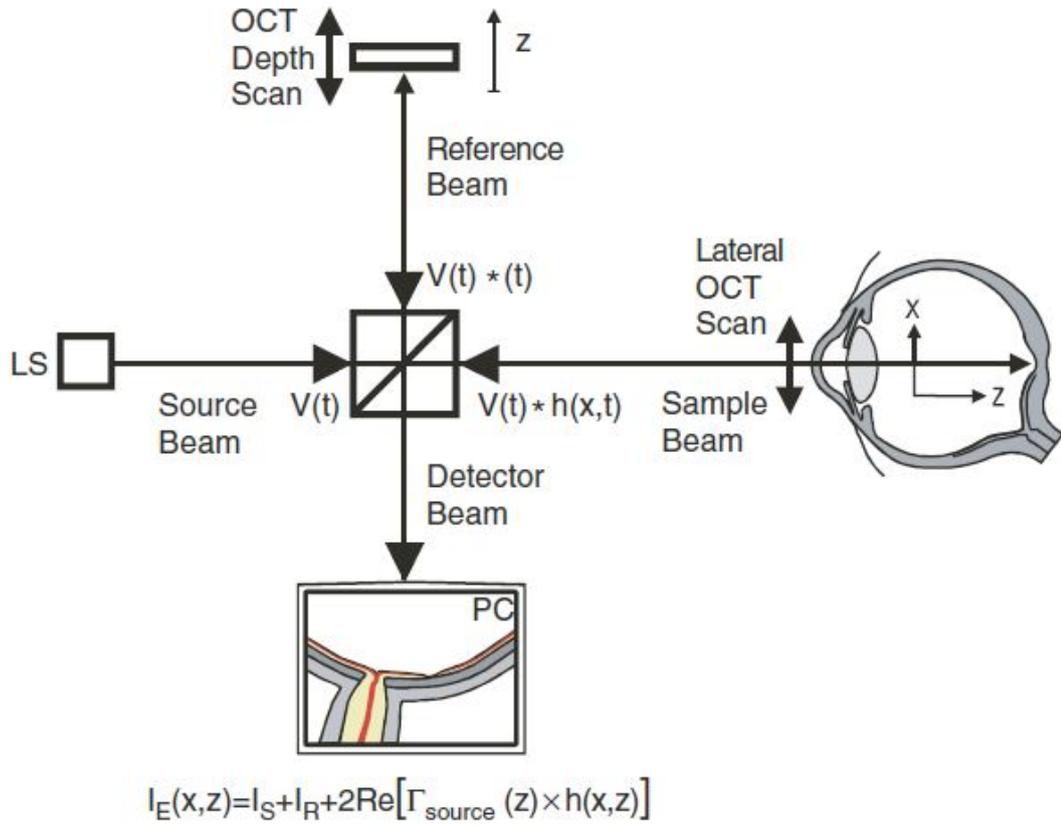


Figure 1.1: OCT scheme based on a low time-coherence Michelson interferometer (adopted from [3]). $V(t)$ = analytic signal, $h(x, t)$ = sample response, $I_E(x, z)$ = total intensity at the interferometer exit.

interference.

1.1.2 Low coherence interference

LCI is based on the occurrence of refractive depth profile (fringes) if the sample and the reference optical path lengths coincide withing the "coherence gate". The coherence gate size coincide with the so-called round trip *coherence length*.

The coherence length is defined as in [3]:

$$l_c = \frac{2 \ln 2}{\pi} \frac{\lambda_0^2}{\Delta\lambda}. \quad (1.1)$$

λ_0 = central wavelength and $\Delta\lambda$ = spectral width, assuming a Gaussian shape of the spectrum of the emitted light. As shown in Eq.1.1, the coherence length is

proportional to the light wavelength squared over the bandwidth.

Almost perfect coherence can be reached using narrow bandwidth light source. Narrow bandwidths generate lot of maxima in the interference signal. Using a wider bandwidth, only one maximum is observable in the interference signal when the path lengths of the sample beam and of the the reference beam are equal.

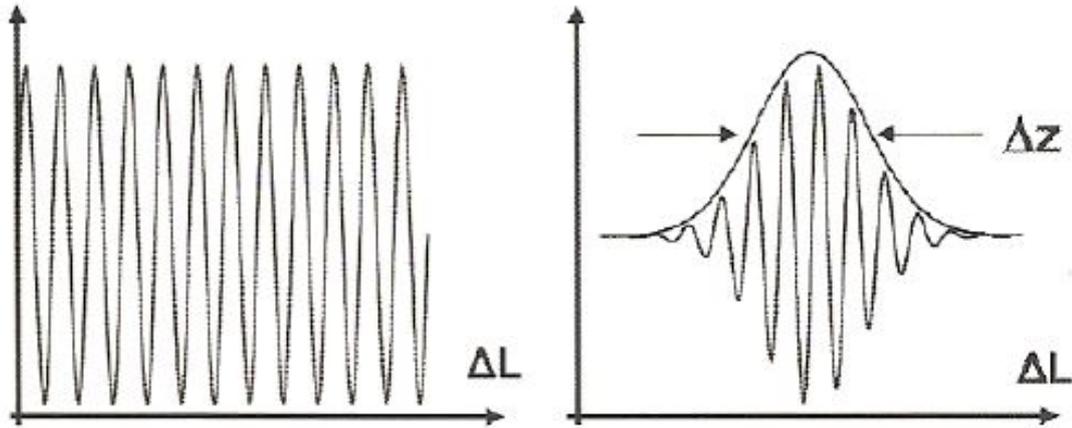


Figure 1.2: Representation of a perfectly coherent light (left) and a low coherent light (right) (adopted from [3]). Δz is the coherence length.

1.1.3 Light sources

Different types of light sources are used in OCT systems: SuperLuminescent Diodes (SLD), Light Emitting Diode (LED), akinetic swept sources (SS), superfluorescence doped fibers, Kerr lens, SLD light sources are the most popular, but the more recent SS technology is taking the scene because of its higher scanning rates.

A careful selection of the light source is important in order to reduce noise, do not damage the sample, reduce the scattering and increase the probing depth (Section 1.3.1). The three main light source properties are: the coherence properties, the wavelength and the shape of the spectrum.

- *Coherence properties:* Depth resolution is defined by temporal coherence, while both lateral and depth resolution are spatial coherence dependent (Section 1.3.3). For this reason, different techniques were developed to measure light sources coherence properties [4].
- *Wavelength:* Source power and emission wavelength determine the penetration

depth (Section 1.3.4). Absorption of biomedical materials is wavelength-dependent [3]: for most of the tissue, by increasing the wavelength the absorption decreases and the scattering (nearly monotonically) too. Since the red end of the visible spectrum provides the so-called "therapeutical window", the light sources wavelengths span from 675nm to 1800nm, depending on the technology adopted.

- *Spectral structure*: By spectrum manipulation, it is possible to control three parameter of the coherence function: width, smoothness and position.
 - *Spectral width*. The spectral width $\Delta\lambda$ defines the depth resolution as well as the spectrometric resolution.
 - *Spectral modulation*. The Point Spread Function (PSF) is defined as twice the real part of the coherence function $\Gamma_{Source}(\tau)$ as shown by Eq.1.3 in Sec.1.2. The coherence function equals the inverse FT of the PSF. Thus, it must have a narrow half-width, no satellites, and smooth shape of the source power spectrum.
 - *Spectral phase*. Position of the coherence function can be shifted through a phase shift of its Fourier components as stated by the *Fourier shift theorem*

1.1.4 Data terminology

A tomogram is a two-dimensional image which shows a slice or a section of a three-dimensional object. In OCT, different terms are used depending on which spatial plane a tomogram is observed. Fig.1.3 shows the three-dimensional organization of the planes.

An *A-scan* corresponds to an axial (or depth) scan in the z direction. In simple terms, it is a one-dimensional scan of the sample. If the laser light source of the system uses a swept source technology, one A-scan is produced every sweep cycle.

A *B-scan* is composed by adjacent A-scans collected over the x -direction. Thus, a B-scan lays on the xz -plane. For better understanding the tracing algorithm this plane was called *B-scan plane*. The scan runs first alongside the x -axis before moving in the y -direction. For this reason, the x and the y -axis are respectively called fast and slow scanning direction.

Adjacent B-scans acquired alongside the y -direction build a three-dimensional structure called volume or *C-scan*.

En-face is the two-dimensional image that can be observed looking at the volume from a xy -plane parallel direction. They are widely used in Optical Coherence Tomography angiography.

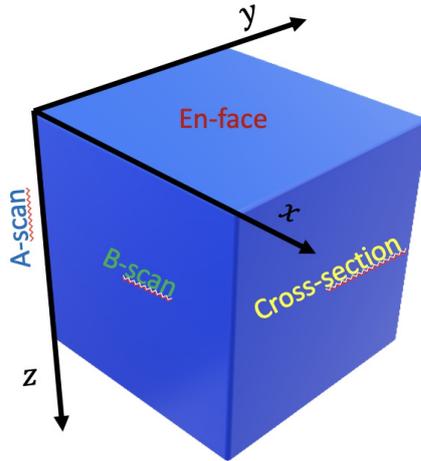


Figure 1.3: Spatial disposal of the OCT planes. B-scans, Cross-sections, and En-faces lay respectively on xz , yz , and xy -plane.

In this work, *cross-section* is used as a term to identify volume slices observed alongside the yz -plane. Cross-sections are nothing but B-scans from another dimension.

1.2 Mathematical treatment

Looking at the exit of the interferometer, the interference signal can be measured as a current by a photodetector. The averaged intensity at the interferometer exit is [3]:

$$\langle I_E(t; \Delta t) \rangle = \langle I_S(t) \rangle + \langle I_R(t) \rangle + G_{SR}(\Delta t). \quad (1.2)$$

$I_S(t)$ and $I_R(t)$ are two intensities referred respectively to the beam arriving at the detector if there would not be the reference mirror but only the scattering object. $G_{SR}(\Delta t)$ is the interference term or cross-correlation term.

Particular interest must be given to the **interference term**. This term is twice the real part of the cross-correlation of the analytic signals of the two interfering beams [3]:

$$G_{SR}(\Delta t) = 2Re\langle V_S^*(t)V_R(t + \Delta t) \rangle = 2Re\{\Gamma_{SR}(\Delta t)\}. \quad (1.3)$$

$\Delta t = \frac{\Delta z}{c}$ is the time delay of the back scattered sample beam on the reflected reference beam. Δz is the path difference between the beams, and c the speed of light.

The cross-correlation function $\Gamma_{SR}(\Delta t)$ is an analytic function, and it can be expressed by analytic continuation as [3]:

$$\Gamma_{SR}(\Delta t) = \frac{1}{2}G(\Delta t) + \frac{i}{2}HT\{G(\Delta t)\}, \quad (1.4)$$

where, $HT\{\}$ is the Hilbert transform, and $G(\Delta t)$ can be measured from the LCI signal.

LCI and OCT signals can be obtained by the photoelectric signal $U_G(t)$ of the interferogram G_{SR} . A photodiode is often used as detector in OCT configurations. The generated photocurrent can be expressed by the Eq.1.5. Via a transimpedance electronic amplifier circuit and due to its proportional relation with the incident light power, the photocurrent can be converted to voltage [3].

$$U_G(t) \propto i_G(t) = \frac{q_e \eta}{h\bar{\nu}} \int_{A_r(r)} G_{SR}(\mathbf{r}, t) d^2\mathbf{r}. \quad (1.5)$$

$i_G(t)$ = photoelectric current, q_e = electronic charge, η = quantum efficiency of the detector, h = Planck's constant, $\bar{\nu}$ = mean optical frequency, and $A_r(r)$ = sensitive detector area.

Frequently, photoelectric ac signal rectification and low-pass filtering are used to obtain the envelope of the LCI signal. If that solution can not be adopted, amplitude and phase of the photoelectric ac signal are measured by a lock-in amplifier.

Since the wavefronts of the interfering beams and the photodetector surface are coplanar, this reduction can be made [3]:

$$G_{SR}(\mathbf{r}, t) = G_{SR}(t) \propto i_G(t). \quad (1.6)$$

Then, the real envelope of the coherence function $\Gamma_{SR}(\Delta t) = A_\Gamma(t)e^{i\Phi_\Gamma(t)}$ and its phase are obtained respectively from Eq.1.7 and Eq.1.8 [3]:

$$A_\Gamma(t) = \frac{1}{2}\sqrt{(G_{SR}(t))^2 + (HT\{G_{SR}(t)\})^2} \quad (1.7)$$

$$\Phi_\Gamma(t) = \arctan \left[\frac{HT\{G_{SR}(t)\}}{G_{SR}(t)} \right]. \quad (1.8)$$

Thanks to the Wiener-Khintchine theorem, it can be useful to pass to the corresponding spectral relations. The power spectrum of a light wave can be obtained by Fourier Transform (FT) its self-correlation [3]:

$$S(\nu) = FT\{\Gamma(\Delta t)\}. \quad (1.9)$$

Furthermore, the Fourier transform of the cross-correlation function is equal to the cross-spectral density function of two waves [3]:

$$W_{SR}(\nu) = FT\{\Gamma_{SR}(\Delta t)\}. \quad (1.10)$$

The corresponding spectral domain of Eq.1.2 is expressed by Eq.1.11[3].

$$S(\nu; \Delta t) = S_S(\nu) + S_R(\nu) + 2Re [W_{SR}(\nu)] \cos(2\pi\nu\Delta t), \quad (1.11)$$

It represents the mathematical expression of the so called spectral interference law. Δt is the interferometric time delay.

1.3 OCT signal properties

The OCT is sensitive to discontinuities of the scattering potential. For this reasons, OCT signal is generated by discontinuities in the attenuation coefficient or in the refractive index of the target tissue. Its quality can be influenced by tissue and system characteristics. While single scattered light retrieves structural information about the underlying scattering potential, multiple scattered light retrieves tissue information like absorption coefficients, scattering anisotropy, scattering coefficients, and particle concentration in a solution. Furthermore, properties like probing depth and Signal-to-Noise Ratio (SNR) are influenced by multiscattered photons.

1.3.1 Probing depth

OCT imaging is based on light backscattering phenomena, but multiple scattered light does not contribute to the object's Fourier spectrum. It generates disturbing background with a consequent deterioration of the image quality (reduction of the image contrast, resolution, and penetration depth). Furthermore, multiple scattered photons contribute to increase the background noise (time-coherence and space-coherence decrease). Thus, it is vitally important to balance single and multiple scattered light.

The probing depth is defined as the depth where multiple scattering phenomena start to dominate and it depends principally on the numerical aperture of the imaging optic, NA. Optics with higher NA have higher resolution, but at the same time they detect an higher number of multiple scattered photons. It is important to find a good compromise between a good lateral resolution (high NA) and a reasonable depth of focus (low NA). The probing depth is determined by the path length of single and double scattered photons. Considering a double scattered photon, its first scattering event must occur within the illuminating probe beam cone. The second scattering event must be located inside or very close to the

coherent volume (coherence condition). The coherent volume is the product of the coherence gate depth and the corresponding beam cross section.

To better understand the geometry of sample and probing beam, can be useful to analyze Fig.1.3. The two double scattered photons do not contribute to the interferogram: photon A is outside the coherence gate, while photon B will miss the fibre or pinhole in front of the photodetector despite its trajectory is within the coherence gate.

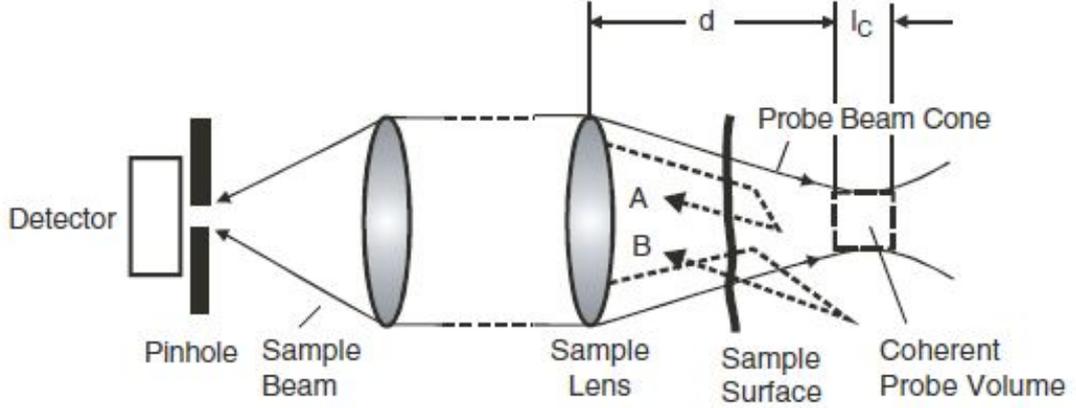


Figure 1.4: Geometry of sample and probing beam (adopted from [3])

1.3.2 Sensitivity

Considering a perfectly reflecting mirror ($R = 1$), the sensitivity is defined as the signal power over the weakest sample reflectivity ($R_{S,min}$) [3]:

$$S = \frac{1}{R_{S,min}} \Big|_{SNR=1} \quad (1.12)$$

From Eq.1.5, the interference term generates an effective signal photocurrent [3]:

$$i_G(t) = \frac{qe\eta}{h\nu} \sqrt{2P_S P_R}, \quad (1.13)$$

where, P_S = sample beam signal power and P_R = reference beam signal power.

In standard OCT devices, the dominating noises are: *shot noise* $\langle \Delta i_{sh}^2 \rangle$, *excess intensity noise* $\langle \Delta i_{ex}^2 \rangle$, and *receiver noise* $\langle \Delta i_{re}^2 \rangle$. Since receiver noise can be extrapolated from manufacturer's specifications, or modelled as thermal noise, the

resulting mean squared photocurrent noise will be [3]:

$$\langle \Delta i_p^2 \rangle = \langle \Delta i_{sh}^2 \rangle + \langle \Delta i_{ex}^2 \rangle = 2q_e B \langle i \rangle + (1 + \Pi^2) \langle i \rangle^2 \frac{B}{\Delta \nu_{eff}}, \quad (1.14)$$

where, B = electrical bandwidth, Π = degree of source polarization, $\langle i \rangle$ = mean detector photocurrent, and $\Delta \nu_{eff}$ = effective optical line width of the source.

The SNR is the ratio of the mean signal power and the noise power [3]:

$$SNR = \frac{2\alpha^2 P_S P_R}{\langle \Delta i_{sh}^2 \rangle + \langle \Delta i_{ex}^2 \rangle + \langle \Delta i_{re}^2 \rangle}, \quad (1.15)$$

with, $\alpha = q_e \eta / h \bar{\nu}$ and $\langle \Delta i_{re}^2 \rangle = 4k_B T B / \rho_L$ (k_B = Boltzmann's constant, T = absolute temperature and ρ_L = load resistor). Now, Eq.1.11 can be modified [3]:

$$S = \frac{\alpha^2}{8B} \frac{P_{Source}^2 R_R}{(\alpha/4) P_{Source} R_R \left[2q_e + \frac{(\alpha/4) P_{Source} R_R (1+\Pi)^2}{\Delta \nu_{eff}} \right] + \left(\frac{4k_B T}{\rho_L} \right)}, \quad (1.16)$$

Two limiting regimes can be identified in the two addenda of the denominator. The first one, where receiver noise dominates, and the second one, where the light source has high power. However, in the intermediate regime, shot noise dominates, the sensitivity can be approximated to [3]:

$$S = \frac{\alpha}{4} \frac{P_{Source}}{q_e} \frac{1}{B}. \quad (1.17)$$

In the range where the shot noise dominates, S is linearly dependent to the source power and inversely proportional to the electronic bandwidth. This is the regime usually realized in OCT systems because:

- a Receiver noise limits sensitivity at lower source power;
- b Excess noise limits the sensitivity gain at higher source power.

1.3.3 Resolution

The OCT resolution is split in *depth* (or *axial*) resolution and *transverse* (or *lateral*) resolution. They depend respectively on the bandwidth of the light source, and on the NA of the lens. Unlike other Computed Tomography (CT) techniques (e.g. x-rays or γ -ray tomography), depth resolution is decoupled from transverse resolution.

Considering a Gaussian amplitude spectrum applied to a FD-OCT, *depth resolution* is defined by the full width half maximum (FWHM) of the source intensity spectrum as in Eq.1.18 [3]:

$$\Delta z_{FWHM} = \frac{2 \ln 2}{\pi} \frac{\bar{\lambda}^2}{\Delta \lambda}. \quad (1.18)$$

$\bar{\lambda}$ = mean wavelength. In TD-OCT, the depth resolution equals half the PSF FWHM-width (l_{FWHM}). l_C is the so-called round trip coherence and it is used as a measure of depth resolution. Since the backscattered light has to travel twice through the sample, l_C is defined as in Eq.1.19 [3].

$$l_C = \frac{l_{FWHM}}{2} = \frac{2 \ln 2}{\pi} \frac{\bar{\lambda}^2}{\Delta \lambda_{FWHM}}. \quad (1.19)$$

Two are the main phenomena which deteriorate depth resolution:

- *Finite width of the depth PSF (Point Spread Function) envelope*: any ripple of the complex depth PSF spectral structure will lead to spurious image structures, and since the complex depth PSF is the inverse FT (Fourier Transform) of the power spectrum $S_{Source}(\nu)$, deconvolution is the natural technique to compensate these distortions.
- *Dispersion*: in order to balance the dispersion of the sample arm, a dispersive arm is usually deployed in the reference arm. That technique adds a frequency-dependent phase to the sample wave [1]:

$$\Phi_{Disp}(\omega) = k(\bar{\omega})z + k^{(1)}(\bar{\omega})(\omega - \bar{\omega})z + k^{(2)}(\bar{\omega})\frac{(\omega - \bar{\omega})^2}{2}z, \quad (1.20)$$

where, ω = frequency, $\bar{\omega}$ = central frequency of the source beam, and $k^{(j)}\bar{\omega}$ = j-th order dispersion. Normally, to compensate the dispersion and correct the phase shift, different digital techniques can be used. Note, until the bandwidth ($\omega - \bar{\omega}$) is small, the phase shift correction can be left out.

The *transverse resolution* depends on the transversal sampling rate and the probe beam diameter. Likewise depth resolution, transverse resolution is defined as the probe beam amplitude distribution's FWHM diameter (Δd_{FWHM}) at the beam waist of the focused probe beam [3]:

$$\Delta d_{FWHM} = 2\sqrt{\ln 2}\omega_0 = 2\sqrt{\ln 2}\frac{\bar{\lambda}}{\pi\sigma_S}, \quad (1.21)$$

where, Σ_S = angular spread of the Gaussian beam, and $\sin \Sigma_S$ = NA of the beam. Note that, using the Abbey's criteria, it is possible to define the transversal resolution by the numerical aperture of the objective lens [5]:

$$\Delta x = 0.61\lambda_0/NA. \quad (1.22)$$

Transverse resolution is not only dependent on the NA, but also to depth. Thus, it is important to introduce the *Depth Of Focus* (DOF). It is defined as twice the Rayleigh range z_R :

$$DOF = 2z_R = 2 \frac{\bar{\lambda}}{\pi \sigma^2} = 2 \frac{\bar{\lambda} n}{NA^2}, \quad (1.23)$$

where, n = refractive index. Hence, the increase of the NA value decreases both DOF and transversal resolution quadratically. *"For example, a transversal resolution of $\delta r = 20 \mu m$ at a mean wavelength of $\bar{\lambda} = 830 nm$ leads to a depth of focus of $DOF = 3 mm$, whereas a transversal resolution of $\delta r = 2 \mu m$ reduces the depth of focus to $DOF = 30 \mu m$."*[3]. A compromise between the transversal resolution and DOF has to be found.

