



**Politecnico
di Torino**

POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

**3D Reconstruction in photoacoustic imaging
assisted by deep learning**



Supervisor

Dr. Bastien ARNAL

Candidate

Ivana FALCO

Cosupervisor

Prof. Kristen MEIBURGER

OCTOBER 2021

Abstract

Photoacoustic imaging (PA) is an emerging biomedical modality consisting in the emission of pulsed laser light, which, when absorbed by the tissue components, generates ultrasound waves. After reception, the PA signals are used to provide an image by a reconstruction algorithm. In the vast field of biomedical imaging, this modality is extremely promising, as it permits to image tissue optical properties at greater depths than other optical imaging modalities, such as optical coherence tomography, with interesting resolution and can provide images of optical absorption with specific molecular contrast which can be enhanced by spectroscopy. In particular, the omnipresence of hemoglobin in living tissues allows the imaging of microvasculature without employing any exogenous contrast agent, which is one of the most important uses of photoacoustic imaging.

However, conventional PA imaging systems are limited by low contrast and visibility artefacts that arise from coherence of PA waves and characteristics of the detection system, such as geometry (limited-view) and frequency bandwidth. Limited bandwidth artefacts occur when the central part of the reconstructed object is not visible because ultrasound detectors filter out low-frequency components of PA waves emitted by large absorbers (large as compared to the detection wavelength range). Limited-view artefacts occur when the coherent acoustic waves that are directional cannot be measured if they don't reach the probe.

A first dynamic technique to solve these visibility problems is the photoacoustic fluctuation imaging that exploits the natural fluctuation of the blood flow to reconstruct the total visibility of the vessels. It was shown on a 2D imaging system that the reconstruction quality could be also enhanced thanks to a deep learning algorithm trained on simulated and experimental data. This approach permits to correct artefacts from a single acquisition, responding to the limit of the fluctuation imaging: the acquisition time. The main objective of this thesis is to transfer these results to a 3D environment by elaborating a

neural network able to process 3D volumes.

In the first part of this project, we focus on programming a user-friendly and reliable real-time visualization and acquisition system to collect time signals from a chicken embryo model. Its chorioallantoic membrane is, in fact, an optimal *in vivo* model to study blood vessels thanks to its visibility, accessibility and rapid developmental growth. From acquired radio-frequency signals, we aim at developing a procedure consisting in the delay and sum beamforming method followed by the singular value decomposition method to create an experimental data-set (1500 volumes per acquisition).

Our deep learning approach, using a convolution network based on the UNET architecture, includes the pre-training the network with simulated data and a subsequent training with the experimental data to find a model that permits, from inputs never trained and not known by the network, to predict an output free of artefacts. Noisy, corrupted by artifacts and altered volumes of the chicken embryo vasculature are eliminated to get a final data-set composed by 105 examples: 90 for the training-set, 10 for the validation-set and 5 for the test-set. Among all the conventional PA volumes, just one of them is chosen as input, while the respective PA fluctuation volume (computed on 1500 conventional volumes, free of its average to eliminate the background noise) is used as ground truth.

The correctness and the robustness of this prediction is finally verified by a correlation measure. By testing our model on the training-set, we find a high index of correlation (greater than 84%): the model learns to correct the artifacts and to extrapolate some structures not visible in conventional PA images. On the other hand, the predictions on the test-set present a lower correlation value, some artefacts, like not existent structures that appear on the output and a poorly defined microvasculature that was well resolved in the fluctuation image. Despite this, the network is able to completely eliminate the noise of the conventional image, to greatly increase the contrast and to reconstruct the volume of the vessels.

Acknowledgements I

This work has been possible thanks to the great deal of support I have received during my internship at LiPhy (Laboratoire Interdisciplinaire de Physique, Université de Grenoble-Alpes) and to Politecnico di Torino that gave me this opportunity.

I would like to thank my supervisor Bastien Arnal, for supporting and guiding me through each stage of our work. Thank you for conveying to me your passion and dedication to this project that I hope to continue in the best way. Your personal and technical qualities have enriched my career and my life.