1.3.4 Penetration depth

The scattering behaviour and the absorption properties are significantly influenced by the morphology of the tissue [5]. The penetration depth is strongly dependent on the light scattering behaviour of the tissue, which is scaled as $1/(\lambda_0^k)$. The value of k depends on several parameters as the scattering size, shape and refractive index.

Due to the presence of hemoglobin and melanin, most biological tissues show high scattering and absorption in the near-IR wavelength range. Between 0.8 and 1.8 μm the imaging penetration depth is limited by scattering and water absorption.

Thus, in the OCT imaging field the central wavelength is usually choose between 500-1500 nm. For dermatological applications, a good compromise between axial resolution and depth penetration has been achieved using wavelengths in the range between 1300-1500 nm. In this range, penetration depths between 1 to 2 mm can be achieved

1.4 OCT techniques

Two main big families of OCT techniques can be found over the literature. They differ from each other by the way they scan the sample.

1.4.1 Time-domain OCT (TD-OCT)

Two scans must be performed: one lateral OCT scan, and a depth-scan. The first one gives information about adjacent sample positions, the second one detect the depth positions of the sample light re-emitting sites using LCI.

The echoes from different depths of the target tissue are acquired by moving the reference mirror between a “zero depth” position and a “maximum depth” position.

As stated, when a low coherence light beam is used, interference signal is obtained when the path length of the backscattered light is equal in both reference and sample arm. As the sample information is included in the interference term, this technique is based on the convolution operation expressed in Eq.1.3.

While depth-scans are acquired moving the reference mirror, lateral-scans are acquired moving the light source.

Because two scans are needed to obtain two or three-dimensional images, this technique is the most time-consuming. However, it is also the one with the highest potential to be adapted to different various further developments (Spectroscopic-OCT, Doppler-OCT, etc. . .).

1.4.2 Fourier-domain OCT (FD-OCT)

This technique has been developed to overcome the problem of the TD-OCT where two scans are needed. FD-OCT needs only the lateral scan because the depth-scan information is obtained through an inverse FT of the backscattered light spectrum. Thus, it necessary the use in combination of a broadband light source and a spectrometer, which usually consists of a grating and a line scan camera. From the innovation brought by the FD-OCT technique, two very used techniques have been developed:

- *Spectral Domain OCT (SD-OCT):*

The interference spectrum is thus spatially encoded on the different detector pixels. This technique is based on the Eq.1.25 [3]:

$$F_S(z) \propto FT\{A_S(K)\} \quad (1.24)$$

where, $A_S(K)$ = spectrum of the backscattered sample light amplitude. The spectrum is obtained using a spectrometer instead of a photodetector. The spectrometer is composed by a grating which separates the light in different wavelengths. Then, the light interference is detected depending on its wavelength by a linear array of photo-sensitive elements (usually CMOS or CCD). As stated by Eq.1.25, the scattering potential $F_S(z)$ is obtained by FT the field amplitude of the sample beam $A_S(K)$, and it is a function of depth. Looking at the interferometer exit, the spectral intensity will be [3]:

$$I_{SR}(K) = I_S(K) + I_R(K) + 2\sqrt{I_S(K)I_R(K)}Re\{\mu(K)e^{i[\Phi_S(k)-\Phi_R(k)]}\}. \quad (1.25)$$

$I_S(K) = |A_S(K)|^2$ = sample beam's spectral intensity (or power spectrum), $A_S(K) = a_S(K)e^{i\Phi_S(K)}$ = $A_S(K)$'s complex amplitude, $I_R(K)$ = reference beam power spectrum, $\mu(K)$ = spectral degree of coherence (=1 in case of

monochromatic light), $\Phi_S(k)$ = sample beam's spectral phase, and $\Phi_R(k)$ = reference beam's spectral phase. Depth resolution is spectral width dependent in both TD-OCT and FD-OCT, but, looking at Eq.1.25, the FT operation needs the spectrum samples are in k-space. Thus, a re-sampling operation is needed.

As stated, depth-scan is not needed in SD-OCT, that means higher speed than TD-OCT, and higher speed means less motion artifacts. Therefore, sensitivity advantages over TD-OCT are offered [6].

High financial cost (expensive detector arrays), small dynamic range and high roll-off (intensity signal loss depending on distance from the zero delay) are the two main disadvantages of SD-OCT. There are some specific disadvantages depending on the technique used to obtain the real amplitude $a_S(K)$. The correlation technique, for example, needs twice the width of the object structure to represent the auto-correlation of the object structure wasting a considerable part of the bandwidth of the signal processing system. This problem can be avoided using a phase shifting LCI technique [3].

- *Swept Source OCT (SS-OCT):*

In this technique the spectrum is encoded in time. Differences can be found in the hardware:

- a The spectrometer is replaced by a single photo detector or a dual balanced detector (DBD) at the exit of the interferometer;
- b The light source used is a wavelength-tunable swept laser. It works by emitting every single wavelength of a wide selected range of wavelengths.

The light source scans its wavelength, and the photo detector records all wavelengths over time. Three main advantages over SD-OCT techniques are offered by SS-OCTs: faster scanning rate, higher spectral resolution, no high roll-off. Nowadays, swept source sweep rates crossed the barrier of the MHz enabling depth scan rates of tens of million axial scans per second [7].

In general, SS-OCT systems are the most used in dermatology and, because of their high scanning rate, the system of this work use this technology.

1.5 OCT system

The OCT system used for all the acquisitions necessary to this work was provided by the *Center for Medical Physics and Biomedical Engineering of the Medical University of Vienna*. Fig.1.5 shows the scheme of the concerning OCT system. Six main components can be identified: the laser source, the reference arm, the sample

arm, the beam splitter, the dual-balance-detector and the workstation. The main parameters of the OCT system are reported in Tab.1.1.

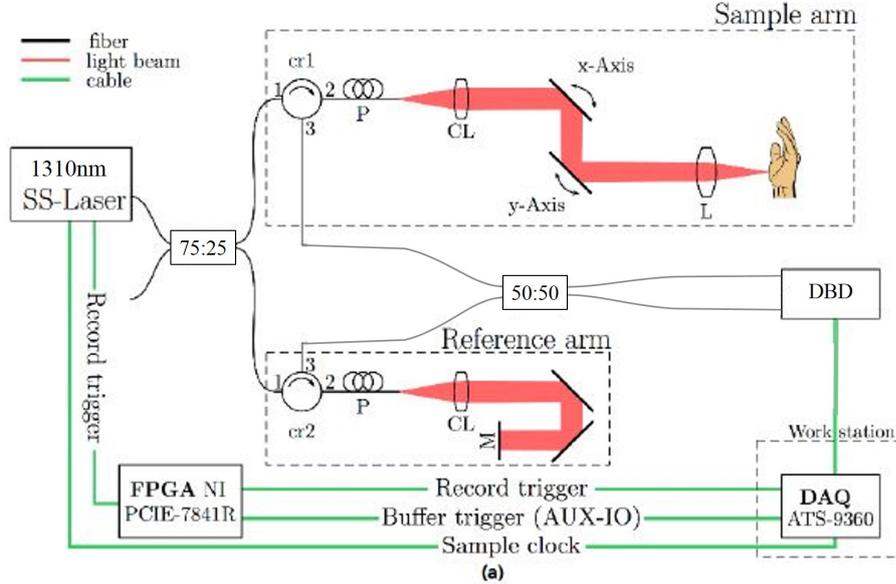


Figure 1.5: OCT system imaging setup used during the acquisition sessions [adopted from [8]]

<i>Central wavelength</i>	1310nm
<i>Bandwidth</i>	29nm
<i>Swept frequency</i>	400kHz
<i>Source power</i>	10mW
<i>Duty Cycle</i>	100%
<i>Coupler ratio</i>	75:25
<i>Axial pixel size</i>	13.7 μ m
<i>Lateral pixel size</i>	9.8 μ m
<i>SNR</i>	105.3dB

Table 1.1: System main parameters

The source power is not the same the patient feels. While covering the path, the beam will dissipate a part of his power and only around 15mW will be provided to the patient skin.

As stated in Sec.1.3.2, the light source is a swept source (SS-OCT 1310, Insight Photonic Solutions, Inc, U.S.). There is no waiting between one sweeping process and the next one so the laser scan the sample continuously (DC = 100%). Because a significant amount of time is spent to jump from one sample point to the next one, and this is a problem, two different solutions can be adopted. The first one is to use the a time interval equal to the bandwidth end frequency to jump to the next sample point. The second one is to use a sweep interval to jump to the next sample point. Because of the short bandwidth, the second solution has been adopted in order to not waist any spectrum point.

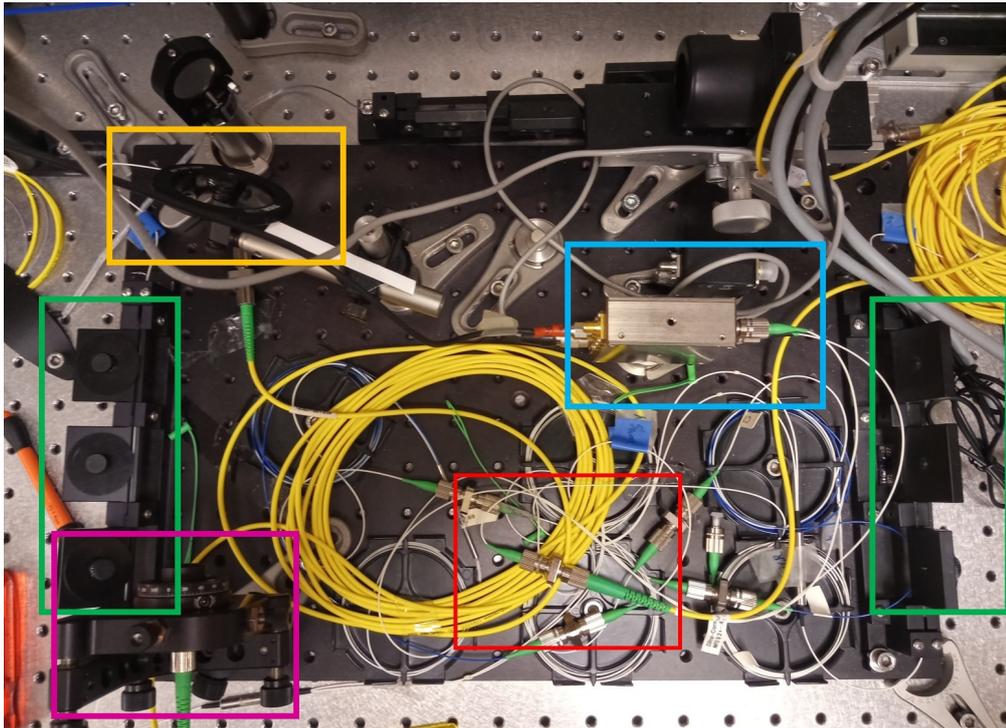


Figure 1.6: Reference arm top view

The coupler ratio indicates that the beam splitter addresses 75% of the beam power to the sample arm and the remaining 25% to the reference arm.

Lateral pixelsize can be calculated in air using a special piece of glass as a sample. The sample had 1mm dot-target on its surface. Since it shows a rectangular pattern with 1mm distance between the dots, it is possible to measure the lateral pixel size on the image.

The axial pixel size can be measured following the same process, but using a different sample. The sample used is a 1mm thick glass.

A top picture of the OCT system is shown in Fig.1.6. The light is carried

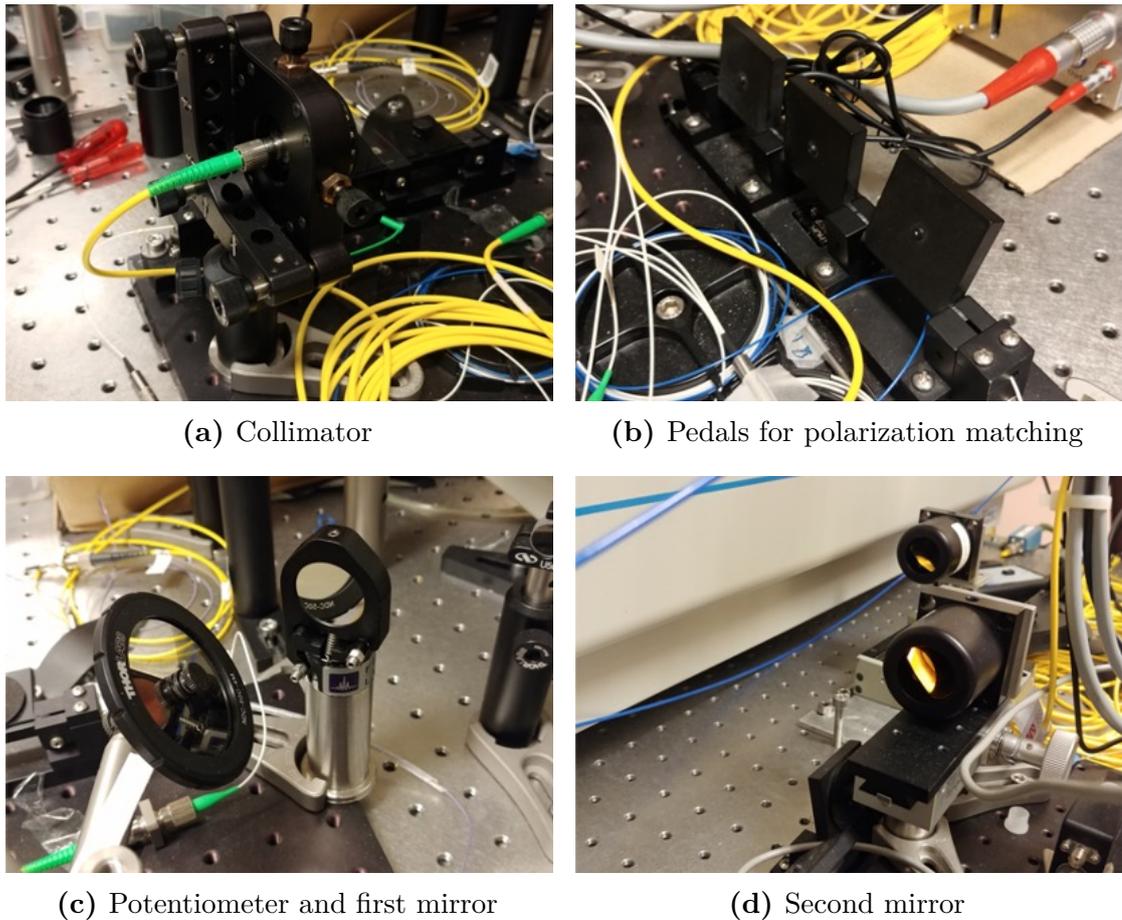


Figure 1.7: Reference arm

through the system by the yellow fiber until it is splitted in reference and sample by the beam splitter (red square).

The reference beam enters the reference arm which is composed by a collimator (purple square, detail in Fig.1.7a), three pedal for polarization matching (green square, detail in Fig.1.7b), a potentiometer (yellow square, detail in Fig.1.7c) and 2 mirrors (Fig.1.7c and Fig.1.7d).

The sample arm (Fig.1.8) is composed by the galvanometer scanner, the scan lens and the three pedal (Fig.1.6). The galvanometer uses a combination of moving mirrors to scan automatically in both lateral directions. The inclination of the scan lens can be changed manually to reduce the side-lobe artifacts which can worsen the image quality.

Fig.1.9 shows an overview of the power supplies and the laser source.

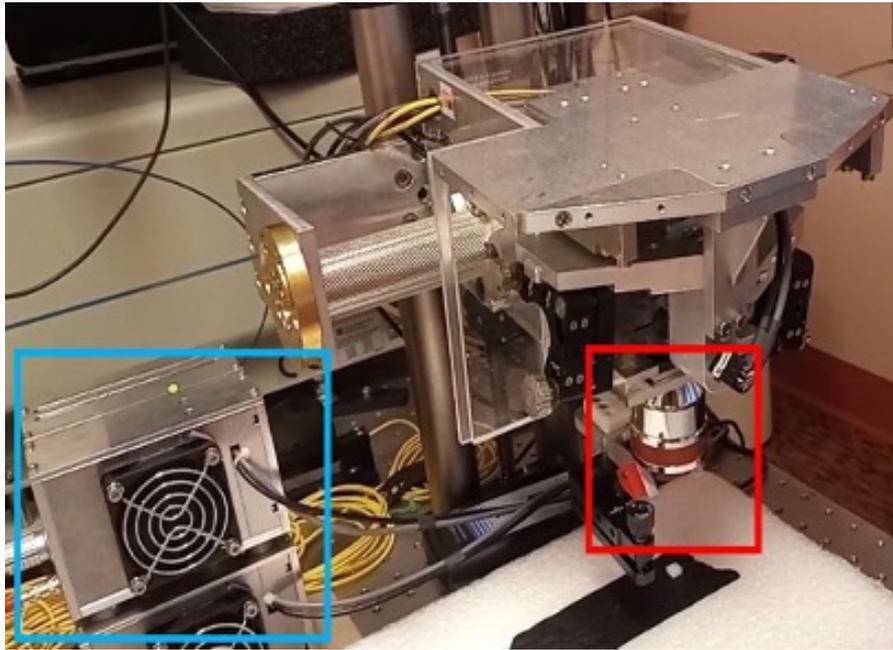


Figure 1.8: Sample arm

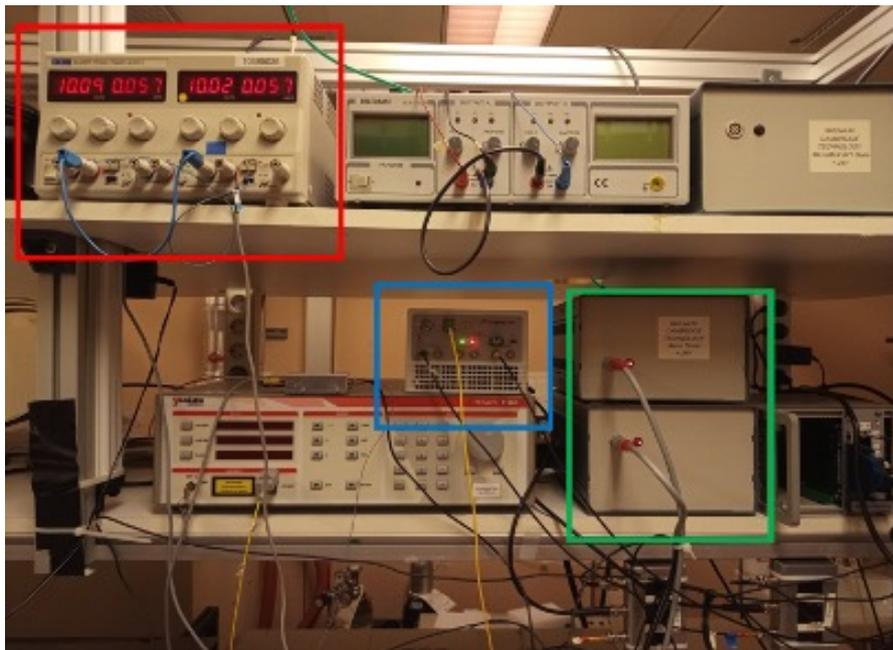


Figure 1.9: Power supplies and laser source

Chapter 2

Image reconstruction

This chapter proposes an algorithm for high speed OCT volume reconstruction. All the proposed steps were designed relying on the evidences found in literature.

The algorithm was developed in Python using the GPU offered by Google Colaboratory (Colab). Colab is an online platform powered by Google which allows to write and compile Jupyter notebooks without any configuration. All the files used were stored in a Google Drive account which can easily communicate with the Colab platform. In order to use data on Google Colab, each volume acquired by the OCT system (1280 axial pixels x 1024 lateral A-scans x 1024 adjacent B-scans) were to be uploaded in a Google Drive folder. Colab offers two GPU models: Nvidia Tesla T4 [1.59GHz memory clock, 16Gbytes memory, 8.1 TFLOPS performance] and Nvidia K80 [0.82GHz memory clock, 12Gbytes memory, 4.1TFLOPS performance]. The GPU model depended on its availability.

Nvidia's compute unified device architecture (CUDA) could only be programmed in C++ programming language. However, it is possible to exploit the GPU acceleration by using simple and user-friendly Python's libraries as CuPy. "*CuPy is an open-source array library for GPU-accelerated computing with Python. CuPy utilizes CUDA Toolkit libraries including cuBLAS, cuRAND, cuSOLVER, cuSPARSE, cuFFT, cuDNN and NCCL to make full use of the GPU architecture. CuPy's interface is highly compatible with NumPy and SciPy; in most cases it can be used as a drop-in replacement.*" (<https://cupy.dev>).

2.1 Volume reconstruction algorithm

Different reconstruction algorithms were found in literature. A basic GPU accelerated algorithm for OCT volume reconstruction was designed by Watanabe and Itagaki [9], the flowchart is shown in Fig.2.1. Other algorithms showed principally different denoising techniques as windowing [10], or different way to perform the FT [11].

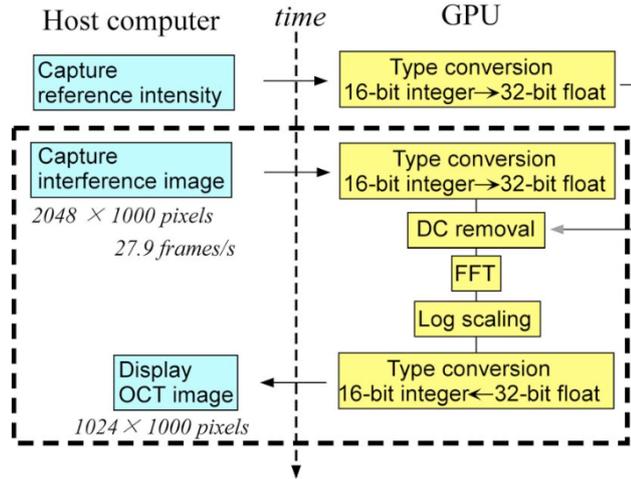


Figure 2.1: Basic OCT reconstruction algorithm [adopted from [9]]

Considering the Watanabe's algorithm, it was possible to define the four essential operations for GPU-accelerated processing of interference images.

1. *16bit integer conversion to 32bit float:* interference image were saved as 16-bit unsigned integer in binary files. In order to prepare the data for the FFT, they must be converted to 32-bit float;
2. *DC removal:* this is the elementary denoising operation to reduce autocorrelation artifacts. It consists in a simple subtraction of the background signal from the interference image. The background signal is produced by only the reference beam and can be obtained by blocking the sample arm. As shown in [9], this procedure is performed externally to the algorithm. The background signal is then uploaded on the GPU memory in order to be subtracted from the interference images;
3. *FFT:* this is the most important step. As explained in Sec.1.2, the interferogram can be obtained by Fourier transforming the spectrum of the backscattered sample light amplitude;

4. 32 bit float conversion to 16bit integer: this step returns the data into a format which can be displayed.

Normally, after the conversion from 16bit integer to 32bit float, a resampling operation must be performed. The data have to be moved from a λ -space to a wave number-space (k -space). This step is necessary since there is nonlinear relationship between λ and k . Van der Jeught et al.[12] offer a comparison between the most popular resampling procedures: nearest-neighbor, linear, and cubic-spline interpolation.

This operation can be really time consuming. It can be avoided using a k -linear spectrometer combined with a diffractive grating and a equilateral prism [9] or using the solution adopted in the OCT system adopted for this work. It uses a combination of a uniform in time and frequency user sample clock (or electronic k -clock) signal with a start-sweep (A-line trigger) signal. With this solution, not only the resampling operation is not even necessary, but also the hardware is significantly simplified [8].

The reconstruction algorithm presented in this work is shown in Fig.2.2.

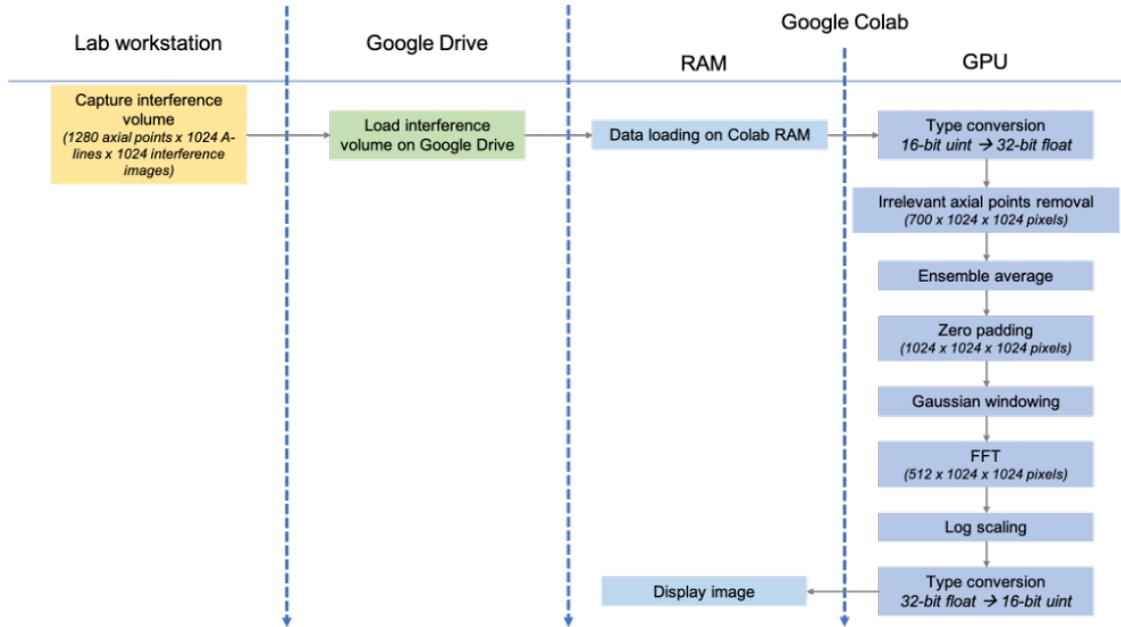


Figure 2.2: Volume reconstruction algorithm

The interference image was acquired and saved by the OCT system as a binary file. Then, it was uploaded on Google Drive in order to be read and processed on Colab. The OCT system saved the spectra of the interference signal as a row

vector of 1342177280 32-bit unsigned integers ($1280 \times 1024 \times 1024$). The data were to be resized to the known dimensions of the final volume. The resizing operation transformed the vector into a known size three-dimensional *numpy* array. Only 700 pixels of each axial scan were relevant. The pixel positions which had to be taken in account were recorded during the automatic calibration of the system and then stored in a ".txt" file. The input array size was 700 axial pixels \times 1024 lateral axial-scans \times 1024 adjacent interference images.

Because of RAM limitations, it was not possible to upload more than one interference image at time on Colab. Thus, instead of performing the DC removal, an ensemble average was performed [13]. The averaged interference image was calculated and then deducted from every interference image in the volume.

Afterward, a zero padding operation was performed for two reason: increase the resolution and exploit as many GPU threads as possible. Normally the number of threads is a power of two. Thus, the axial pixels were padded to the closest power of two bigger than the actual A-line length (1024 pixels).

Because of its small bandwidth, to do not decrease the resolving power, a non-Gaussian spectrum light source was used (Fig.2.3).

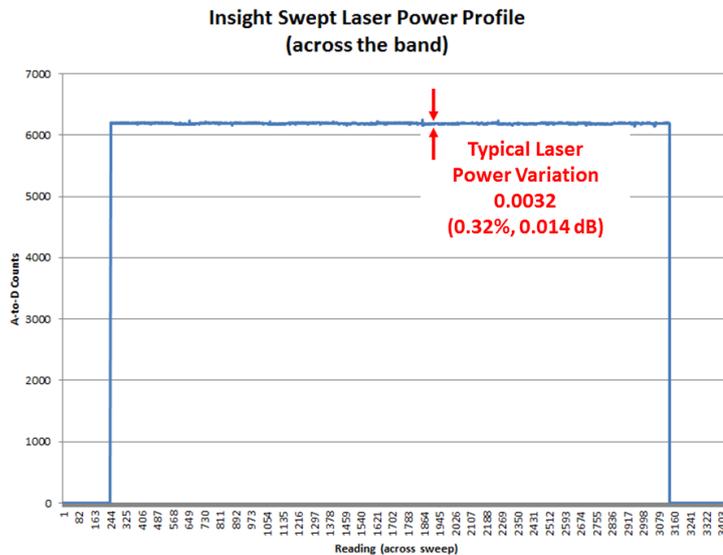


Figure 2.3: Light source spectrum profile adopted in the OCT system [adopted from <https://www.sweptlaser.com/flexible-stable-power>]

Sidelobe artifacts are a natural byproduct of non-Gaussian spectrum light sources

[14]. Since the presence of a needle increases the noise related to that typology of artifact, a Gaussian window was applied to every A-line of the interference images (Eq.2.1).

$$w = e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad (2.1)$$

where $\mu = 512$, and $\sigma = 315$.

The needle dimension is several order of magnitude bigger than the axial and the lateral resolution. Thus, a small decrease in the optical resolution will not affect significantly the efficiency of the tracing algorithm. A comparison between non-windowed and windowed B-scan is shown in Fig.2.4.

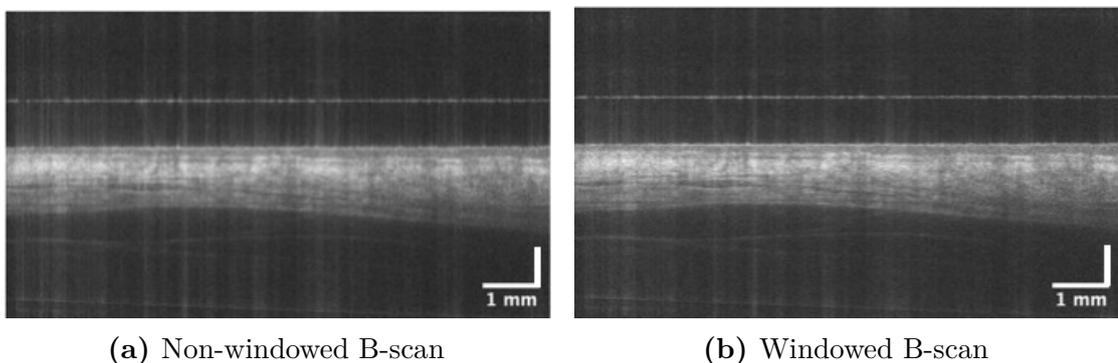


Figure 2.4: Comparison between non-windowed (a) and windowed (b) B-scan. The two figures represent the same B-scan of an OCT volume before and after the signal windowing.

The FFT was computed on a number of points equal to the A-line length. All the complex pixel values were taken back to real space by taking their absolute value. Then, the spectral replica was removed reducing the A-line length to 512 pixel.

FT produces symmetrical images around the 0-delay line, that's why the spectral replica needs to be removed. Because of the Fourier transformation, FD-OCT systems are characterized by a sensitivity roll-off due to limited spectrometer resolution [15]. The greater the distance from the 0-delay line, the lower the sensitivity and the finer the interference signals are. Moving closer to the 0-delay line increases the possibility that the visible half of the symmetric image is crossed by the other half. When this happens, mirror artifacts can be seen in the image (Fig.2.5). In this work, eventual mirror artifacts were removed by cropping 64 B-scans from each side of the volume. The final volume size was 512 axial pixel \times 896 lateral A-scans \times 896 adjacent B-scans.

In OCT, logarithmic scaling operation is used to exploit the whole image dynamic. In other words, it works almost as an image equalization. After the logarithmic

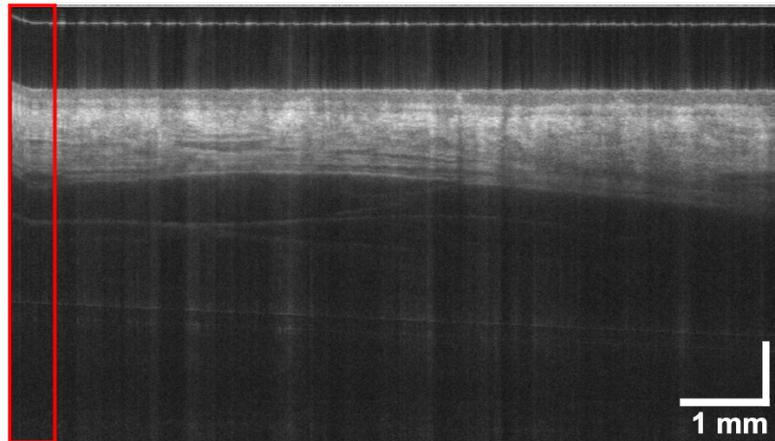


Figure 2.5: Mirror artifact (red box).

scaling, it was necessary to add a 1 to all the pixel values or the 0-value pixel would be resulted equal to '-inf' after the logarithm calculation.

Before typing back the data format from 32-bit float to 16-bit integers, a mix-max scaling was performed in order to obtain pixel values between 0 and 1.

Fig.2.6 shows an acquired volume from the forearm skin.

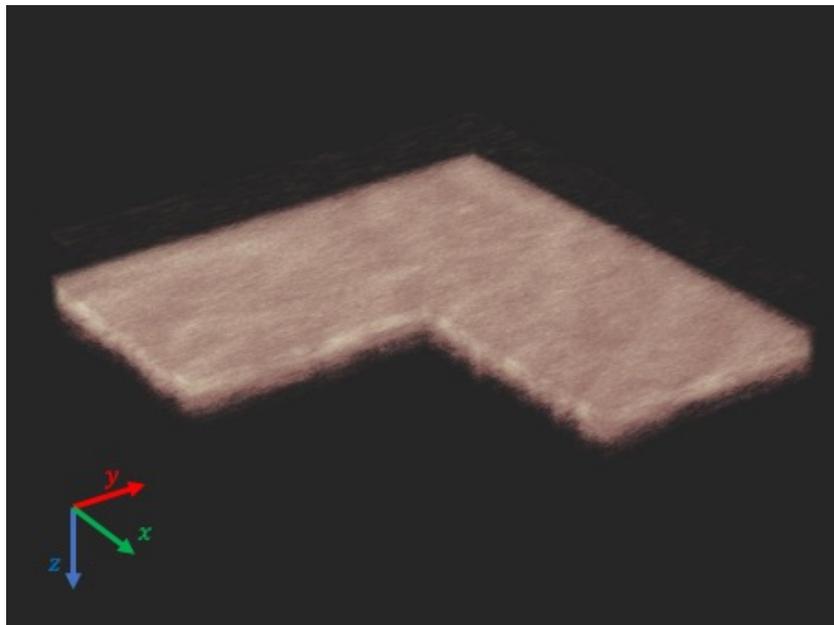


Figure 2.6: Example of an acquired and processed OCT volume from the forearm skin. The volume was displayed on VolumeViewer (Matlab Image Processing Tool).

2.2 Volume registration

Another way to decrease the sidelobe artifacts was to incline the scan lens. The inclination was translated as a rotation of the acquired volume. This rotation prevented a correct En-face visualization. Thus, a volume registration was necessary.

As showed in Fig.2.7, only one between the B-scan and the cross-section frames are tilted. When the B-scans are straight but not aligned, the cross-section are tilted but aligned, and vice versa.

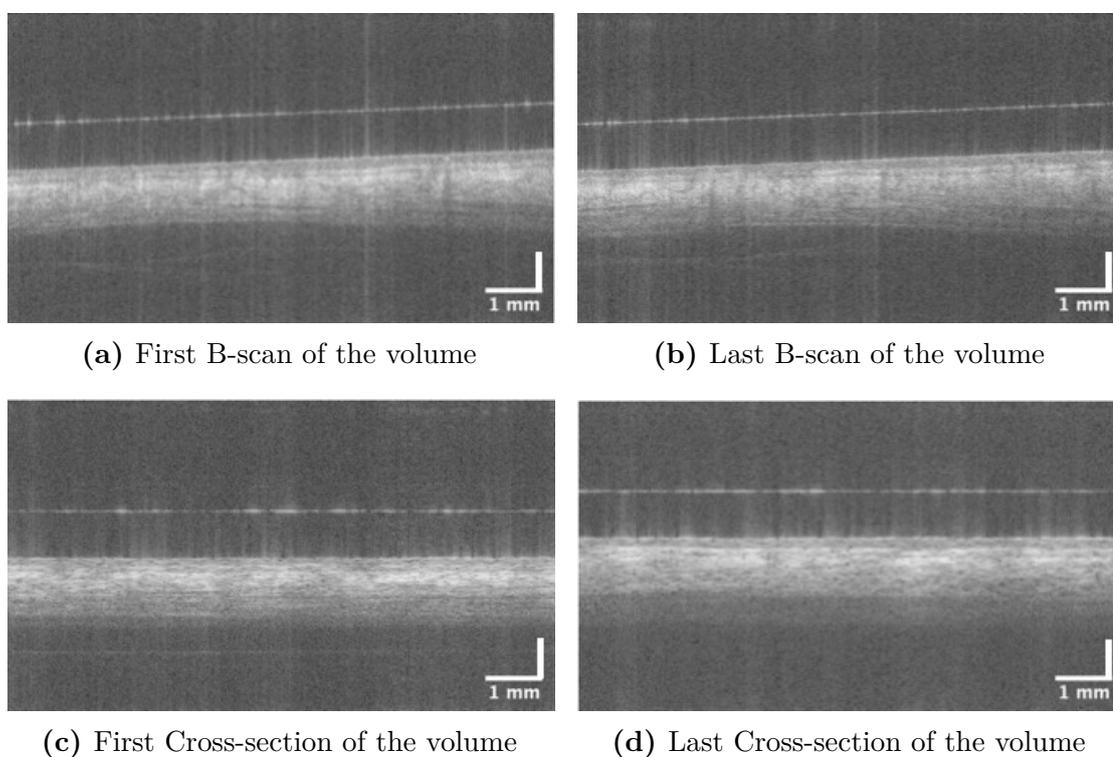


Figure 2.7: Volume inclination: (c) and (d) are tilted but aligned, (a) and (b) are straight but not aligned. The inclined spatial plane was know a priori by the way the scan lens was tilted.

The first inclined frame of the volume was considered to measure the inclination of the volume. The registration was carried out following 5 steps:

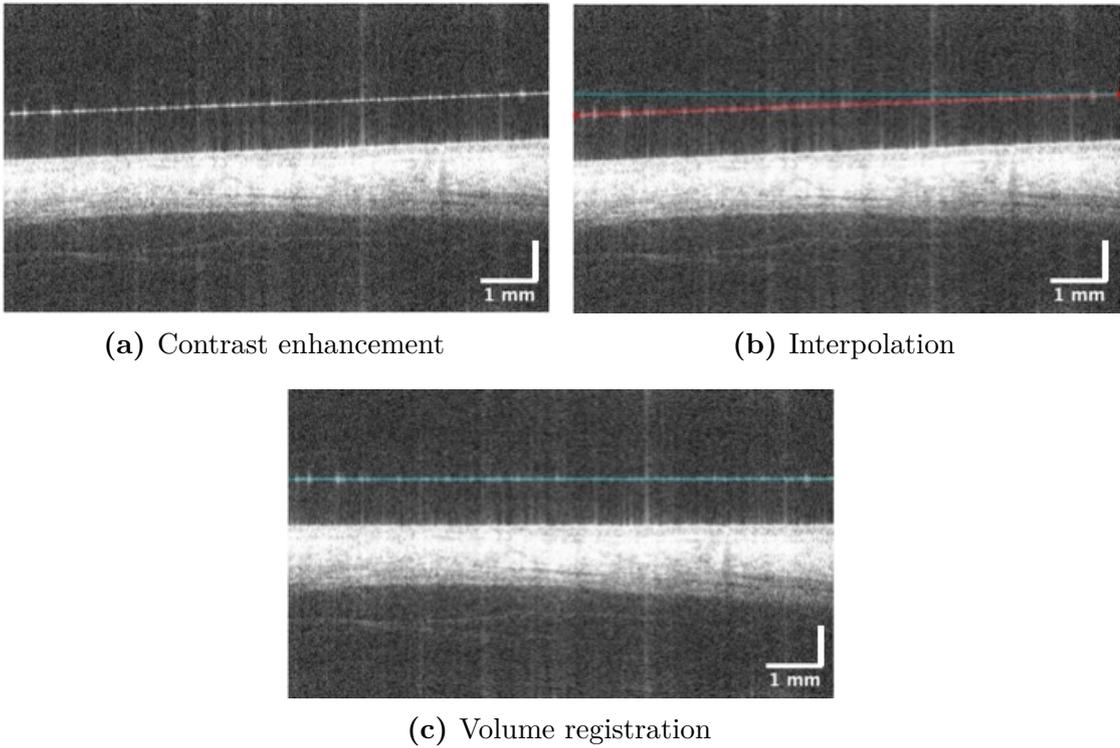


Figure 2.8: Volume registration steps.

1. *Brightness reduction*: the brightness of the selected frame was reduced in order to do not saturate the image during the subsequent contrast enhancement;
2. *Contrast enhancement* (Fig.2.8a): a simple contrast enhancement was performed by increasing every pixel deviation from the luminance (average pixel intensity) by the 30% of it. Then all the pixel with intensities greater than 1 or smaller than 0 were equalled respectively to 1 and 0;
3. *Interpolation* (Fig.2.8b): the first pixel with intensity greater than 0.8 was searched in the first and the last A-line of the frame. The line equation passing through the two points was calculated.
4. *Variance from horizontal line calculation*: for each point of the interpolated line, its distance from the horizontal line was measured. The horizontal line was made to pass through the highest y-coordinate point between the two founded during interpolation;
5. *Registration* (Fig.2.8c): in the previous step was measured the distance between each B-scan (or Cross-section) and the horizontal plane. The registration was performed by shifting each of these frame by its related distance.

2.3 Timings

In order to exploit the GPU-acceleration and reconstruct the volumes real-time, it was necessary to allocate the GPU memory. The memory allocation was the most time consuming operation and several seconds were necessary to run the algorithm every time the Colab runtime was restarted. Thus, the algorithm was compiled twice. The first for the memory allocation and the second one to appreciate the GPU acceleration. Once the GPU memory was allocated, it was possible to delete the volume and processed it as many time it was needed.

The algorithm processing timings were measured over 50 acquisitions. The results were collected on an Excel file and displayed on bar diagrams as shown in Fig.2.9,2.10,2.11. The GPU model used during the timing was a Nvidia Tesla K80.

The reconstruction algorithm showed stable and fast processing time all over the records. The average processing time, considering the memory allocation on the GPU, was $954,317ms$. Memory allocation required an average interval of $418,612ms$. Thus, the processing lasted an average of $535,705ms$. While the reconstruction algorithm showed stable timings, the registration ones were not stable as well. The highest recorded registration time was $637,643ms$ and the lowest one was $133,394ms$. The most time consuming operation was the frame rolling. It was observed that if the cross-sections had to be rolled, the average registration time was around 0.6s. As stated, which frames type (B-scans or cross-section) had to be rolled could be controlled by rotating the lens scan. Thus, this drawback could be avoided by adjusting the lens scan orientation.

In the worst situation, all the volume A-lines (1048576) were processed in $1.637s$. In other ways, more than 640×10^3 A-lines were processed every second, that was more than 420 B-scans/s. The fastest recorded processing time was $1.022s$. That meant, more than 10^6 A-lines and more than 10^3 B-scans processed each second.



Figure 2.9: Total processing timings for the volume reconstruction.

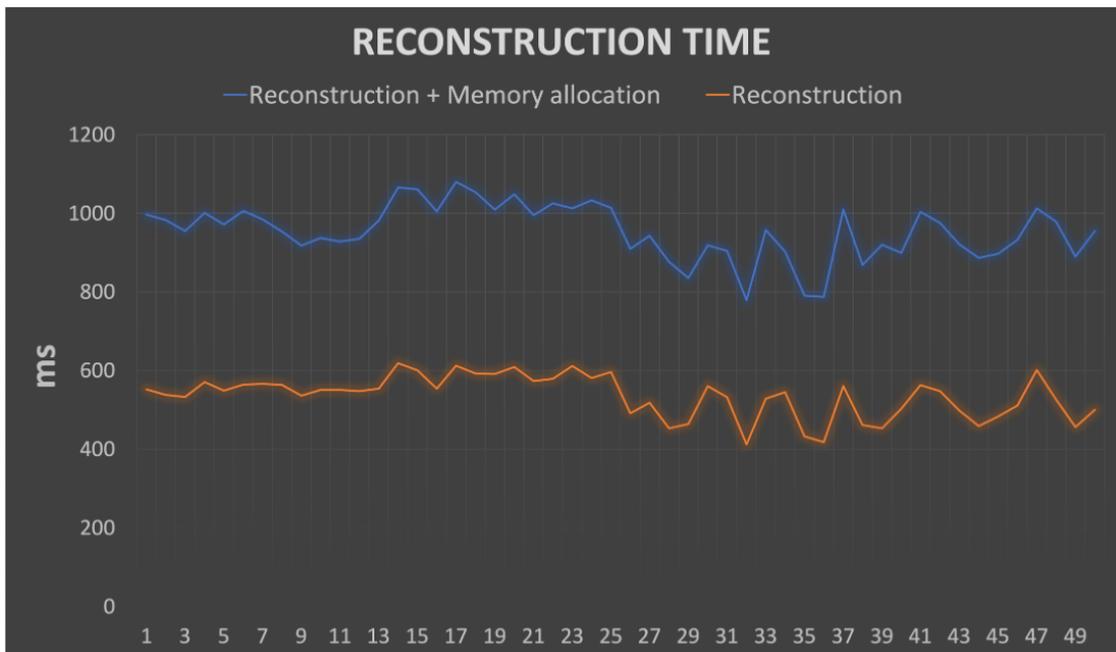


Figure 2.10: Reconstruction algorithm timings

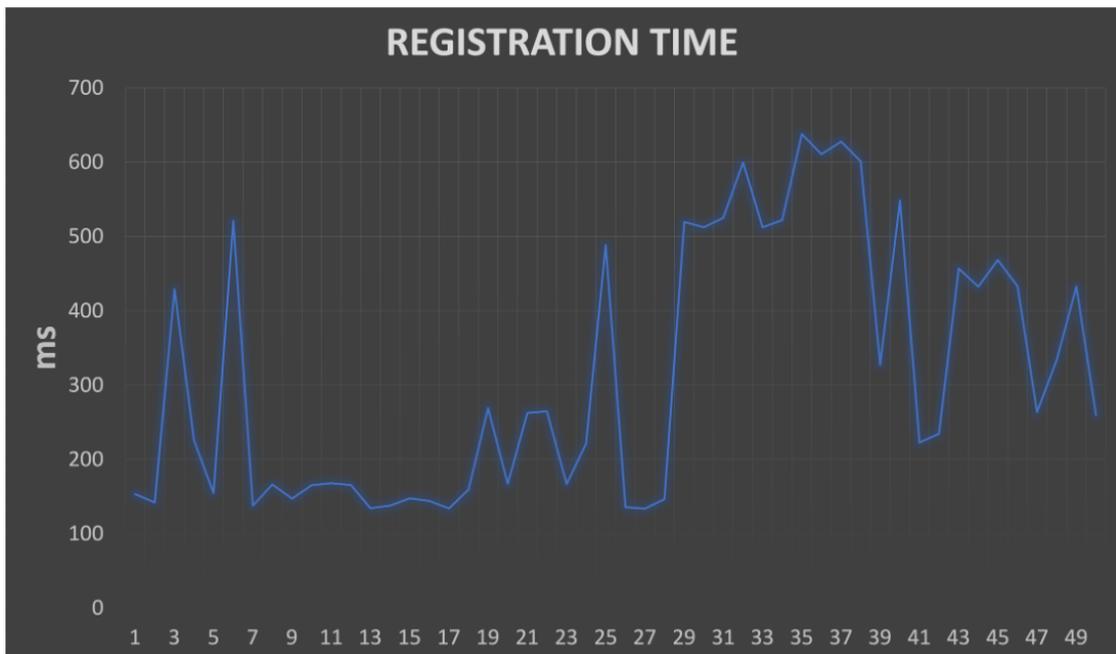


Figure 2.11: Registration timings

Chapter 3

Tissue-mimicking phantom

For the development of biomedical instruments and the validation of new methods, technologies and applications, tissue-mimicking phantoms are vitally important.