Kind thanks to Guillame, Adrien, Camille and the whole Optima team for welcoming me and for making me feel part of the group from the first day.

Finally, I would like to acknowledge my co-supervisor, Prof. Kristen Meiburger, for her kindness, her help and her precious advice in following this work.

Acknowledgements II

A Torino, il mio punto di partenza. La città che mi ha fatto vivere le giornate più burrascose della mia vita e allo stesso tempo mi ha dato portici per coprimi. Che per ogni esame mi ha invitato ad una festa o ad un concerto per berci su. Dove, dopo ogni “Non ce la faccio”, alla fine ce l’ho fatta con accanto le migliori amiche e coinquiline del mondo.

Proprio una di loro, prima di andar via, mi ha scritto: “Parti. Torna con idee nuove.”.

L’ho fatto. Questi due anni, per me, sono stati la prova di quanto sia bello cambiarle, le proprie idee. E poi cambiarle ancora. Ed aggiungerne di nuove. E dire “Ma dove sono finita?”, “Che sfiga, la Francia. . .” e poi innamorarmene.

Innamorarmene grazie a Compiègne e ai miei amici della “Resistenza” che l’hanno vissuta e scoperta insieme a me. Perché nel periodo più buio, io su quel fiume e dentro me stessa ho visto un’altra luce. Perché in quel miscuglio di nazionalità e lingue, gnocchi a fine mese e ricordi di feste composti come un puzzle il mattino dopo, ho imparato a vivere. Ed ho imparato a vivere seguendo quella sorta di *carpe diem* francese, che loro devono tradurre proprio tutto: “bisogna approfittarne”, “il faut profiter”.

Questi due anni sono stati la prova di quanto sia bello posare lo sguardo su orizzonti differenti. E quando l’orizzonte non l’ho più visto ero appena scesa dal treno per Grenoble e mi sono ritrovata in mezzo alle montagne per l’ultimo atto, quello della tesi, della passione per il mio progetto, della prima stretta di mano al futuro. Grazie alla mia “Little Italy” ho vissuto una vera e propria *randonnée* di sei mesi ed ora ho proprio voglia di restare con loro in cima, per godermi la vista.

D'altra parte, il bello del cambiare idea risiede anche nell'accorgersi, dopo tanto tempo, che alcune di esse sono rimaste invariate. Se sono arrivata fin qui è anche grazie alle mie "idee fisse".

Alla mia Calabria, con quelli di sempre, con chi sa aspettarmi ed aspetta anche l'alba insieme a me.

A Calopezzati, ai suoi paesaggi che conosco a memoria e che mi stupiscono ancora.

Alle volte in cui il mio nervosismo è stato scacciato da una passeggiata "tranquilla" con Runa.

Ai miei zii e i miei cugini.

A Christian, Gaetano e la nonna Elena.

A mamma, papà e Marilena, a cui devo tutto.

Voi non apparterrete mai ad un solo posto perché abitate in ogni stanza del mio cuore.

*Ad Alex,
la cui assenza fa da linea-guida ai miei obiettivi,
diventando presenza.*

Contents

List of Figures	VIII
1 INTRODUCTION	1
1.1 History of photoacoustic effect	1
1.2 Description of the photoacoustic effect	5
1.2.1 Photoacoustic generation and physical description	5
1.2.2 Photoacoustic imaging	6
1.3 Visibility problems	8
1.3.1 Origin of visibility problems	8
1.3.2 State of the art	9
2 PHOTOACOUSTIC FLUCTUATION IMAGING	13
2.1 Methodology	13
2.1.1 Experimental device - Hardware components	13
2.1.2 Vasculature model	17
2.1.3 Software	20
2.2 Photoacoustic image reconstruction	30
2.2.1 Delay and sum beamforming	30
2.3 Photoacoustic fluctuation imaging	32
2.3.1 Principles and theory	33
2.3.2 Noise sensitivity	34
2.3.3 Parasitic fluctuations and SVD	35
2.4 Results	38
3 DEEP LEARNING APPROACH	44
3.1 Introduction	44