In biomedical applications as OCT, the main challenge is to design durable phantoms which can replicate optical properties and mechanical behaviour of the biological tissues [16].

3.1 Materials

Different materials can be used for this purpose. The most common ones are silicone, fibrin, poly(vinyl alcohol) cryogels (PVA-C) [17], polyurethane [18] and polydimethylsiloxane (PDMS) [19].

For this application, the mechanical properties, the ease of preparation, and the stability of the material three-dimensional structures, played an important role for the choice of the material.

Since the needle had to be inserted repeatedly in the phantom, the material had to be tough enough to do not overly ruin the three-dimensional structure.

As wavelength increases, optical properties like scattering and absorption decrease nearly monotonically [3]. No evidences were found about the behaviour of the listed materials in the wavelength range used in this work (central wavelength = 1310nm, bandwidth = 30nm). Thus, it was necessary to choose a material with highly controllable optical properties.

The choice felt to the silicone. It is tough, highly compatible with a wide range of scattering and absorbing agents, and its mechanical properties are wide range controllable. In addition, the refractive index of the silicone matrix (1.4) is close to that of tissues (1.35 ÷ 1.55) [20].

Because of the incompatibility of the material with organic materials, titanium

dioxide (TiO_2) and black pigment were used in order to increase respectively the scattering and the absorption of the phantom [21].

The main challenge of using silicone was to achieve an homogeneous distribution, without scattering and absorbing agent aggregates, sedimentation, and air bubbles. To minimize these problems, solutions as sonication [22], and degassing under vacuum [23] were adopted.

The silicone used for the fabrication was the Dragon Skin™ 30, 1:1 mixing ratio, hardness 30A, cure time 16h [<https://www.smooth-on.com/product-line/dragon-skin/>]. The black pigment used was the Silc Pig™ [<https://www.smooth-on.com/products/silc-pig/>].

3.2 Design

As stated in Sec.3.1, no characterization of the optical properties for silicone phantom at 1310nm wavelength were found in literature. Different TiO_2 and black pigment concentrations were proposed by Fu et al.[21], whom compared absorption and scattering coefficients at 633, 750 and 900nm. Starting from these concentration values, lot of trials were conducted before an acceptable compromise between the profile of the average A-scan of a forearm skin volume and the one acquired from the phantom was achieved. The goal was to replicate as much as possible the B-scan showed in Fig.2.4b.

The equipment required for the fabrication was:

- 60mL silicone matrix;
- 60mL silicone catalyst;
- TiO_2 ;
- black pigment
- 3x 50mL Falcon™ conical tubes;
- 3x 50mL syringes;
- 1x Petri dish;
- Vacuum chamber;
- 1x brush;
- black plastic ball head pins.

The fabrication process started with the preparation of three mixtures of the catalyst with different concentrations of TiO_2 and black pigment. The mixtures were then placed in three different Falcon™ conical tubes.

1. *First mixture*: 8mg/mL of TiO_2 and 0.2% in weight of black pigment were mixed in 20mL of catalyst (less viscous than the silicone matrix);
2. *Second mixture*: a concentration of 0.2% in weight of black pigment was mixed in 20mL of catalyst;
3. *Third mixture*: a concentration of 1% in weight was mixed in 20mL of catalyst.

The concentrations were calculated considering the later addition of the silicone matrix.

The mixtures were placed in ultrasonic bath for about 2h [22] at 40 Celsius degrees. Sonication is more effective on less viscous fluids [24], and, since this operation requires quite a long time, if both the silicone matrix and the catalyst were mixed, the curing process would begin during the bath. Consequently, a good degassing of air bubbles would have been compromised. The sonication process broke most of the aggregates. The bigger ones were let to settle at the bottom of the Falcon™ for 24h.

The silicone matrix was added to the mixtures in a 1:1 ratio. Since the mixture volume was doubled, the concentrations of TiO_2 and black pigment were halved.

Then, the mixtures were placed in syringes in order to better control their distribution over the Petri dish.

The first mixture was deposited at the bottom of the Petri dish, in such an "side-down top-to-bottom" fabrication technique. Approximately 1.5mm thick layer of the first mixture was poured over the Petri dish. This layer was designed to simulate the epidermis. In order to recreate a similar inhomogeneity as in the biological tissues, a very thin layer of the second mixture was poured over first one.

Plastic ball head pins were carefully inserted in the silicone. The reason pins were used is explained in Sec.3.2.1. The pins were blocked to the Petri dish to restrict the slight movements that could happen during the degassing process.

The Petri dish was then placed in a vacuum chamber until the silicone is completely degassed. During this step, the air bubbles went bottom to top of the Petri dish.

A thick layer of the third mixture was then applied. This layer was designed to simulate the high tissue absorption at approximately 2mm of depth (Fig.3.1b).

After the silicone completely cured (24h), the phantom was removed from the Petri dish, turned upside down, and brushed with a really thin layer of transparent silicone. The goal was to recreate on the phantom top surface a stratum corneum-like layer, the first and easily discernible skin layer in OCT skin images [25].

The final result is shown in Fig.3.3. The presented fabrication required more than 48h to complete all the steps needed. However, using 40mL of silicone for each mixture, four phantoms could be produced.

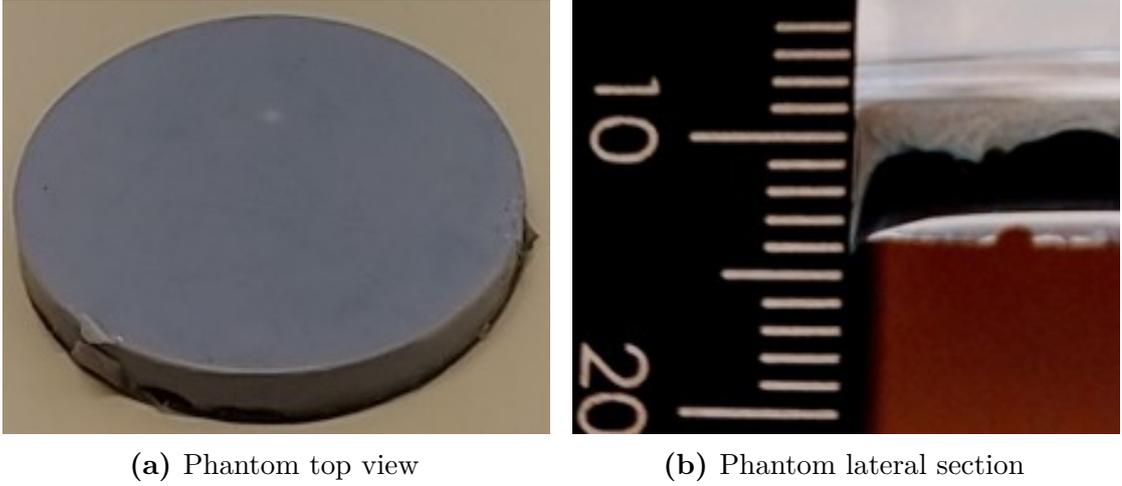


Figure 3.1: Pictures of the fabricated phantom: (a) Phantom top view, (b) lateral view of a section of the phantom.

3.2.1 Biopsy area-mimicking plastic inclusion

In OCT images, non-melanoma skin cancer (NMSC)[26] and malignant melanoma (MM)[27] were widely investigated. The characteristic layering proper of normal skin is lost in NMSC, MM and in various benign lesion [28]. While MM shape is characterized by large vertical, icicle-shape structures, a deeper investigation of other NMSC characteristics is needed to avoid confusion with benign lesions. For example, focal changes of thickening of epidermis or dark rounded areas, sometimes surrounded by a white structure can be found respectively in actinic keratosis (AK) and in basal cell carcinoma (BCC)[29]. Despite its invasivity, skin biopsy is still one of the reference standard for skin diseases diagnosis [30].

One of the task of this work was to trace real-time the distance of the needle tip from the area where to extract the biopsy.

As showed in Sec.3.2, the insertion of a plastic ball head pins was intended to circumscribe a region within the phantom. This area was used to simulate the tissue region to collect by biopsy.

Is important to remember that the solution proposed was not intended for the simulation of a skin lesion. Skin lesions have so many different features and huge variety of shapes to be replicated that specific studies should be conducted for the realization of a skin lesion-mimicking inclusion.

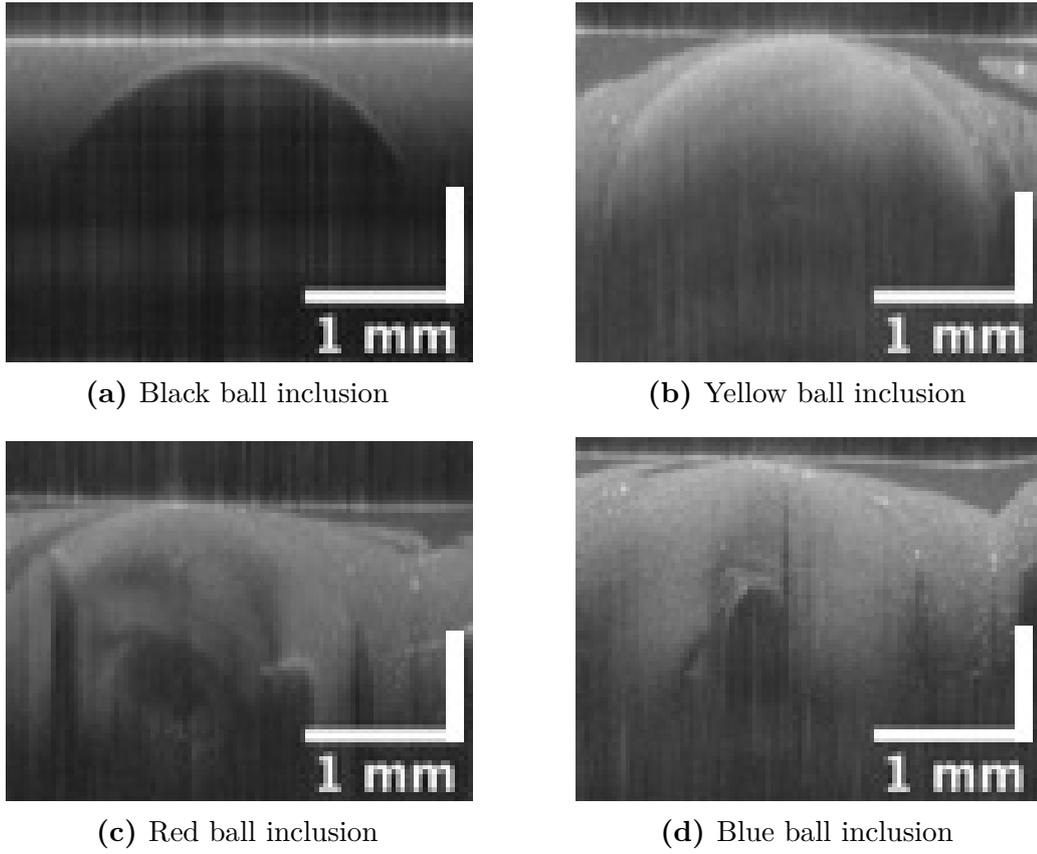


Figure 3.2: Comparison of different colored inclusion: (a) black ball inclusion, (b) yellow ball inclusion, (c) red ball inclusion, (d) blue ball inclusion. All the phantoms were imaged with same OCT system and laser parameters.

Once inserted in phantoms, different colored plastic balls were examined under OCT (Fig.3.2). It was supposed that black plastic ball should be easier to segment by the network. In addition, it is the most similar to a BCC because of its dark rounded shape (Fig.3.2a). Therefore, black plastic ball inclusions were used for the task.

3.3 Phantom VS reality

The average A-scan was calculated over the whole phantom volume. In order to evaluate its similarity with the skin tissue, it was compared with the average A-scan of the skin volume taken as reference (Fig.3.3b).

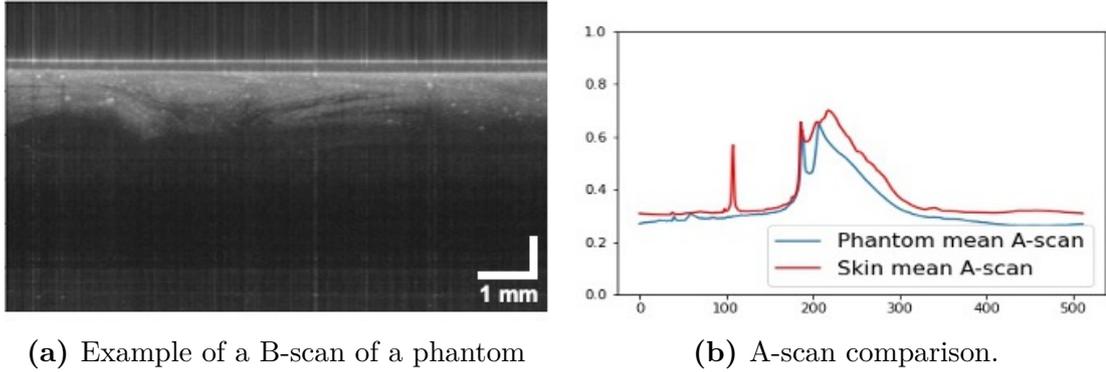


Figure 3.3: B-scan comparison between phantom and skin: (a) this B-scan belonged to one of the fabricated phantoms, its cosine similarity score was 99.74%. (b) Comparison between the average A-scan of the phantom and the reference one. The spike on the left of the red profile corresponded to the backscattered light produced by a thin mirror positioned over the skin. The mirror was necessary to reduce the sidelobe artifacts.

The cosine similarity between the reference average A-scan and the average A-scan of each phantom was measured. Phantoms with a score lower than 70% were not considered. The phantoms similarity was $91.48\% \pm 8.26\%$.

A B-scan from one of the produced phantoms is shown in Fig.3.3a.

Chapter 4

Needle tracing algorithm

Needle tracing algorithms can offer different benefits in an image-guided surgery scenario. The first main challenge is to track the needle within the tissue, then trace it over the time.

Two evidences were found about needle tracking in ophtalmologic examinations using OCT [31][32]. Basically, needle tracking can be carried out by two methods: exploiting the image morphological features, or by the semantic segmentation of the needle.

The algorithm described in this work was designed for a dermatological use of the OCT imaging. Since the skin has completely different morphological features from eye tissues, the morphological approach was not considered. Thus, the algorithm was based on a semantic segmentation provided by the training of a customized FCN specifically designed for biomedical image segmentation, called U-net [33].

The network architecture was customized using the Python library Keras powered by Tensorflow. It exploits the GPU acceleration both for the training and the prediction processes.

The U-net was vitally important in order to find the frames of the volume which contained needle sections. Then, the deepest section of the needle was identified and its orientation was evaluated using a Principal Component Analysis (PCA) based method on the segmentation map. This step was necessary to find the closest inclusion pixel to the needle tip.

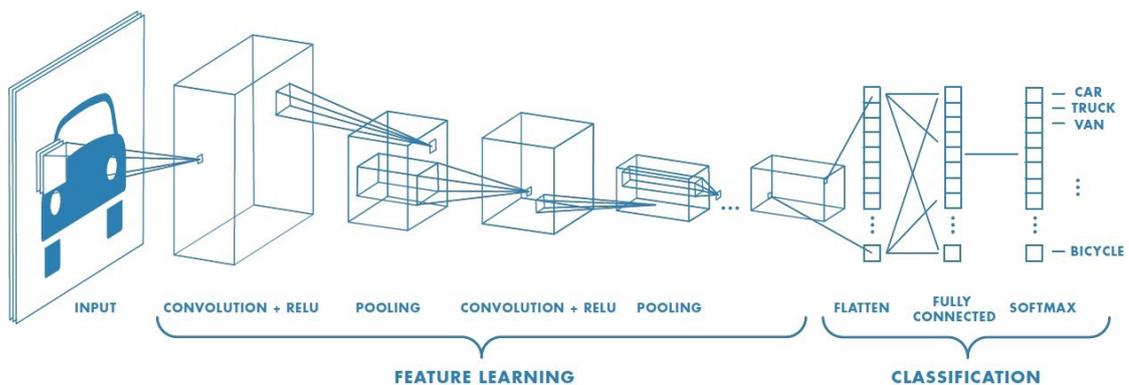
The guidelines of Percutaneous Needle Biopsy (PNB)[34] offered information about which needles are commonly used during biopsies. Fine-Needle Aspiration Biopsy (FNAB) and Core Biopsy (CB) are the two main techniques. This two technique use both hollow needle, but with different gauge. FNAB uses thin, hollow needle (18-25 gauge), CB use hollow needle (9-20 gauge). In this work, a 20 gauge

hollow needle was used in order to simulate both the techniques.

This chapter shows the U-net architecture, the network training and test results, the segmentation performance, the tracing algorithm concept and some tracing simulations.

4.1 Convolutional Neural Network (CNN) VS Fully Convolutional Network (FCN)

Before analyze the U-net, it is important to understand the concept of CNN and the implementations needed to transform it to a FCN. Fig.?? shows the concept of a CNN.



[Concept of a CNN]Concept of a CNN [<https://towardsdatascience.com>]

Let's have a look to the basic operations performed by a CNN [35]:

1. *Convolution*: let's consider an RGB image (width and height of 32×32 pixels, and a depth of 3 RGB channels). To fully connect n neurons to the input layer, $32 \times 32 \times 3$ weight connections are needed. It emerged that the number of connections drastically reduced as the neurons were regional connected (Fig.4.1) and weights were fixed for all the neurons of the next layer. Keeping fixed weights for the local connections is almost the same as sliding a kernel in the input neurons. This windowing-like operation maps the generated output to the corresponding place and allows features detection and recognition regardless of their position in the image. They are called convolutional networks for this reason.

The kernels are squared matrices (usually 3×3 , 5×5 or 7×7). They are also called filters because they act like filters in image processing.

Adding more layers after the input layer makes this method more beneficial. It is possible to associate each layer with different kernels.

Eq.4.1 shows how the convolution is calculated for one pixel in the next layer [35]:

$$net(i, j) = (x * w)[i, j] = \sum_m \sum_n x[m, n]w[i - m, j - n] \quad (4.1)$$

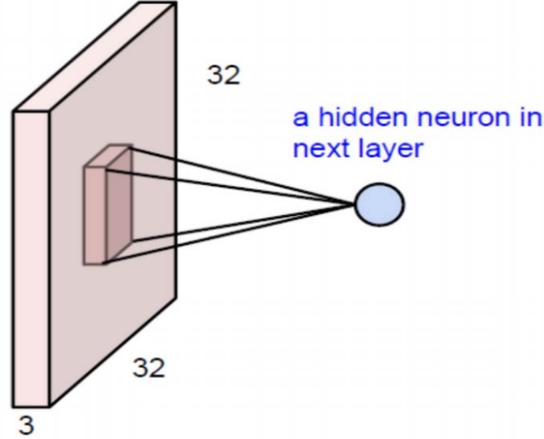


Figure 4.1: Convolution as alternative for fully connected network [adopted from [35]]

2. *Stride*: two adjacent layers can have different overlaps, and stride controls them. A stride equal to 1 means that the filter moves 1 node every time. Considering a 3×3 kernel and stride 1 with a 7×7 pixels image, the output will be 5×5 dimension with different overlaps. Using stride 2, the output size will be 3×3 and the number of overlaps will decrease.
3. *Padding*: the image edges are only captured when the kernel slides, involving a loss of information in these regions. To resolve this issue, a simple zero-padding operation can be used (Fig.4.2).

0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

Figure 4.2: Zero-padding [adopted from [35]]

4. *Nonlinearity*: an activation function is used to saturate or limiting the generated output. The most used activation functions is the Rectified Linear Unit (ReLU), but also sigmoid and tanh can be adopted. ReLU has simple definitions (Eq.4.2), constant gradient, and creates a sparser representation (a zero in the gradient leads to obtain a complete zero). On the other hand, sigmoid and tanh cause problems in the back propagations and their gradients have always non-zero results, which might not be the best solution for training. Fig.4.3 shows different types of activation functions.

$$\begin{cases} ReLU(x) = \max(0, x) \\ \frac{d}{dx} ReLU(x) = \{1 \text{ if } x > 0 ; 0 \text{ otherwise}\} \end{cases} \quad (4.2)$$

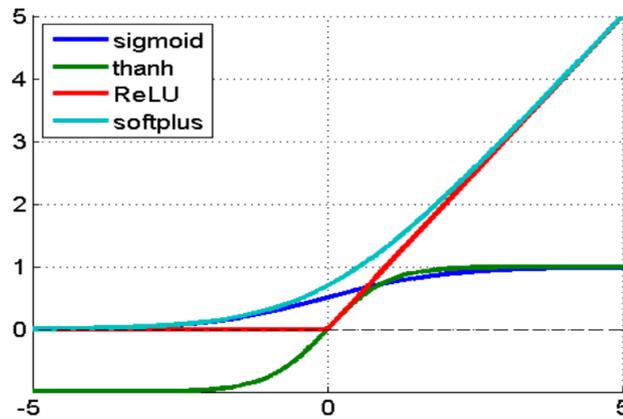


Figure 4.3: Common types of activation functions [adopted from [35]]

5. *Pooling*: this is a down-sampling operation that reduces the complexity for further layers like a sort of resolution reduction. The most used pooling operator is the Max-pooling (Fig.4.4). It returns the maximum value of image sub-region rectangles. The most used max-pooling size is 2x2.
6. *Fully-connected layer*: is similar to the neurons arrangement in traditional neural network. Each node on the last frames in the pooling layer is connected as a vector to this layer (Fig.4.5)

Considering data layers in a ConvNet, their shapes are all defined as a three-dimensional $h \times w \times d$ matrix, where h and w are spatial dimensions, and d is the number of channels. Thus, the first layer is a $h \times w$ image with d color channels. Each location in higher layers is path-connected to locations in the image (receptive fields).