3.1.1	State of the art - 2D results	47
3.2	Methodology	48
3.2.1	Network architecture	48
3.2.2	Network training	49
3.3	Network performances - Results	51
4	CONCLUSIONS AND FUTURE EXPERIMENTS	56
	Bibliography	57

List of Figures

1.1	Photophone designed and built by Alexander G. Bell and Charles S. Tainter [1]	2
1.2	Time-line of various significant events in the history of photoacoustics from 1880 toward the modern times. Red arrows indicate the low frequency intensity modulated variant; blue arrows indicate the short pulsed variant. [1]	4
1.3	The basic principle of photoacoustic imaging.	5
1.4	Spectrum of the absorption coefficient for different components of the tissue. [12]	7
1.5	The two types of visibility artefacts. We can note that in the reconstructed object we lost (a) the vertical part of the blood vessel because of the limited view, (b) the central part of the body vessel because of the limited bandwidth.	10
2.1	Experimental Set-Up	13
2.2	3D imaging probe.	15
2.3	Signals measured with oscillator. a : Photodiode signal for a shot (green) and gated integrator (blue). b : Result of integration for about ten impulses. This signal is then converted by the DAQ.	16
2.4	Plot of the normalized signals coming from the PD (in blue) and the pyroelectric detector (in orange). The zoom on the plot at each of the six wavelengths (x-axis: 300 laser pulses for [700 730 760 790 820 850] nm) allows to compare them and validate the synchronisation.	17

2.5	Chicken embryo and associated structures. a) At day 4 post-fertilization CAM and YSM are highly vascularized but distinguishable (YSM is in the background). b) At day 12 post-fertilization in a petri dish: YSM is under the embryo and it is only associated with the yolk. [24]	18
2.6	Embryo vasculature at day 10 after incubation.	18
2.7	Imaged embryo placed in a plastic container with a central hole containing water. The container was filled with water maintained at 37°C thanks to a hot plate.	19
2.8	Main interface of our acquisition software where the buttons LiveMode, DisplayRF, USorUSHR,MIPPA, MIPUS,Save, DispPD and the sliders XZ,YZ,XY and LenscorrPA represent the custom functions that we added.	21
2.9	ReceiveBuffer composition for LiveMode. M is equal to the number of samples for US and USHR acquisitions and P is equal to the number of samples for PA acquisition. na and na2 are the numbers of flash angles for USHR and US acquisitions respectively.	22
2.10	Real time reconstruction of the imaged object.	24
2.11	Interface of the volumic reconstruction in Live Mode. The red circle underlines the button used to display the MIPs.	25
2.12	Interface of radio-frequency signals control. The button DisplayRF permits to show the max. of RF signals to avoid saturation. In can be decrease by modifying the TGC value.	25
2.13	Sequence of an acquisition US-PA with $M_\lambda = 6$ wavelengths, $PRF_{US} = 200$ Hz, $PRF_{PA} = 100$ Hz and $na2 = 10$ (number of emissions to form an US image.	27

2.14	ReceiveBuffer composition for SaveMode. M is equal to the number of samples for US and USHR acquisitions and P is equal to the number of samples for PA acquisition. na and $na2$ are the numbers of flash angles for USHR and US acquisitions respectively. nPA and nUS are the number of acquisitions for PA and US ($nPA = 50$, $nUS = 100$). The number of frames used was 30.	27
2.15	Display of photodiode signal trend after acquisition.	28
2.16	ReceiveBuffer composition for SaveMode without US storage.	29
2.17	Summary of the different steps to obtain the final photoacoustic image. . .	30
2.18	A typical radio-frequency (RF) frame available at the output of the acquisition machine as a result of PA acquisition with multielement linear array. [25]	31
2.19	Delay-and-sum beamforming applied to a $N=50$ RF data to obtain the conventional PA image (displayed as MIPs).	32
2.20	In the left section (RF signals domain) a peak P of the interface at the entrance and exit of the vessel and a fluctuation level F can be defined. When a noise n is added, it affects the entire signal, so the fluctuation F must be extrapolated from $\sqrt{F^2 + n^2}$. In the right part of the figure (Image domain), on the average image there are no fluctuations while the fluctuation image present the level of fluctuation increased by the noise level. [29]	34
2.21	The upper section of the figure was obtained thanks to a simulation set. The first line shows a fluctuation to noise ratio (FNR) equal to 0.2 (noise louder than the fluctuation). With enough images, a good reconstructed image is obtained. When the FNR increases, fewer images needed to obtain a comparable image quality. In experience a FNR of 0.8 has been founded. [29]	35

2.22	Fluctuation image $\sigma[A](r)$ dominated by the average image $ \langle A_k(r) \rangle $, before SVD.	36
2.23	Fluctuation image after SVD, free of artefacts.	37
2.24	Result 1. The fluctuation PA image points out a big vertical vessel and a really complex vascular net that were absolutely not visible in the conventional PA image.	39
2.25	Result 2. The "Limited bandwidth" problem has been properly corrected. Blue flashes reveal a vessel whose just upper and lower edges are reconstructed, while green flashes show the vessel successfully reconstructed in its inner part.	39
2.26	Result 3. As in Fig.2.25, the internal part of the object is well reconstructed.	40
2.27	Result 4. The "Limited view" problem has been solved too. Blues flashes point the absent vertical structure which became visible and well defined after SVD.	40
2.28	Result 5. As in Fig.2.27, green flashes show the correction of the visibility problem that did not allowed to see the entire structure.	41
2.29	Result 6. Error in reconstruction: absorbing structure pointed by the blue flash that disappeared in fluctuation PA image.	41
2.30	Result 7. Presence of bright point that absorb more than the other important structures and do not allow to have a good visibility of them.	42
2.31	Result 8. 3D Visualisation	43
3.1	Perceptron architecture and process. X is the input vector, the W values represent the weights. An integration and a linear activate function are applied on the product between X and W to obtain the output value. . . .	45
3.2	Forward propagation process.	46
3.3	Backward propagation process.	46
3.4	Max pooling operation applied on feature maps to decrease the number of data and accelerate the computation time. [30]	48

3.5	Example of deep learning application on simulated data: the first column represent the MIPs of the conventional image are showed, the second column the MIPs of the simulated ground truth and the third one shows the results of the accurate deep learning prediction. [<i>G.Godefroy</i>]	50
3.6	Result 1. MIPs of the input, ground truth et output on the three axes. Correlation between ground truth and output: 86%.	52
3.7	Result 2. Correlation between ground truth and output: 89%.	52
3.8	Result 3. Correlation between ground truth and output: 87%.	53
3.9	Result 4. Correlation between ground truth and output: 84 %.	53
3.10	Result 1. Prediction of the network tested on an example not known (out of the training set). We note a contrast enhancement and a good reconstruction of the internal part of the vessel (blues fleshes in fluctuation image and green fleshes in predicted image). Therefore, compared to the ground truth, there are some not existent structures that appear on the output (surrounded by the yellow circle).	54
3.11	Result 2. As in the previous example, vessels are well reconstructed (blue and green fleshes) but the predicted image presents a poorly defined microvasculature that was well resolved in the fluctuation image (yellow circle).	55

1. INTRODUCTION

1.1 History of photoacoustic effect

The original photoacoustic effect was observed for the first time in 1880 when Alexander Bell and his assistant Charles Sumner Tainter, transmitted an acoustic signal for 213 meters through a device that they called “Photophone”. The most successful apparatus consisted of a transmitter, which was a thin silvered glass disk mounted on a frame with a flexible rubber hose whose free end comprised the mouthpiece. Sunlight was focused on the mirror and the associated optics so arranged that the reflected light could be collected by the remotely placed receiver—a parabolic mirror at whose focus a selenium cell incorporated into a conventional telephone circuit was arranged (Fig.1.1). [1]

Speech articulated into the mouthpiece caused the mirror to vibrate, producing fluctuations in the intensity of light collected at the receiver. The voice-modulated intensity of light available at the receiver as a modulated battery current was converted into sound in the telephone circuit. The selenium-cell photophone can be considered the first practical implementation of wireless telephony or, in fact, the first optical communication device. In addition to this innovative concept, there was a further important outcome of this investigation by Bell, and that was his discovery that illumination of different solid substances with a rapidly interrupted beam of light energy resulted in the emission of acoustic energy at the same frequency as the modulation frequency. Bell’s results were revolutionary for that period and they caused great excitement in the scientific community; many scientists began to be interested in both experimental and theoretical study of photoacoustic effect.