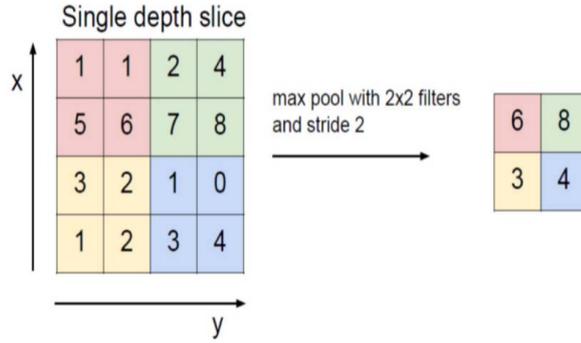


Figure 4.4: How max-pooling is performed with a 2x2 filter and stride 2 [adopted from [35]]

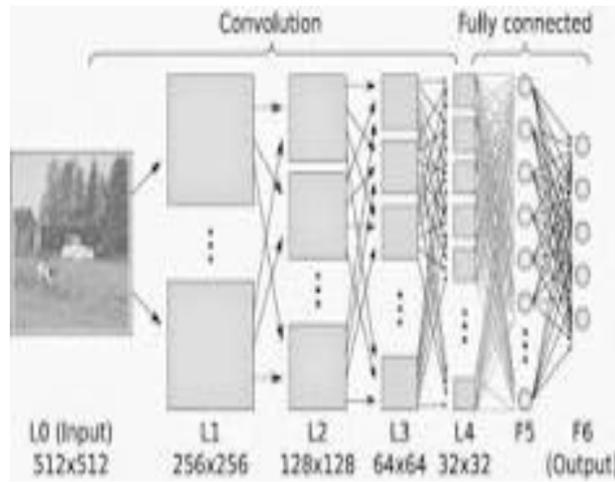


Figure 4.5: Fully-connected layer [adopted from [35]]

Convnets are built on translation invariance [36]. Spatial coordinate dependent basic components (convolution, pooling, and activation) operate on local input regions. The outputs y_{ij} are computed by the Eq.4.3 from the data vector located at the same position (x_{ij}) in a particular layer [36]:

$$y_{ij} = f_{ks}(\{x_{si+\delta i, sj+\delta j}\}_{0 \leq \delta i, \delta j \leq k}) \quad (4.3)$$

where k = kernel size, s = stride (or subsampling factor), and f_{sk} determines the layer type: "a matrix multiplication for convolution or average pooling, a spatial max for max pooling, or an elementwise nonlinearity for an activation function, and so on for other types of layers"[36]. As long as kernel size and stride obey

to the transformation rule in Eq.4.4, the functional form in Eq.4.3 is maintained under composition [36].

$$f_{sk} \circ g_{k's'} = (f \circ g)_{k'+(k-1)s',ss'} \quad (4.4)$$

If all the layers of the net come in this form, the net computes a nonlinear filter which can be called deep filter or fully convolutional network. *"An FCN naturally operates on an input of any size, and produces an output of corresponding (possibly resampled) spatial dimensions."*[36].

Typical ConvNets [37][38][39][40] have fixed dimension fully connected layers and throw away spatial coordinates. These classifiers can be adapted for dense prediction by transforming their fully connected layers into convolution layer with kernels that cover the entire input regions. This operation enables the transformation of ConvNets to FCNs (Fig.4.6). Because of their spatial output maps, FCNs are a valid choice for dense problems like semantic segmentation.

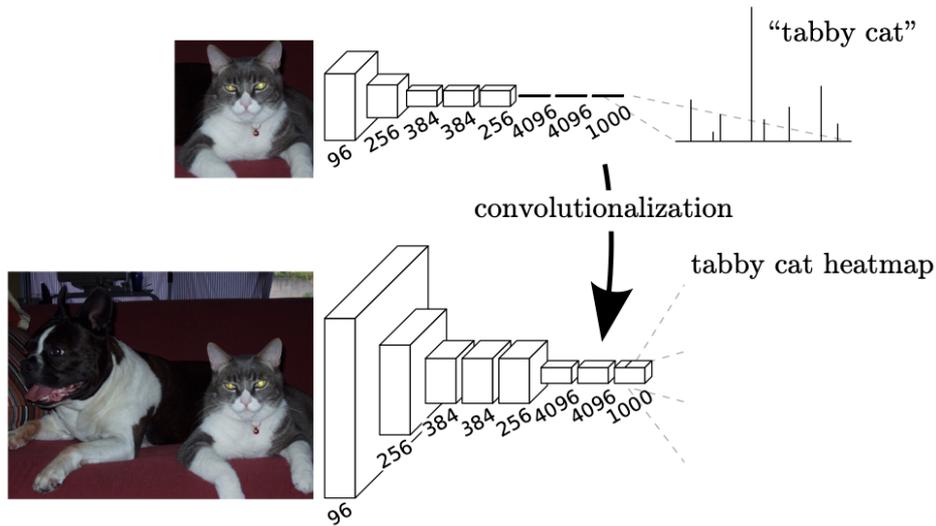


Figure 4.6: CNN to FCN transformation: fully connected layers are transformed in convolution layers enabling a classification net to output an heatmap [adopted from [36]].

4.1.1 U-net

The U-net is an FCN architecture extension that works with very few training images and provides more precise segmentations [33]. This network adds layers where pooling operators are replaced by upsampling operators in order to increase

the output resolution. Hence, U-net consists in one contracting path and one expansive path. Localization derives by the combination of the contracting path high resolution features with the upsampled output. Based on the output information, more precise assembling of the output is obtained by training a successive convolution layer. The large number of feature channels coming from the expansive path propagates context information to higher resolution layer. "As a consequence, the expansive path is more or less symmetric to the contracting path, and yields a u-shaped architecture. The network does not have any fully connected layers and only uses the valid part of each convolution, i.e., the segmentation map only contains the pixels, for which the full context is available in the input image. This strategy allows the seamless segmentation of arbitrarily large images by an overlap-tile strategy... To predict the pixels in the border region of the image, the missing context is extrapolated by mirroring the input image. This tiling strategy is important to apply the network to large images, since otherwise the resolution would be limited by the GPU memory."[33].

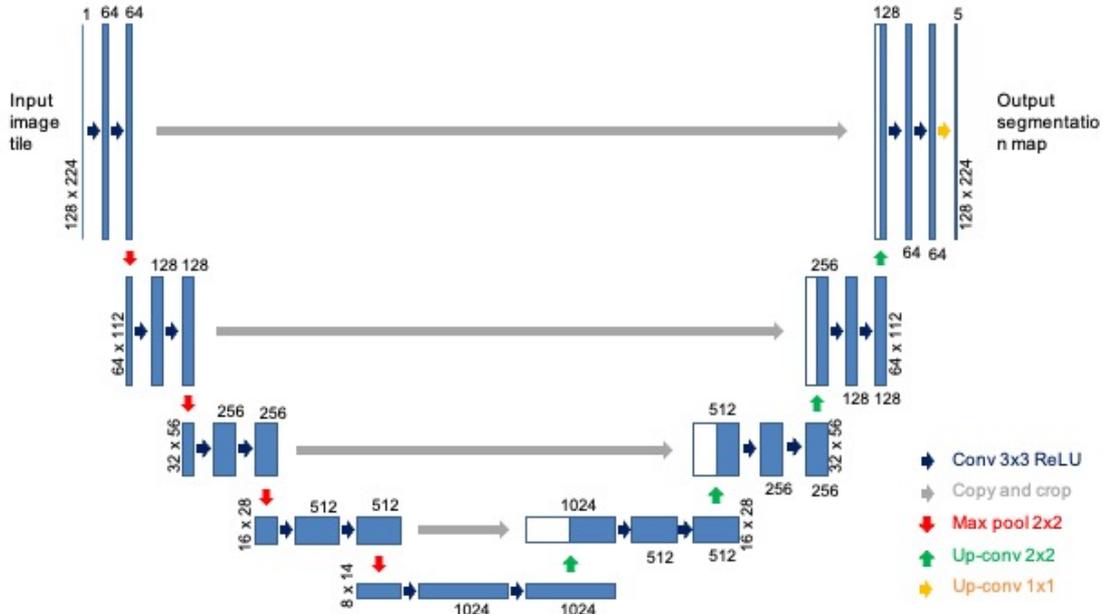


Figure 4.7: Custom U-net architecture. Multichannel-feature map can be identified by the blue-boxes, copied feature maps by the white ones. The layer dimensions are denoted on the left edge (x-y-size), and on top (number of channels) of the boxes. The color of the arrows denotes the operation (legend on right bottom corner).

The custom U-net architecture is shown in Fig.4.7. Each layer of the contracting path consists of a double step 3×3 padded convolution followed by ReLU and

2×2 max pooling operation with stride 1. The number of feature channels is doubled at each downsampling step. Each layer of the expansive path consists of an upsampling operation, a 2×2 "up-convolution", a concatenation, a 3×3 padded convolution, and ReLU activation. The "up-convolution" halves the number of feature channels. The "up-convolution" output is then concatenated with the corresponding contracting path cropped feature map. Cropping operation is needed because of the loss of border pixels after each convolutions step. Then, each of the 64-component feature vector is mapped to the desired number of classes by a 1×1 convolution. A total of 23 convolutional layer are used to design this network.

4.1.2 Inputs and network training

For the network training, a substantial amount of images were needed. Before the acquisition, the system made it possible to detect the central B-scan position by emitting a IR-light beam alongside that plane. The needles were kept as parallel as possible to the x or y-axis in order to maximize the visible needle section in each image.

For this work, 1286 Maximum Intensity Projection (MIP)[41] grayscale images (512×896 pixel) were recorded in a database with their respective segmentation maps. MIP images are constructed by the projection of 8 aligned B-scans. Each pixel of a MIP image correspond to the maximum intensity signal among the pixels of the B-scan packet. MIP was performed in order to maximize the needle section. To double the dataset dimension, all the images were flipped on the z-axis. In order to increase the variety of distances from the 0-delay line assumed by the phantom, MIP images were shifted up or down by different quantities. All the images were resized by a factor of 4 in order to reduce the network size[32].

All the images were segmented manually using ImageJ in 5 different labels: the needle, the inclusion, the tissue, the needle shadow, and the background. Four possible real situations were detected: images with only the tissue, images with inclusion and tissue, images with tissue and needle, and images with tissue, needle and inclusion. For the network training, the number of the images representing the four different situations have been balanced. 2000 images were selected as training set (500 images for each situation) and 200 as test set. The validation set was created from the training set extracting 400 images from the training set (20%).

The network was trained for 50 epochs. The training could be stopped by an "EarlyStopping" callback when the validation loss score did not vary more than 0.01 points for 13 consecutive epochs. Network performance with different loss functions were investigated using always the same network optimizer (*Adam*) and learning rate (10^{-4}). During the training a callback was used to reduce the learning rate up to 10^{-7} when the validation loss score did not vary more than 0.1 points for

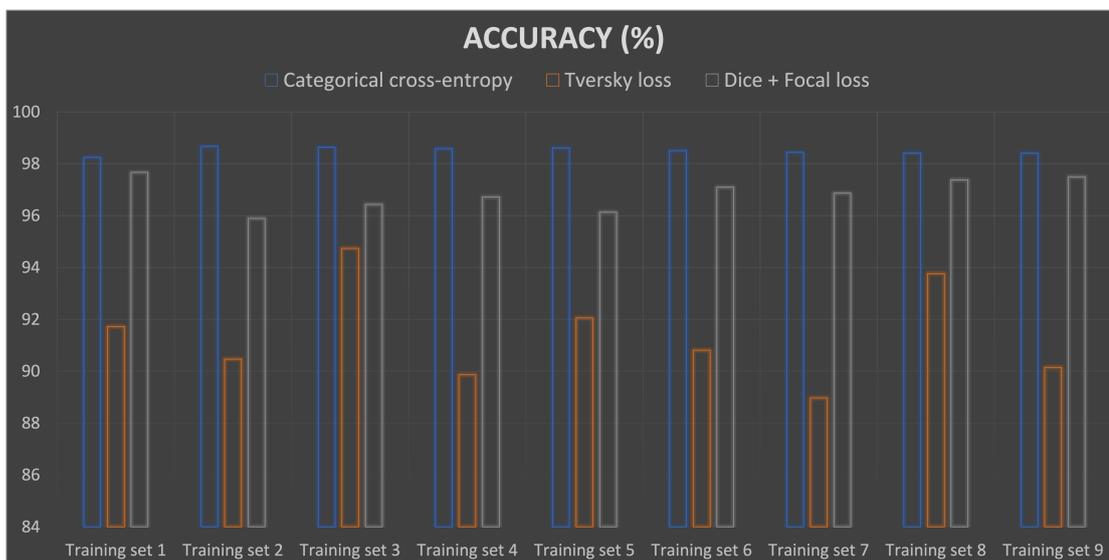
3 consecutive epochs. The loss functions used were: *categorical cross entropy*[42], *Tversky Loss*[43], and a combination of *dice loss* and *focal loss*[44]. The accuracy and the Intersection over Union (IoU) scores were crucial to decide which loss function is the best for the purpose. The IoU is an evaluation metric used to measure the accuracy of an object detector on a particular dataset. It can be defined as the ratio between the overlap area and the union area of the predicted bounding box and the ground-truth bounding box(Fig.4.8).

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

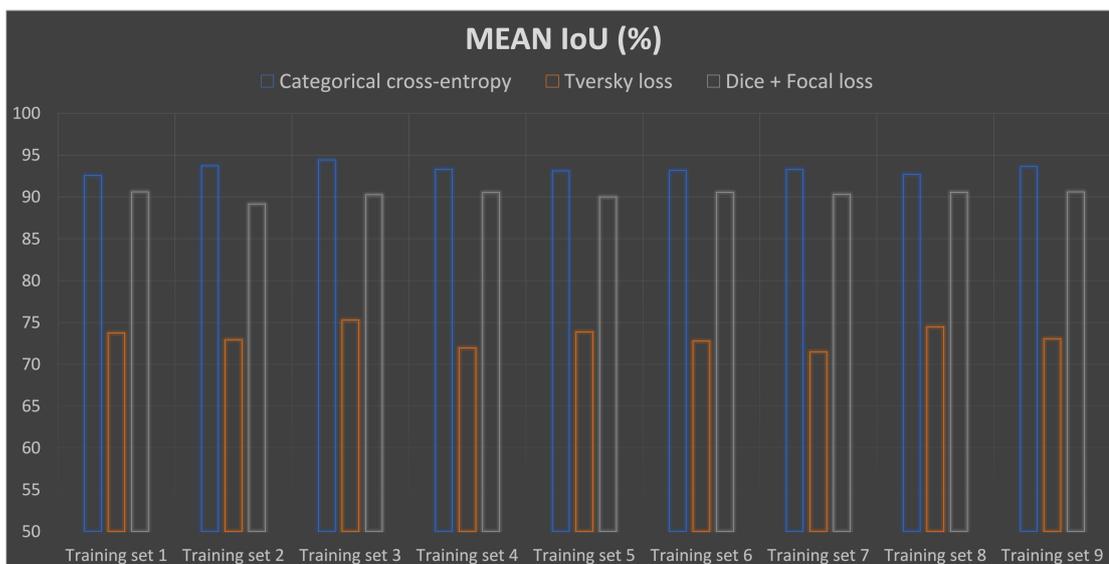

Figure 4.8: Intersection over Union equation.

Fig.4.9 shows a comparison of the accuracy and the mean IoU of the trained U-net over the nine different training sets with the three loss functions.

All the loss functions showed stable performance, but *categorical cross entropy* was the one with the best one.



(a) Accuracy scores comparison over the different loss functions



(b) IoU scores comparison over the different loss functions

Figure 4.9: Accuracy (a) and IoU (b) score comparison over the nine training sets with the three different loss functions.

4.1.3 Outputs and performance

The average number of epochs required for the training was 35 for an average training time of 39 minutes. An example of training and validation loss score and training and validation accuracy score trend are reported in Fig.4.11 and Fig.4.10.

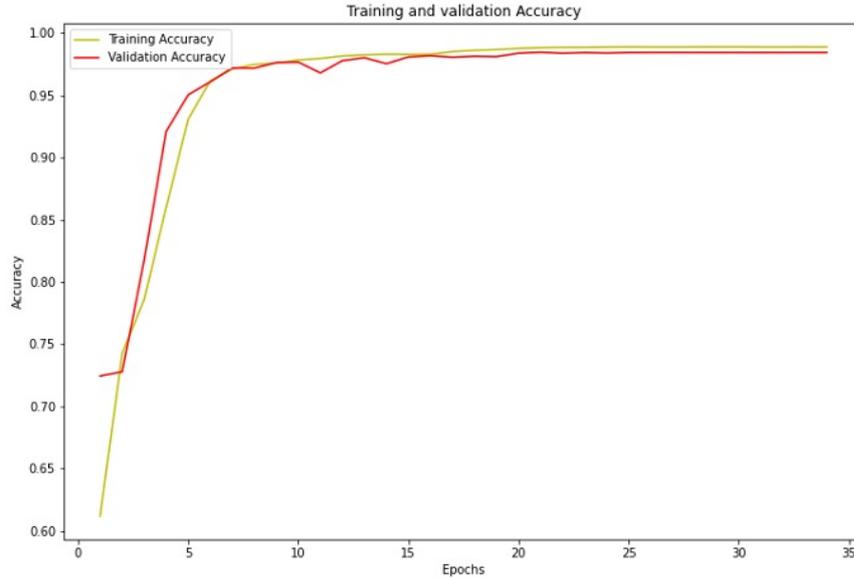


Figure 4.10: Example of a training and validation accuracy trend over the epochs. The yellow trend is for the training, the red one is for the validation.

In order to evaluate the network performance, eleven parameters were calculated: accuracy, mean IoU score, IoU score for each label, precision, recall, F1-score, and dice score. The results are reported in Fig.4.12.

The mean accuracy and the mean IoU scores were respectively 98.49% and 93.34%.

The mean IoU performance referred to the single labels were: 79.96% (needle), 96.98% (inclusion), 97.10% (tissue), 92.80% (needle shadow) and 99.34% (background).

The precision, recall, F1-score, and dice score mean value were respectively 98.31%, 97.85%, 98.02%, and 97,23%.

Comparison between the U-net output and the corresponding mask are shown in Fig.4.13.

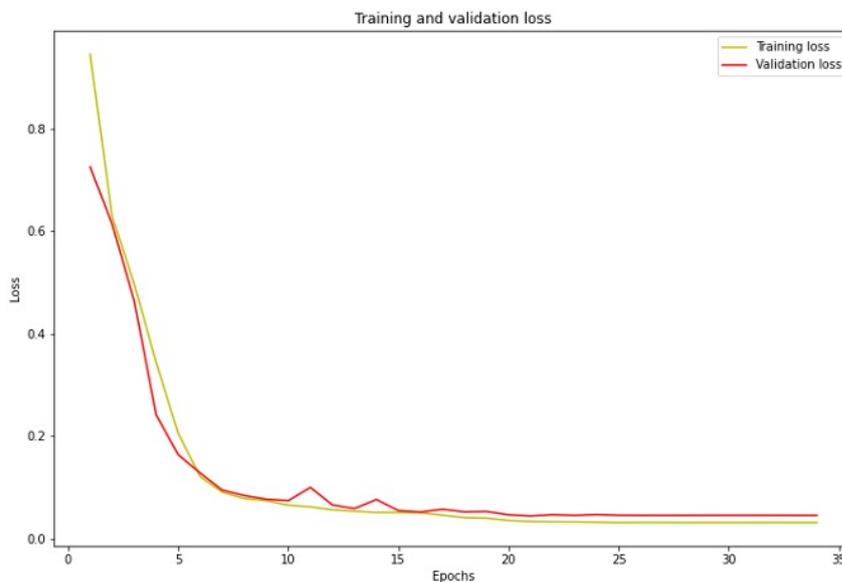


Figure 4.11: Example of a training and validation loss trend over the epochs. The yellow trend is for the training, the red one is for the validation.

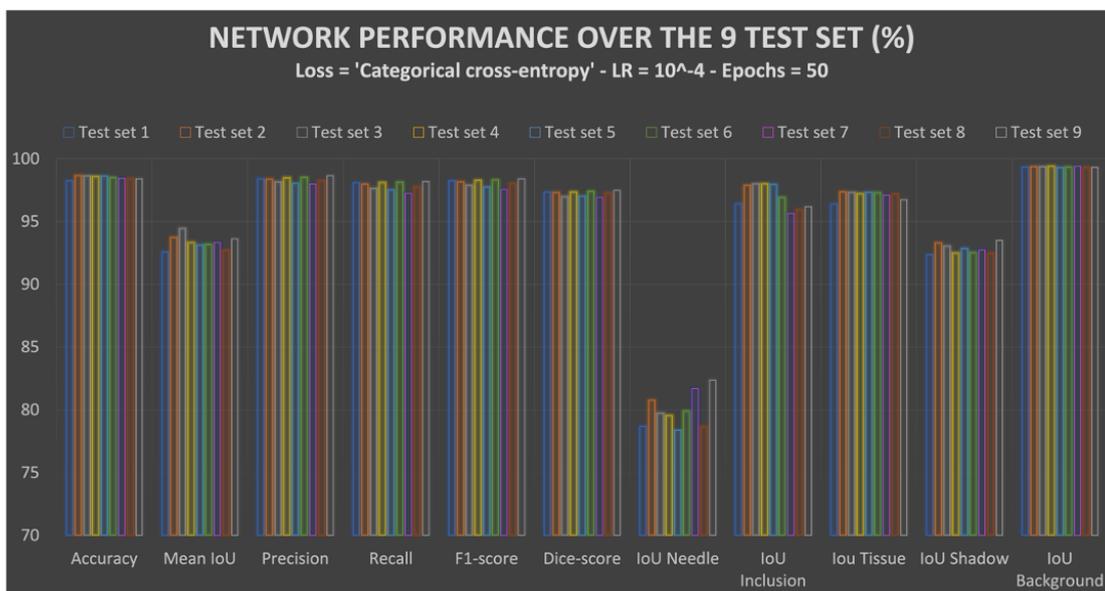


Figure 4.12: Network performance: Accuracy, mean IoU, label IoU, precision, recall, F1-score, and dice score over the 9 training sets

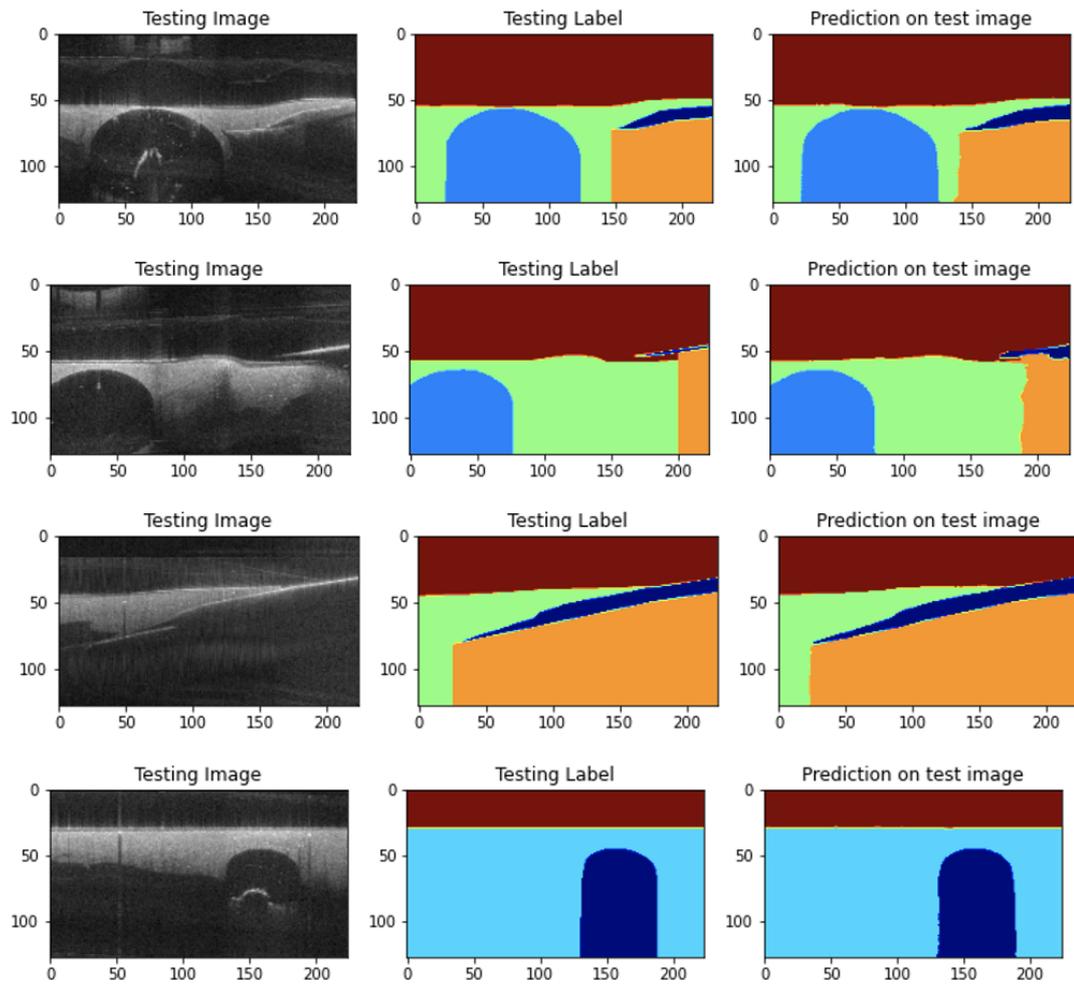


Figure 4.13: U-net output example.

4.2 Algorithm

The needle within the phantom can be observed by looking at the B-scans or the cross-sections of the volume (Fig.4.14). In both cases, only a section of the needle can be seen. Sagittal or frontal needle sections can be inverted as the needle is inserted alongside the B-scan or the cross-section axis. The best needle visualization is reached when the needle direction is as parallel as possible to one of this two planes.

Unlike the evidences founded about FCN for needle segmentation[31], in this work the U-net was used to segment sagittal needle sections. Considering a 20 gauge needle (0.9mm diameter), if it was inserted perfectly parallel to the B-scan axis or to the cross-section axis, at least 80 tomograms contained a sagittal needle section.

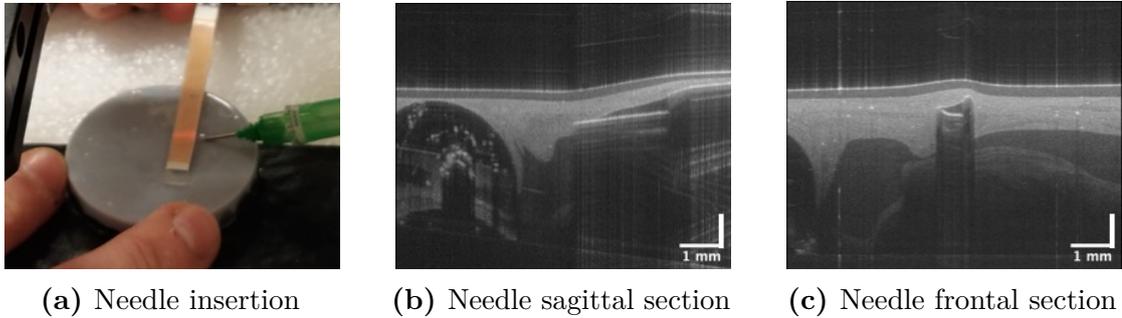


Figure 4.14: Needle visualization: (a) the sagittal plan of the needle was inserted as parallel as possible to the B-scan axis. (b) Needle sagittal section, (c) needle frontal section.

The tracing algorithm was designed as a two steps scanning process over the volume. Fig.4.15 shows the concept of the tracing algorithm.

The first scanning step was designed to find the frame of the volume which contained a needle section.

MIP was performed over packets of 8 adjacent cross-sections. This operation worked as a volume compression from $512 \times 896 \times 896$ to $512 \times 112 \times 892$ pixels. Furthermore, by compressing frames together the needle section was maximized (Sec.4.1.2) and the number of frames to segment were reduced. If the uncompressed volume contained at least 80 needle section, after the MIP the needle could be found in at least 10 frames. Thus, to be sure to find a needle section, a U-net prediction must be performed every 10 B-scans.

The MIP volume was resized by a factor of four in order to fit the data into the U-net.

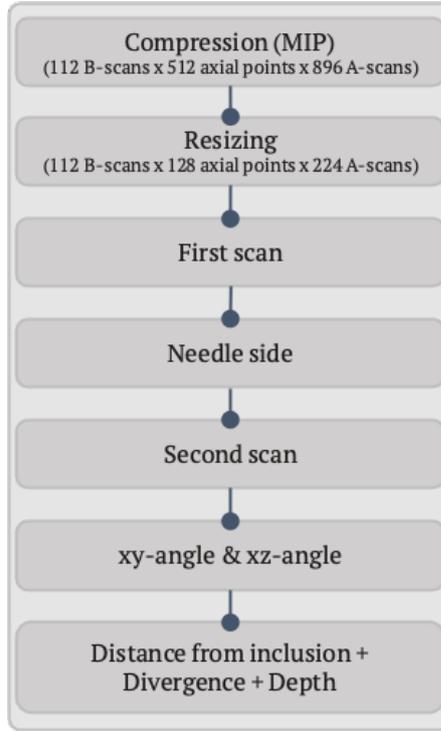


Figure 4.15: Tracing algorithm concept.

Because the number of frames to analyze was a multiple of 8, the network predictions of the first scanning step were performed every 8 frames. Starting from the first frame and only for the first iteration, if at least one pixel belonged to the needle label was found, the number and the coordinates of these pixels were counted and saved. A mask with only the pixels belonging to this label was created.

Noisy needle pixels areas were removed by a combination of a binary erosion operator, a small object removal, and a binary dilatation operator. The erosion separated these areas from the edges of the image or from the needle shape. Then, if smaller than half of the number of needle labeled pixels, these areas were removed. Finally, binary dilatation operator was applied to the image in order to restore the resolution lost during the binary erosion. Then, the amount of needle labeled pixels was measured and saved with their related coordinates for the next iteration.

For the next iterations, almost the same process was repeated until all the volume was scanned. The only difference was that the needle labeled pixels were counted only once at the end of the iteration (not also after the prediction). If the amount of these pixels was bigger than the previous iteration, its area value, the pixel coordinates, the binary mask, and the frame position were saved and the older ones deleted (Fig.4.16).

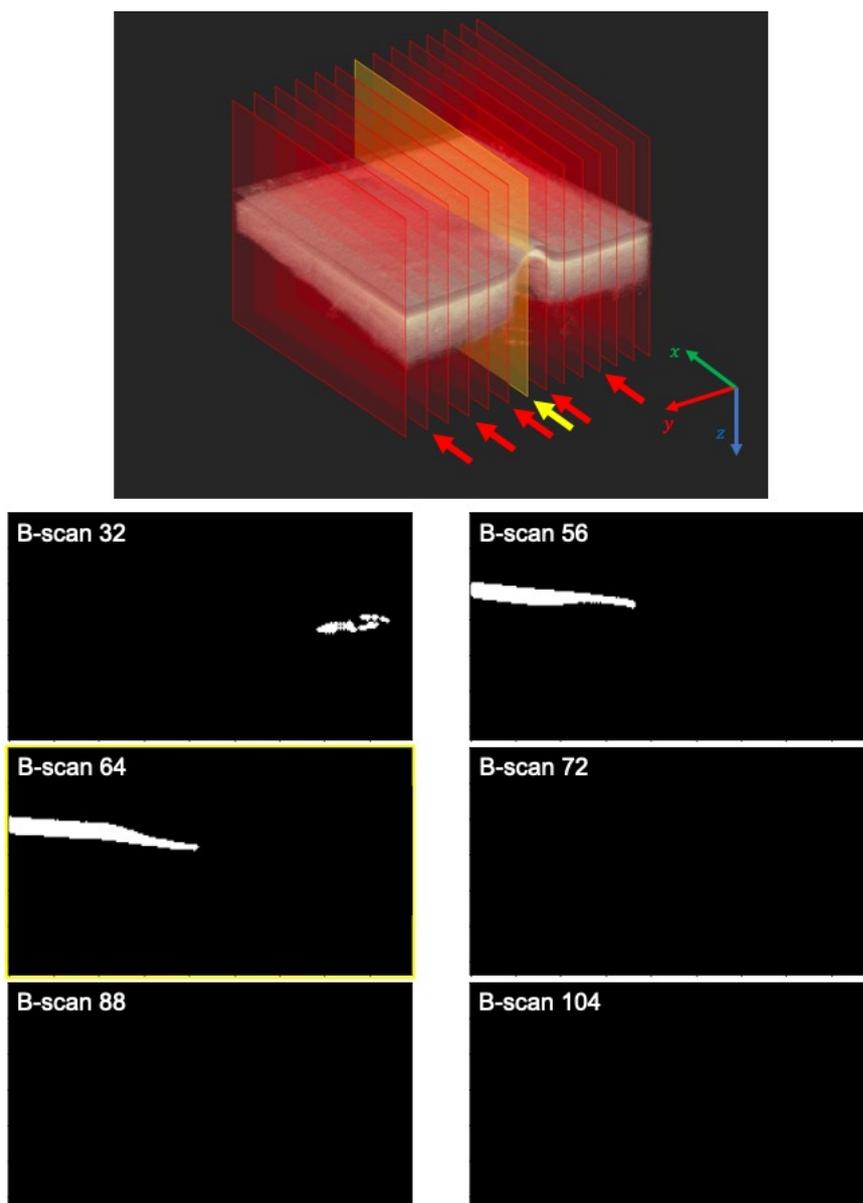


Figure 4.16: First scanning step. Red-highlighted B-scans show where the first scan made a prediction. Red arrows show where the prediction found some needle labeled pixels. The needle masks of these predictions, with their related position, are reported under the volume image. Totally black masks refer to insignificant predictions. The yellow-highlighted B-scan shows where the first section of the needle was found.

Then it was possible to find the needle tip coordinates from the binary mask previously saved. If the needle was inserted left-to-right in the phantom (Fig.4.17a),

the needle tip x-coordinate was the maximum one over the x-coordinates saved during the first scanning step. By its index position, the related needle tip y-coordinate could be found. The same reasoning was made to search and save the coordinates of the first needle labeled pixel. Since the needle could be inserted also from the other side of the phantom, the algorithm must be able to recognize between these two situations. If the needle was inserted right-to-left (Fig.4.17b), these two points would be confused with each other. In this case, this two pixel coordinates were to be exchanged.

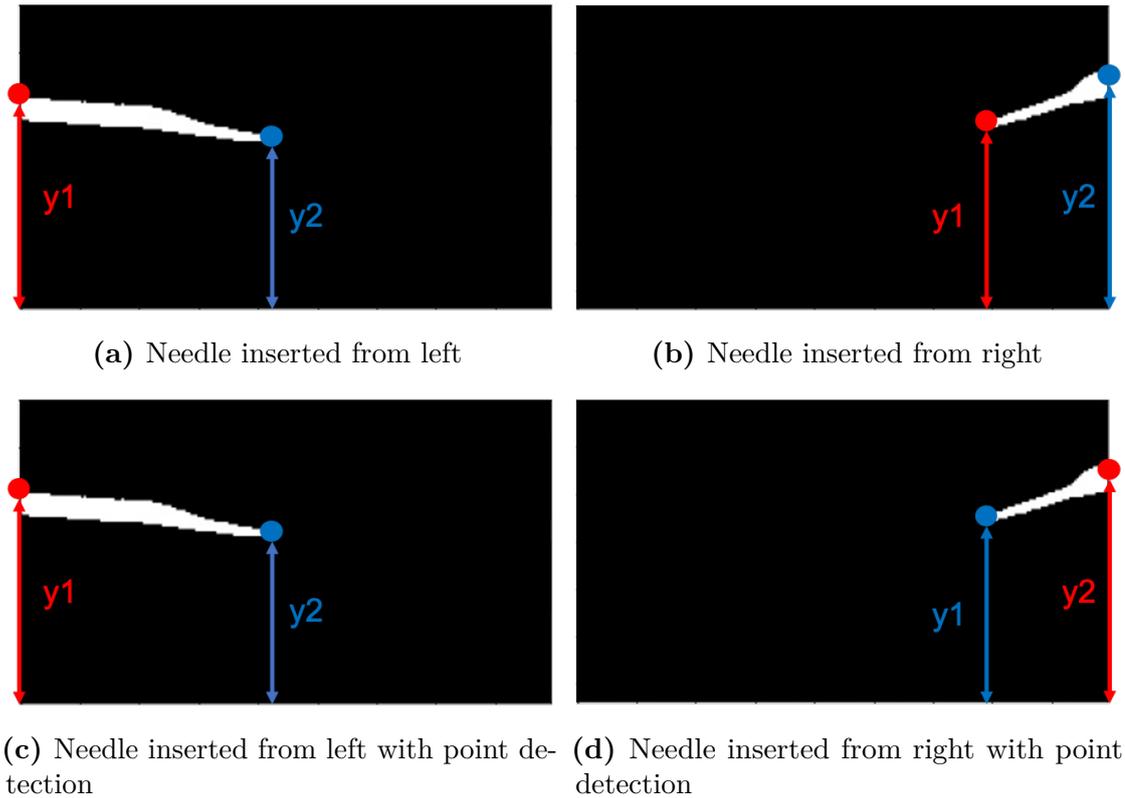


Figure 4.17: Visual comparison between different side needle insertion. The blue highlighted pixel is the one detected as the needle tip, the red one is the first needle labeled pixel detected. When the needle is inserted right-to-left these two points must be exchanged.

Considering Fig.4.17a and Fig.4.17b, the blue point was used to highlight the pixel classified as the needle tip, the red one to highlight the first classified needle labeled pixel. In the situation showed in Fig.4.17a, the blue point y-coordinate was smaller than the red point one. The opposite happened when the needle was inserted from the right side (Fig.4.17b). Thus, knowing this relation, it was possible

to solve this problem with a simple control over the two points y-coordinates (Fig.4.17).

Because there was no certainty the frame founded at the first scan showed the deepest needle section, a second scanning step was necessary. The inputs for the second scanning step were the tip coordinates, the size of the needle area and the frame position calculated at the previous scan.

The output frame of the first scan was taken as a reference from where to start the second step. The second and the fourth frames before and after the reference one were predicted. It was necessary to move this sort of a window if the reference frame was too close the right or the left edge section of the needle. This was done by predicting both the two frames at the ends of the window. If no needle labeled pixels were found in one of the two predicted images, the window was moved by two positions further to this frame, but closer to the reference frame. If both the frames did not show any needle labeled pixel, the window was kept in the original position.

Starting from the first window frame, the needle mask was created as in the first scanning step. After each iteration, the frame position, the tip coordinates and the needle area size were saved only if two conditions were verified:

1. If, after removing the small needle labeled areas, any needle shape was still visible;
2. if the actual tip x-coordinate was greater (needle inserted left-to-right), or lesser (needle inserted right-to-left), than the one saved at the previous iteration.

The occurrence of the first condition did not mean the absolute absence of a needle section within the frame. It could mean that needle section was detected, but it was removed because too small. This step was designed just to find the deepest needle section, not the a generic one (as done for the first scanning process). If no deeper needle section were found, the output frame would be the same calculated after the first scanning step.

Then, a last control was made on the frame before and after the one with the deeper needle tip position (Fig.4.18). The outputs of the second scanning step were the deepest needle section frame position, the mask of the needle, the needle tip coordinates and the related prediction image.

Using a PCA based method on the output binary mask[45], it was possible to calculate the inclination of the needle ' m_{xy} ' to the En-face plane (xy). The equation of the oriented line passing through the needle tip was calculated as in Eq.4.5. By knowing the xy-inclination of the needle, the correspondent angle in

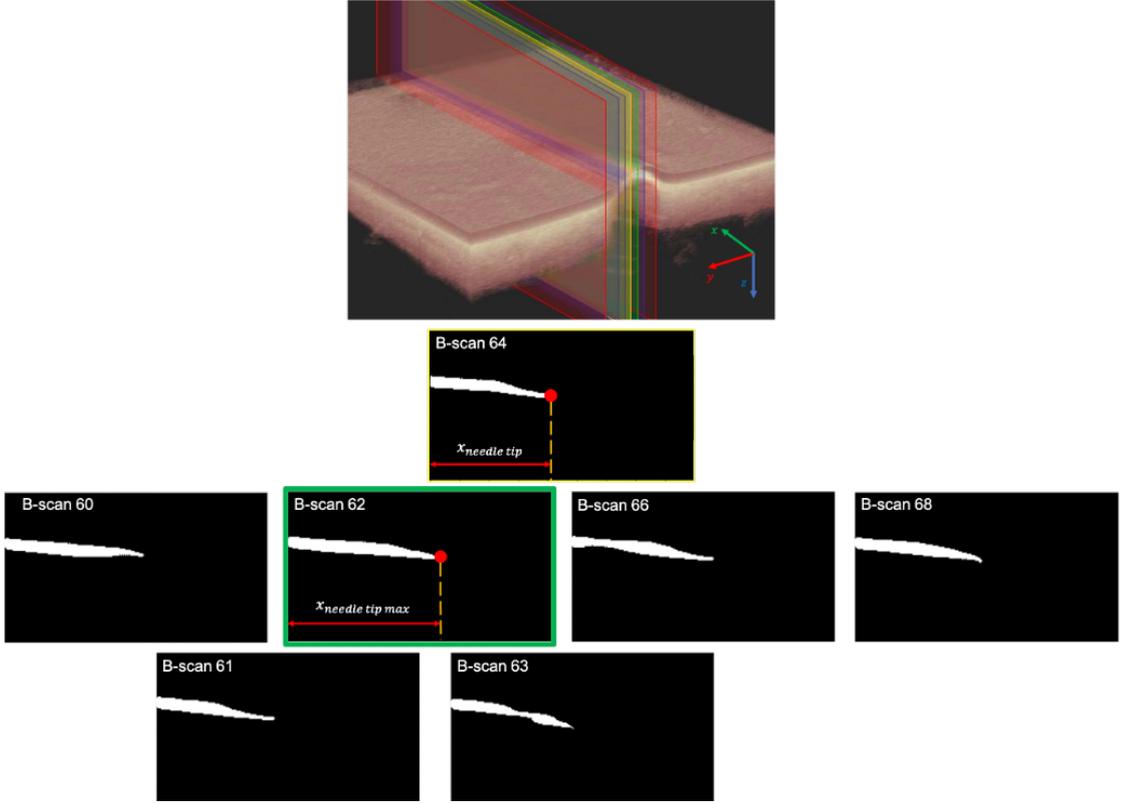


Figure 4.18: Second scanning step. All the blue highlighted B-scans were predicted and the mask with the highest x-value of the needle labeled pixels was saved. Starting from this mask, the two adjacent B-scans were predicted and another check was made in order to find a deeper needle section. Finally, the B-scan corresponding to the highest x-value mask was considered as the output of the second scan. The "closest needle" section is the one highlighted in green.

degrees could be measured.

$$\begin{cases} m_{xy} = \arctan(PC_1) \\ y = m_{xy}(x - x_{tip}) + y_{tip} \\ \phi = m \frac{\pi}{180} \end{cases} \quad (4.5)$$

where, PC_1 was the first PCA component, x_{tip} was the x-coordinate of the needle tip, and y_{tip} was the y-coordinate of the needle tip. The first inclusion pixel was searched over the line equation. The distance of the needle tip to the inclusion was measured as an Euclidean distance (Eq.4.6).

$$d = \sqrt{(x_{tip} - x_{inclusion})^2 + (y_{tip} - y_{inclusion})^2} \quad (4.6)$$

It was assumed the needle moved in a parallel direction of its sagittal plane. If the needle was not positioned such parallel to the B-scan plane, the angle formed by its sagittal plane and the xz-plane were to be taken into account.

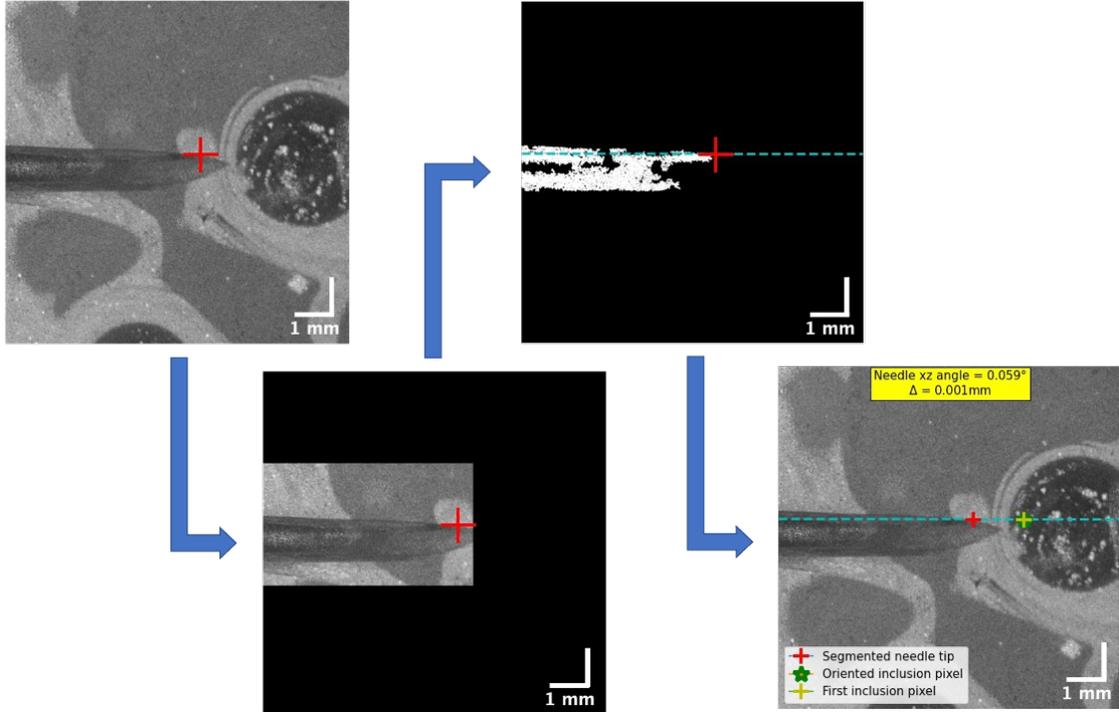


Figure 4.19: Correction of the distance and xz-angle calculation.

The needle tip depth was known by the output y-coordinate of the second scanning step. Thus, it was possible to visualize a slightly deeper En-face image of the phantom. On this image, the mask of the needle was created by global thresholding a ROI which contained the needle shadow. The ROI dimension was automatically generated. If the needle was inserted left-to-right, the ROI width was defined by all the columns between 0 and the needle tip x-coordinate plus forty. In the other condition, it was defined by all the columns between the needle tip x-coordinate minus forty and the second dimension size of the En-face (896). The ROI height was defined considering 160 rows before and after the second scan output frame. All the pixel outside the ROI were equalled to 0. The mask was generated by increasing the contrast of the image as done for the volume registration (Sec.2.2), and thresholding the image. The threshold value was 0.1. All the pixels with intensity greater than the threshold were equalled to 1, the other ones were equalled to 0.

The needle sagittal plane direction, and its related angle θ , were measure by

the same PCA methods used for the B-scans. The distance from the inclusion was corrected by a simple trigonometric relation as the inclusion pixel was projected on the needle sagittal plane. In addition, also the divergence between the inclusion pixel and its projection was measured. That provided a quantitative error related to the divergence of the needle direction from the B-scan plane (Eq.4.7). A visual example of the tip correction was shown in Fig.4.19.

$$\begin{cases} d' = \frac{d}{\cos |\theta|} \\ \Delta = d' \sin |\theta| = d \arctan |\theta| \end{cases} \quad (4.7)$$

The final output of the algorithm is shown in Fig.4.20. On the output B-scan, the needle tip, the distance from the inclusion and the needle depth were marked. The En-face was displayed in order to provide information about needle xz-orientation and the error Δ committed by only segmenting the needle on the B-scan plane. Δ depended on the xz-needle inclination. As θ approached to zero (needle perfectly parallel to the x-axis), Δ was minimized.

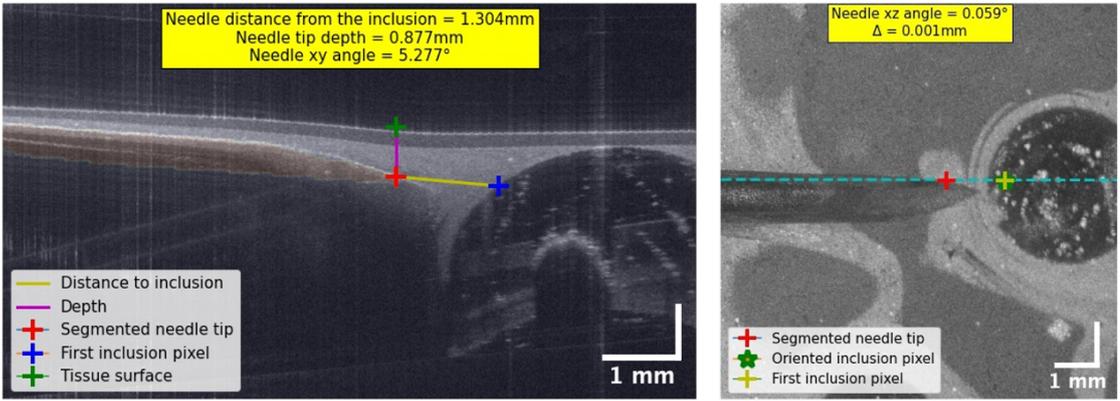


Figure 4.20: Example of the final algorithm output.

4.3 Simulations

Nine tracing simulations were recorded. For each simulation, three subsequent positions of the needle were acquired. The efficiency of the tracing algorithm was analyzed by inserting the needle in different positions. Three examples are reported in this section.

Once the needle was inserted in the desired position, the phantom was blocked with pins to a polystyrene block which was anchored to the OCT system support table by four metal blocks screwed on it (Fig.4.21). Therefore, it was possible to insert the needle deeper without moving the phantom.

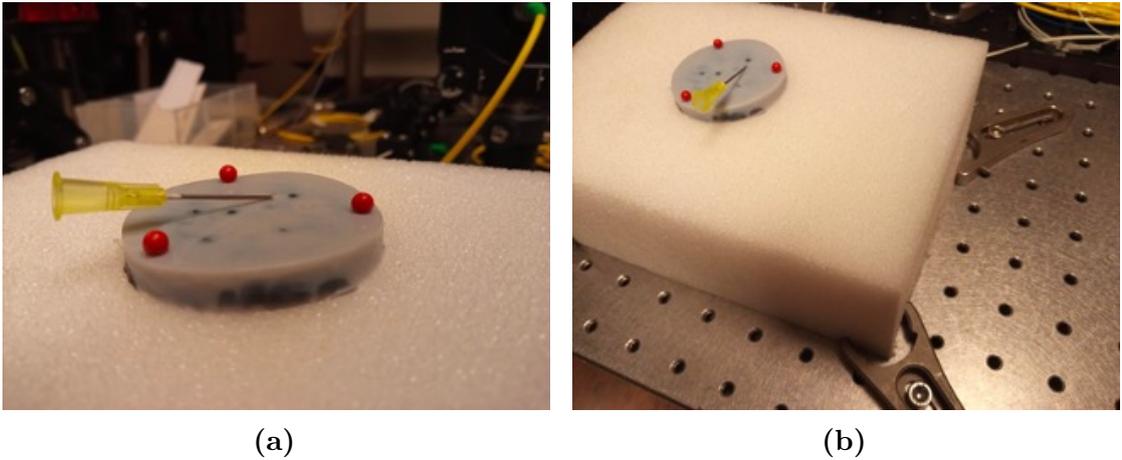


Figure 4.21: Phantom locking system. (a) The phantom was pinned to the polystyrene block (b) which was kept stable by four metal blocks screwed in the OCT support table.

In the first example, a tracing simulation was recorded by inserting the needle left-to-right alongside the B-scan plane and acquiring three different positions of the needle (Fig.4.22). In the first position, the needle was not detected because an area of the inclusion was labeled as a needle section by the network. That area was such big that the small needle area was removed by the algorithm. Because of that, the ROI on the En-face was not positioned correctly and, consequently, the measured xz-angle was not referred to the En-face needle mask. When this situation occurred over the simulations, that position was classified as *"No detection"*. The second position of the needle was correctly traced, but the traced needle tip did not belong to the needle because of a wrong segmentation. Moreover, this wrong segmentation led to an high error measurement of the needle tip depth. Situations like this one were labeled as *"Noisy detection"*. Also the needle in the third position was correctly traced. However imprecise the segmentation was, this time the detected

needle tip belonged to the needle section ("*Good detection*").

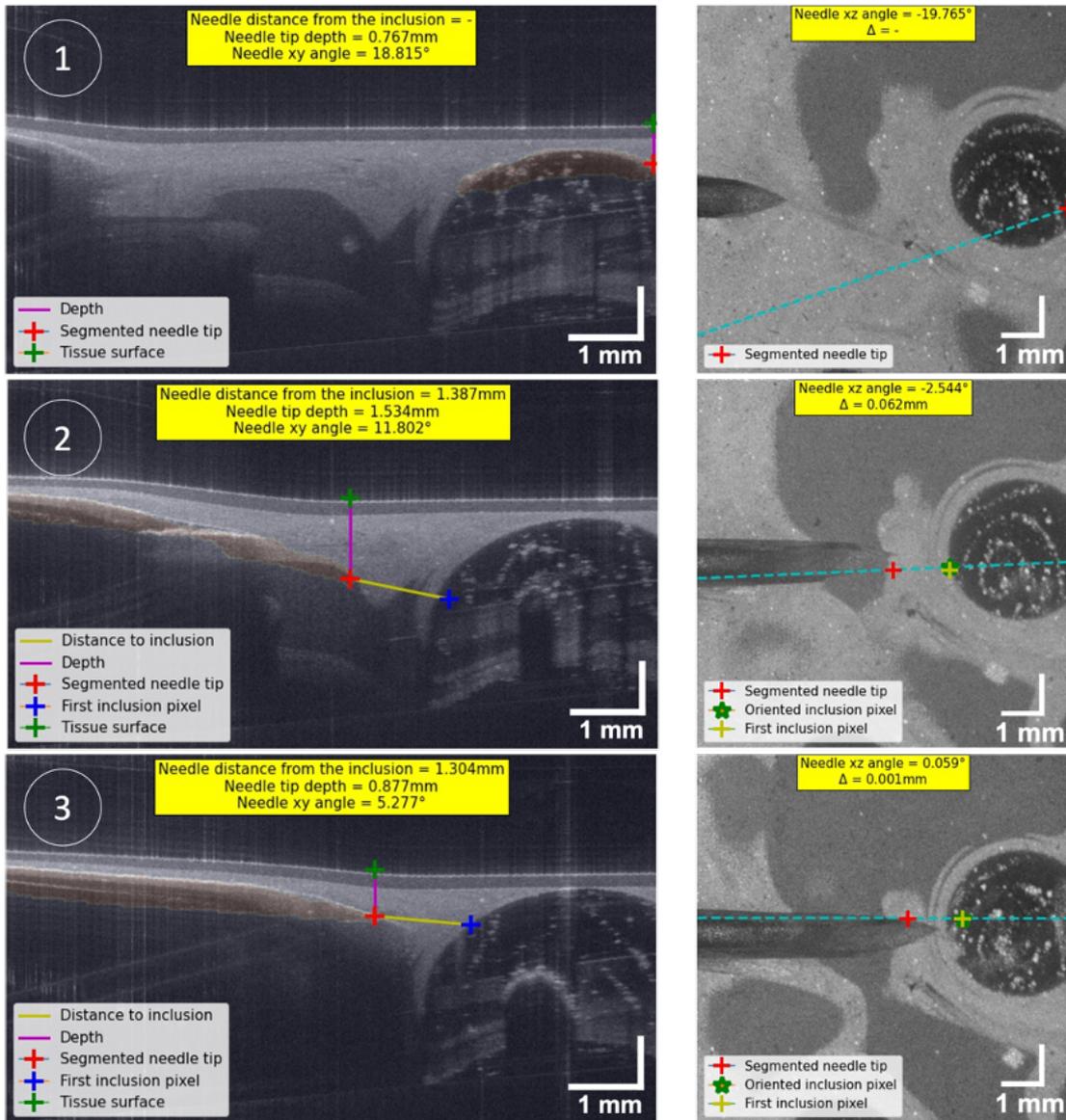


Figure 4.22: First simulation. The needle was inserted left-to-right alongside the B-scan plane.

For the second simulation, the needle was inserted right-to-left alongside the B-scan plane (Fig.4.23). This simulation wanted to verify the efficiency of the algorithm when the needle was inserted on the other plane side. The needle was always detected on the correct side and all three the positions were "*Good detection*".

In the third simulation, the needle was inserted with a significant xz-angle.

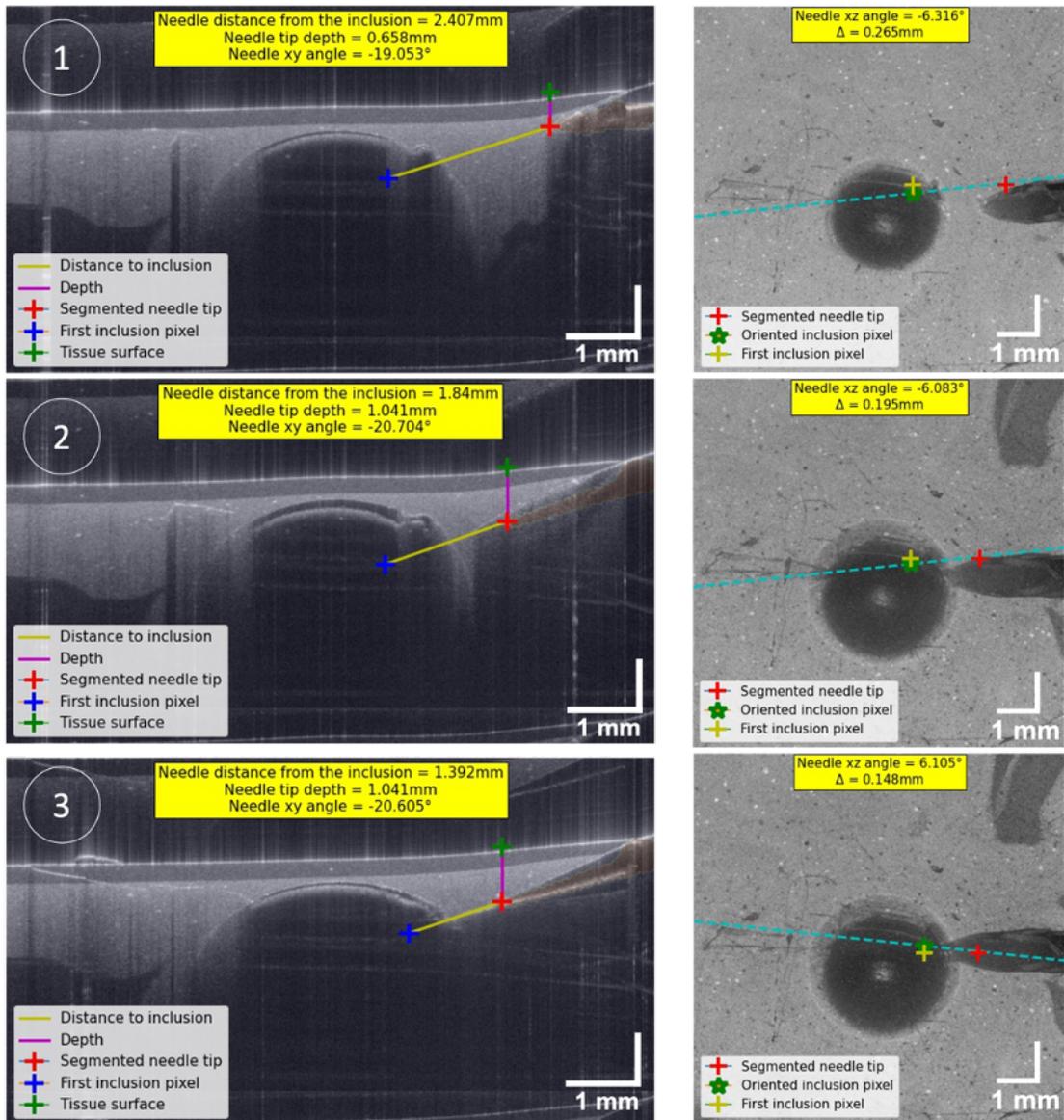


Figure 4.23: Second simulation.

In addition, the angle was changed during each position in order to see if the PCA-method was precise. This simulation wanted to highlight if the algorithm could trace the needle also when it was not positioned parallel to the xz or to the yz-plane. The MIP was performed over the B-scan axis. However the first and the second position were "Noisy detection", the third was a good one, and all the angle measures were correct.

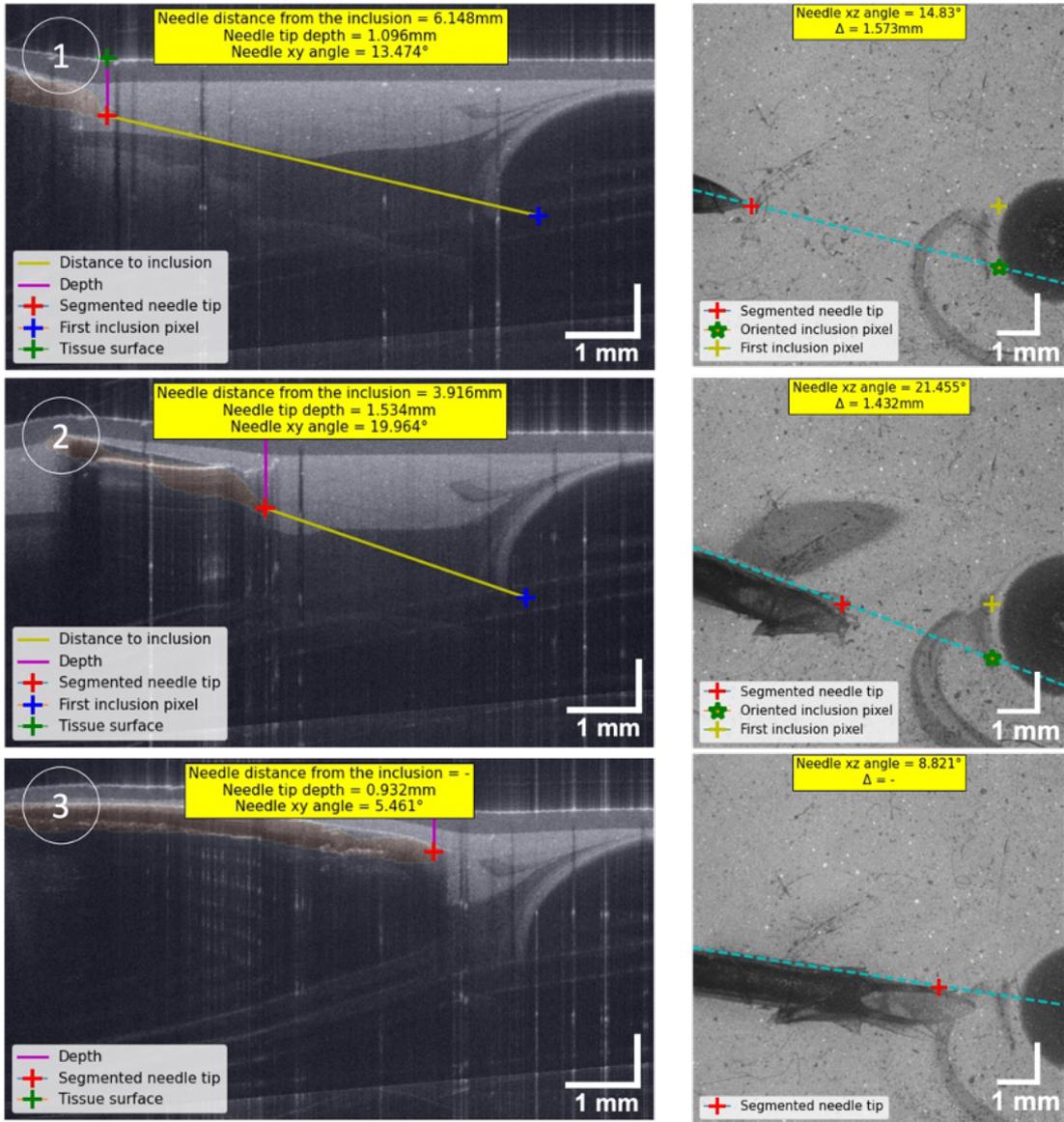


Figure 4.24: Third simulation.

4.4 Errors

Between all the simulations, 3 "No detection", 7 "Noisy detection", and 17 "Good detection" were recorded. Considering both the noisy and the good detections, the needle was detected in almost the 90% of the simulations. However, the tip coordinates founded by the algorithm were not so precise. The errors in the three directions Δx , Δy , and Δz were measured as shown in Fig.4.25.

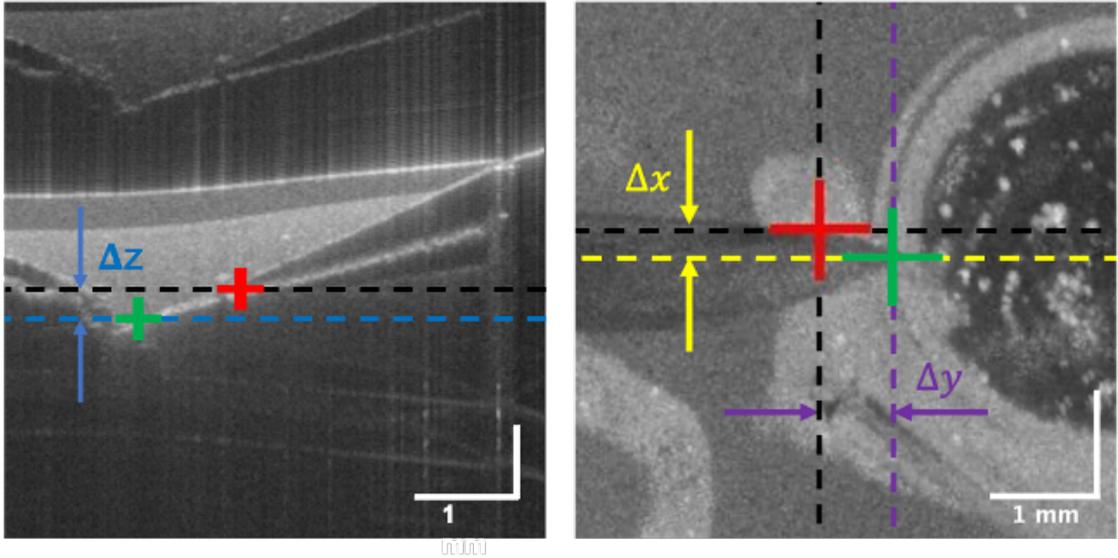


Figure 4.25: Needle tip position error measurement.

Each segmented needle tip was compared with the frame of the volume which had the deepest needle section. This frame was selected manually. On it, the output tip position was projected so it was possible to measure the error Δz . Δx and Δy were measured directly on the output En-face. All the errors were measured on ImageJ.

The simulation positions where the needle was not detected at all were not considered. The averaged errors were $\Delta x = 0.498 \pm 0.242mm$, $\Delta y = 0.324 \pm 0.135mm$, $\Delta z_{good} = 0.237 \pm 0.081mm$, and $\Delta z_{noisy} = 0.572 \pm 0.057mm$ (Fig.4.26).

The inclusion was detected in 74% of the position where the needle was detected. For the remaining 26%, it was not possible to measure the distance from the inclusion and therefore also its divergence from the needle direction.

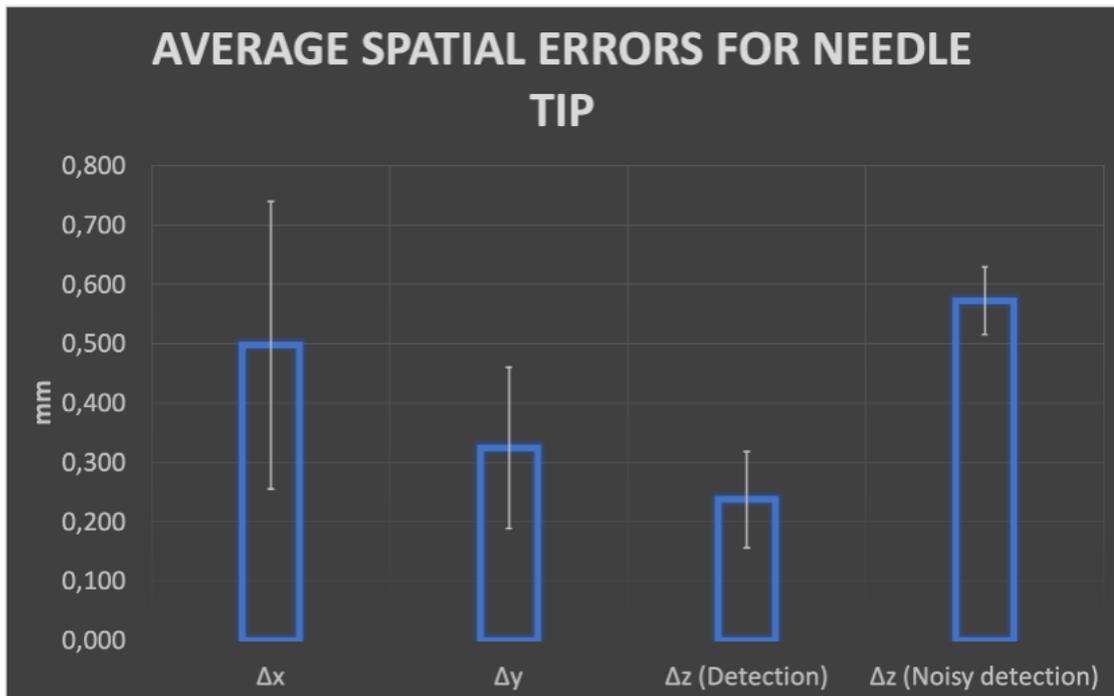


Figure 4.26: Tracing needle tip error on x, y, and z-direction.

4.5 Timings

The processing time depended on the GPU model provided by Colab. As shown in Fig.4.27, as was to be expected, the algorithm was significantly faster on the Nvidia Tesla T4 than the Tesla K80 model because of its better performance.

Anyway, the processing speed was effected not only by the GPU model specs. Depending from which side the needle was inserted, different amount of time was necessary for the MIP compression. It was recorded that MIP over the cross-sections required a considerable higher amount of time than the same process over the B-scans. The algorithm speed was measured 50 times for both the GPU models and on the two different MIP direction (Fig.4.28).

The average processing time recorded on the y-direction were respectively 1.589s and 2.576s on Tesla T4 and Tesla K80. The fastest run on Tesla T4 lasted 1.533s, the slower one 1.643s. The fastest run on Tesla K80 lasted 2,52s, the slower one 2,63s.

On the other hand, the average timing on the x-direction was 2.303s using Tesla T4, and 3.291s using Tesla K80. The fastest run on Tesla T4 lasted 2.289s, the slower one 2.317s. The fastest run on Tesla K80 lasted 3,277s, the slower one 3,305s.

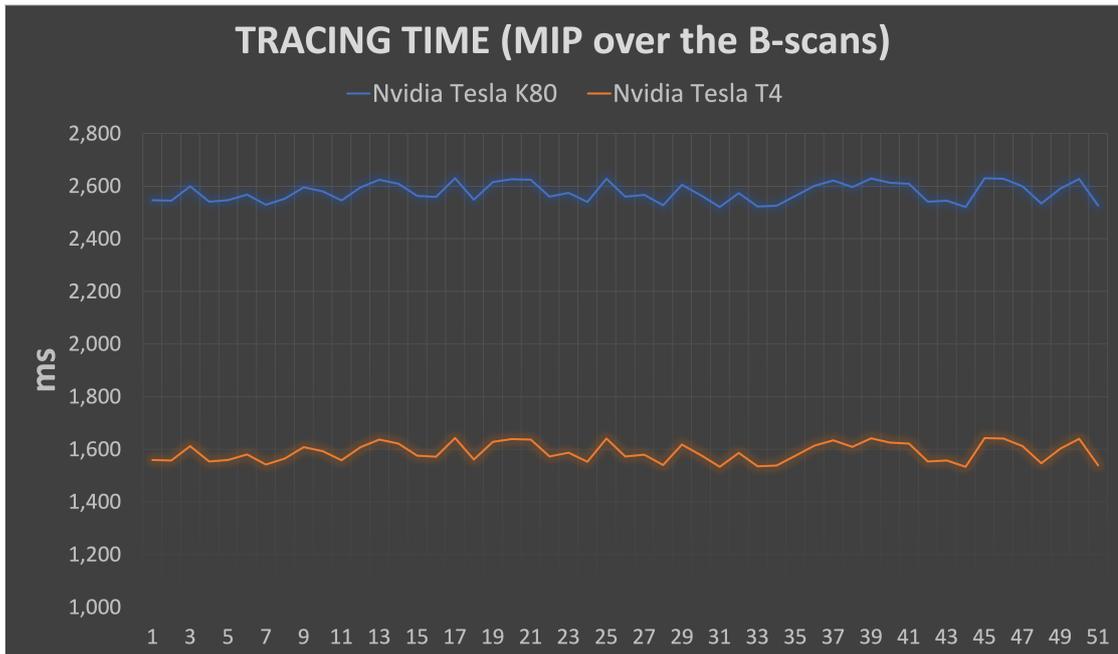


Figure 4.27: Tracing algorithm performed over the y-direction. The MIP was performed over the B-scans.

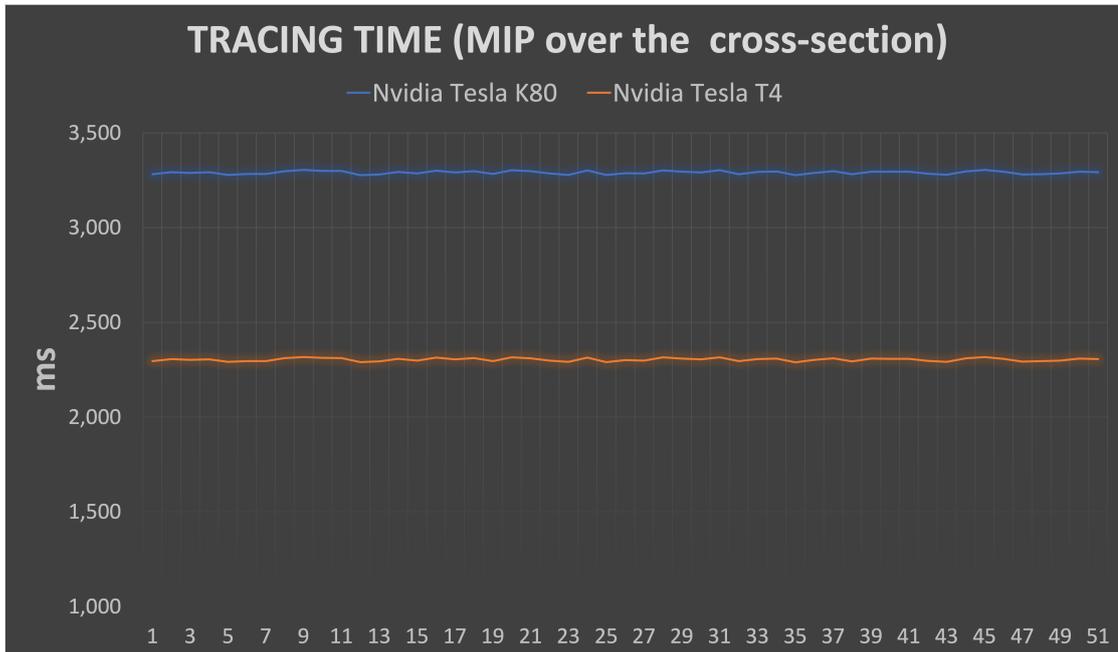


Figure 4.28: Tracing algorithm performed over the x-direction. The MIP was performed over the cross-sections.

Chapter 5

Conclusions

This work showed how the development of GPU-accelerated algorithms could make the OCT processing fast enough for real-time applications.

Using a Nvidia Tesla T4, it was possible reconstruct the volume and perform the needle tracing algorithm in $2,583 \pm 0,142s$. Translated in A-scan/s, in the worst timing ($2,952s$) more than 350×10^3 A-scan/s were processed, while more than 460×10^3 A-scan/s were processed in the best ($2,337$). The processing with a Nvidia Tesla K80 was significantly slower ($3,571 \pm 0,193s$), with $3,939s$ (more than 260×10^3 A-scan/s) and $3,324s$ (more than 315^3 A-scan/s).

The recorded timings are comparable with high-speed OCT swept lasers (sweep rate from tens to hundreds of KHz). Thus, it is reasonable to think this algorithm appropriate for a real-time applications as long as the swept source sweep rate is less than or equal to 300KHz (Nvidia Tesla T4) or 200KHz (Nvidia Tesla K80). The tracing algorithm presented a good needle detection in almost 90% of the simulation positions. Nevertheless, the tracing accuracy of the needle tip position was not so precise. Thus, the distance from the inclusion and the depth measure could provide only a rough spatial information of the needle, but they could not be taken as accurate.

Overall, the feasibility of this algorithm for a real-time application was verified. It showed good needle detection performance, but the tracing accuracy of its tip must be perfected. Anyway, this work can be a good and optimistic starting point for future improvements.

5.1 Limitations

The huge amount of high-speed computation needed for this work, required the usage of a high-performance GPU. Google Colab offered a very useful and well

designed development environment with considerable GPU performance. However, the limited RAM size (12Gbytes) prevented the possibility to upload and process on the platform more than one interference volume at a time. For the same reason, it was not possible to upload interference volume with more than 1280 axial points or the background signal for the DC removal.

From 2.5 to 4 times higher registration times were recorded when the volume was registered over the cross-sections and not the B-scans. The registration side is determined by the inclination of the scan lens. Thus, faster processing can be obtained by inclining the lens in order to force the registration over the B-scans (tilted cross-section). This constrain can be viewed as a limitation since it could be useful to position the sample with inclined B-scans to improve the image quality.

For what it might concerns the needle detection, the algorithm showed speed comparable to high-speed OCT, but only if MIP compression was performed over the B-scans. In fact, squeezing the volume over the cross-sections lead to higher processing time up to 1.45 times slower than MIP over the B-scans. Thus, the needle insertion side is limited to the xz -plane if the maximum speed is required.

The tracing algorithm presented a good needle detection in almost 90% of the simulation positions. Nevertheless, the tracing accuracy of the needle tip position was not so precise. Thus, the distance from the inclusion and the depth measure could provide only a rough spatial information of the needle, but they could not be taken as accurate.

The black plastic inclusions were useful to simulate the biopsy area, but sometimes portions of this area were labeled as needle section by the network. When this happened in a frame with a needle section, when the mask of the frame was generated, the correct labeled area of the needle could be removed but not the inclusion area. This usually happened if just a small portion of the needle was inserted and the wrong labeled area was bigger than the needle section. Different times, the inclusion was segmented partially or not at all. Since the inclusion was one of the label with the best IoU score, it was not clear why that happened.

In this work, only 20G needles were used. Because of it, at least one section of the needle could be found if the compressed volume was scanned every 8 B-scans. For this reasoning, if the needle had higher Gauge factor (smaller diameter), this algorithm could not find the needle.

5.2 Challenges and future work

U-net showed great performance despite the really small dimension of the database. But the needle tip tracing accuracy depends on the segmentation performance of this FCN on the needle label, which was the label with the worst IoU score. This work wanted to show the feasibility of this concept, which was verified, but the database size must be enlarged with a huge amount of new images for a future work. This will probably increase the prediction performance for the needle label.

The images in the database represent principally situation where the needle was positioned alongside the B-scan and the cross-section axis. In all the simulations where the needle was inserted with a significant xz -angle, the needle was well detected and its tip was traced with an error lower than $0.5mm$ in all the directions. However, two simulations are not enough to say if the algorithm can trace correctly the needle in these situations. Thus, another challenge for a future work is to untying the algorithm performance from the direction of the needle insertion. The network must be trained to segment the needle also when images represent a non perfectly sagittal section of it.

The algorithm speed may be not sufficient for real-time applications that use a swept source sweep rate greater than 300KHz. In a future work, the reduction of the processing time will be another hot topic. The CuPy library allowed to exploit the GPU acceleration without writa a line of code the C++. Nevertheless, custom CUDA kernels can probably optimize all the processing steps and speed up the algorithm.

Once all this refinements will be done, there will be the necessity to test the algorithm on a local runtime and with local GPU. That's the last task to achieve before start thinking on the final goal: the direct implementation on an OCT system.

Bibliography

- [1] Christoph K Hitzenberger. «Optical measurement of the axial eye length by laser Doppler interferometry.» In: *Investigative ophthalmology & visual science* 32.3 (1991), pp. 616–624 (cit. on p. 2).
- [2] Gerd Hausler, Michael Walter Lindner, et al. «" Coherence radar" and" spectral radar"-new tools for dermatological diagnosis». In: *Journal of biomedical optics* 3.1 (1998), pp. 21–31 (cit. on p. 2).
- [3] A F Fercher, W Drexler, C K Hitzenberger, and T Lasser. «Optical coherence tomography - principles and applications». In: *Reports on Progress in Physics* 66.2 (Jan. 2003), pp. 239–303. DOI: 10.1088/0034-4885/66/2/204. URL: <https://doi.org/10.1088/0034-4885/66/2/204> (cit. on pp. 3–14, 30).
- [4] C. K. Hitzenberger, M. Danner, W. Drexler, and A. F. Fercher. «Measurement of the spatial coherence of superluminescent diodes». In: *Journal of Modern Optics* 46.12 (1999), pp. 1763–1774. DOI: 10.1080/09500349908231370. eprint: <https://www.tandfonline.com/doi/pdf/10.1080/09500349908231370>. URL: <https://www.tandfonline.com/doi/abs/10.1080/09500349908231370> (cit. on p. 4).
- [5] 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* 19.7 (2014), pp. 1–34. DOI: 10.1117/1.JBO.19.7.071412. URL: <https://doi.org/10.1117/1.JBO.19.7.071412> (cit. on pp. 11, 12).
- [6] Murtaza Ali and Renuka Parlapalli. «Signal Processing Overview of Optical Coherence Tomography Systems for Medical Imaging». In: Texas Instrum. (Jan. 2010) (cit. on p. 14).
- [7] Wolfgang Wieser, Benjamin R Biedermann, Thomas Klein, Christoph M Eigenwillig, and Robert Huber. «Multi-megahertz OCT: High quality 3D imaging at 20 million A-scans and 4.5 GVoxels per second». In: *Optics express* 18.14 (2010), pp. 14685–14704 (cit. on p. 14).

- [8] Zhe Chen, Mengyang Liu, Michael Minneman, Laurin Ginner, Erich Hoover, Harald Sattmann, Marco Bonesi, Wolfgang Drexler, and Rainer A. Leitgeb. «Phase-stable swept source OCT angiography in human skin using an akinetic source». In: *Biomed. Opt. Express* 7.8 (Aug. 2016), pp. 3032–3048. DOI: 10.1364/BOE.7.003032. URL: <http://www.osapublishing.org/boe/abstract.cfm?URI=boe-7-8-3032> (cit. on pp. 15, 21).
- [9] Yuuki Watanabe and Toshiki Itagaki. «Real-time display on Fourier domain optical coherence tomography system using a graphics processing unit». In: *Journal of Biomedical Optics* 14.6 (2009), p. 060506 (cit. on pp. 20, 21).
- [10] Janarthanan Rasakanthan, Kate Sugden, and Peter H Tomlins. «Processing and rendering of Fourier domain optical coherence tomography images at a line rate over 524 kHz using a graphics processing unit». In: *Journal of biomedical optics* 16.2 (2011), p. 020505 (cit. on p. 20).
- [11] Kang Zhang and Jin U Kang. «Graphics processing unit accelerated non-uniform fast Fourier transform for ultrahigh-speed, real-time Fourier-domain OCT». In: *Optics express* 18.22 (2010), pp. 23472–23487 (cit. on p. 20).
- [12] Sam Van der Jeught, Adrian Bradu, and Adrian Gh Podoleanu. «Real-time resampling in Fourier domain optical coherence tomography using a graphics processing unit». In: *Journal of Biomedical Optics* 15.3 (2010), p. 030511 (cit. on p. 21).
- [13] Ruikang K Wang and Zhenhe Ma. «A practical approach to eliminate autocorrelation artefacts for volume-rate spectral domain optical coherence tomography». In: *Physics in Medicine & Biology* 51.12 (2006), p. 3231 (cit. on p. 22).
- [14] Yu Chen, Jeff Fingler, and Scott E. Fraser. «Multi-shaping technique reduces sidelobe magnitude in optical coherence tomography». In: *Biomed. Opt. Express* 8.11 (Nov. 2017), pp. 5267–5281. DOI: 10.1364/BOE.8.005267. URL: <http://www.osapublishing.org/boe/abstract.cfm?URI=boe-8-11-5267> (cit. on p. 23).
- [15] Joseph Ho et al. «Clinical assessment of mirror artifacts in spectral-domain optical coherence tomography». In: *Investigative ophthalmology & visual science* 51.7 (2010), pp. 3714–3720 (cit. on p. 23).
- [16] Brian W Pogue and Michael S Patterson. «Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry». In: *Journal of biomedical optics* 11.4 (2006), p. 041102 (cit. on p. 30).

- [17] Guy Lamouche, Brendan F Kennedy, Kelsey M Kennedy, Charles-Etienne Bisailon, Andrea Curatolo, Gord Campbell, Valérie Pazos, and David D Sampson. «Review of tissue simulating phantoms with controllable optical, mechanical and structural properties for use in optical coherence tomography». In: *Biomedical optics express* 3.6 (2012), pp. 1381–1398 (cit. on p. 30).
- [18] Theodore P Moffitt, Yin-Chu Chen, and Scott A Prahl. «Preparation and characterization of polyurethane optical phantoms». In: *Journal of Biomedical Optics* 11.4 (2006), p. 041103 (cit. on p. 30).
- [19] Maciej S Wróbel, Alexey P Popov, Alexander V Bykov, Valery V Tuchin, and Małgorzata Jędrzejewska-Szczerska. «Nanoparticle-free tissue-mimicking phantoms with intrinsic scattering». In: *Biomedical optics express* 7.6 (2016), pp. 2088–2094 (cit. on p. 30).
- [20] Frank P Bolin, Luther E Preuss, Roy C Taylor, and Robert J Ference. «Refractive index of some mammalian tissues using a fiber optic cladding method». In: *Applied optics* 28.12 (1989), pp. 2297–2303 (cit. on p. 30).
- [21] Jianwei Fu, Guotao Quan, and Hui Gong. «A simple method for prediction of the reduced scattering coefficient in tissue-simulating phantoms». In: *Journal of Innovative Optical Health Sciences* 3.01 (2010), pp. 53–59 (cit. on p. 31).
- [22] Charles-Etienne Bisailon, Guy Lamouche, Romain Maciejko, Marc Dufour, and Jean-Pierre Monchalain. «Deformable and durable phantoms with controlled density of scatterers». In: *Physics in Medicine & Biology* 53.13 (2008), N237 (cit. on pp. 31, 32).
- [23] Daniel Martijn M de Bruin, Rolf H Bremmer, Vitali M Kodach, Roy de Kinkelder, Jan van Marle, Ton G van Leeuwen, and Dirk J Faber. «Optical phantoms of varying geometry based on thin building blocks with controlled optical properties». In: *Journal of biomedical optics* 15.2 (2010), p. 025001 (cit. on p. 31).
- [24] Hubert Monnier, A-M Wilhelm, and H Delmas. «Influence of ultrasound on mixing on the molecular scale for water and viscous liquids». In: *Ultrasonics Sonochemistry* 6.1-2 (1999), pp. 67–74 (cit. on p. 32).
- [25] Julia Welzel, Eva Lankenau, Reginald Birngruber, and Ralf Engelhardt. «Optical coherence tomography of the human skin». In: *Journal of the American Academy of Dermatology* 37.6 (1997), pp. 958–963 (cit. on p. 32).
- [26] M Ulrich, E Stockfleth, J Roewert-Huber, and S Astner. «Noninvasive diagnostic tools for nonmelanoma skin cancer». In: *British Journal of Dermatology* 157 (2007), pp. 56–58 (cit. on p. 33).

- [27] Thilo Gambichler, Philipp Regeniter, Falk G Bechara, Alexej Orlikov, Remus Vasa, Georg Moussa, Markus Stücker, Peter Altmeyer, and Klaus Hoffmann. «Characterization of benign and malignant melanocytic skin lesions using optical coherence tomography in vivo». In: *Journal of the American Academy of Dermatology* 57.4 (2007), pp. 629–637 (cit. on p. 33).
- [28] Mette Mogensen, Lars Thrane, Thomas Martini Jørgensen, Peter E Andersen, and GBE Jemec. «Optical coherence tomography for imaging of skin and skin diseases». In: *Seminars in cutaneous medicine and surgery*. Vol. 28. 3. WB Saunders. 2009, pp. 196–202 (cit. on p. 33).
- [29] Mette Mogensen, Thomas Martini Joergensen, Birgit Meincke Nürnberg, Hanan Ahmad Morsy, Jakob B Thomsen, Lars Thrane, and Gregor BE Jemec. «Assessment of optical coherence tomography imaging in the diagnosis of non-melanoma skin cancer and benign lesions versus normal skin: observer-blinded evaluation by dermatologists and pathologists». In: *Dermatologic Surgery* 35.6 (2009), pp. 965–972 (cit. on p. 33).
- [30] Mette Mogensen, Lars Thrane, Thomas M Jørgensen, Peter E Andersen, and Gregor BE Jemec. «OCT imaging of skin cancer and other dermatological diseases». In: *Journal of biophotonics* 2.6-7 (2009), pp. 442–451 (cit. on p. 33).
- [31] Brenton Keller, Mark Draelos, Gao Tang, Sina Farsiu, Anthony N Kuo, Kris Hauser, and Joseph A Izatt. «Real-time corneal segmentation and 3D needle tracking in intrasurgical OCT». In: *Biomedical optics express* 9.6 (2018), pp. 2716–2732 (cit. on pp. 36, 49).
- [32] Mingchuan Zhou, Hessam Roodaki, Abouzar Eslami, Guang Chen, Kai Huang, Mathias Maier, Chris P Lohmann, Alois Knoll, and Mohammad Ali Nasseri. «Needle segmentation in volumetric optical coherence tomography images for ophthalmic microsurgery». In: *Applied Sciences* 7.8 (2017), p. 748 (cit. on pp. 36, 43).
- [33] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. «U-net: Convolutional networks for biomedical image segmentation». In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241 (cit. on pp. 36, 41, 42).
- [34] Andrea Veltri, Irene Bargellini, Luigi Giorgi, Paulo Alexandre Matos Silva Almeida, and Okan Akhan. «CIRSE guidelines on percutaneous needle biopsy (PNB)». In: *Cardiovascular and interventional radiology* 40.10 (2017), pp. 1501–1513 (cit. on p. 36).
- [35] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. «Understanding of a convolutional neural network». In: *2017 International Conference on Engineering and Technology (ICET)*. Ieee. 2017, pp. 1–6 (cit. on pp. 37–40).

- [36] Jonathan Long, Evan Shelhamer, and Trevor Darrell. «Fully convolutional networks for semantic segmentation». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 3431–3440 (cit. on pp. 40, 41).
- [37] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. «Backpropagation applied to handwritten zip code recognition». In: *Neural computation* 1.4 (1989), pp. 541–551 (cit. on p. 41).
- [38] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. «Imagenet classification with deep convolutional neural networks». In: *Advances in neural information processing systems* 25 (2012), pp. 1097–1105 (cit. on p. 41).
- [39] Karen Simonyan and Andrew Zisserman. «Very deep convolutional networks for large-scale image recognition». In: *arXiv preprint arXiv:1409.1556* (2014) (cit. on p. 41).
- [40] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. «Going deeper with convolutions». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9 (cit. on p. 41).
- [41] Hiroyuki Saigusa, Yoshihiro Ueda, Akihiro Yamada, Masato Ohmi, Makoto Ohnishi, Mitsuo Kuwabara, and Masamitsu Haruna. «Maximum-intensity-projection imaging for dynamic analysis of mental sweating by optical coherence tomography». In: *Applied physics express* 1.9 (2008), p. 098001 (cit. on p. 43).
- [42] Zhilu Zhang and Mert R Sabuncu. «Generalized cross entropy loss for training deep neural networks with noisy labels». In: *32nd Conference on Neural Information Processing Systems (NeurIPS)*. 2018 (cit. on p. 44).
- [43] Seyed Sadegh Mohseni Salehi, Deniz Erdogmus, and Ali Gholipour. «Tversky loss function for image segmentation using 3D fully convolutional deep networks». In: *International workshop on machine learning in medical imaging*. Springer. 2017, pp. 379–387 (cit. on p. 44).
- [44] Lu Wang, Chaoli Wang, Zhanquan Sun, and Sheng Chen. «An improved dice loss for pneumothorax segmentation by mining the information of negative areas». In: *IEEE Access* 8 (2020), pp. 167939–167949 (cit. on p. 44).
- [45] Hafiz Zia Ur Rehman and Sungon Lee. «Automatic image alignment using principal component analysis». In: *IEEE Access* 6 (2018), pp. 72063–72072 (cit. on p. 53).