- Lord Rayleigh concluded in 1881 [2] that the explanation for the sounds was a vibration due to an unequal heating of the diaphragms or plates when illuminated with intermittent light.
- Mercadier and Preece [3] studied the effect in a variety of materials and postulated

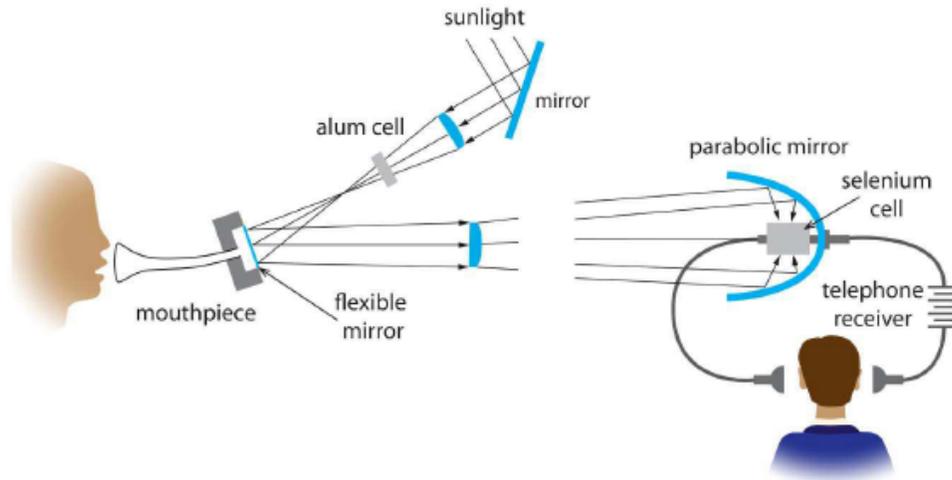


Figure 1.1: Photophone designed and built by Alexander G. Bell and Charles S. Tainter [1]

that the cause of the sounds was not the vibration of the disk, but rather the expansion and contraction of air in contact with the intermittently heated disk.

- Röntgen [4] and Tyndall [5] independently also shared the view that periodic heating and cooling of air in thermal contact with the disk caused the sound production. They also showed that the phenomenon was not restricted to solid bodies, and studied the effect in absorbing gases and vapors.

After the initial flurry of excitement and activity, the radiophonic effect was set aside, principally because experiments were difficult to quantify, as they depended largely on human hearing for signal detection. [1] It waited for five decades before making a re-emergence with the work of Veingerov [7], who used state-of-the art charged capacitive microphone diaphragms and a Nernst glower as an intense blackbody infrared source, to be able to detect CO_2 concentrations in N_2 with a sensibility of 0.2 vol.%. A year later, Pfund [8] reported on a similar system in use at the Johns Hopkins Hospital for detection of CO and CO_2 gases. Instead of observing pressure changes, this system directly measured the corresponding changes of gas temperature using a thermopile shielded from direct radiation,

avoiding the requirement for acoustic noise isolation. In 1938 Luft introduced a commercial infrared gas analyzer based on microphone detectors. A theoretical foundation was developed for the optico-acoustic effect, and various implementations soon followed. In addition to gas concentration analysis, the method was also used in the investigation of vibrational relaxation rates and molecular energy transfer in gases. [9] With the development of the laser in 1960, there was yet another revival in gas analysis, thanks to its important advantages as light source came from the high degree of spectral purity, high stability, and reproducibility.

First approaches to biomedical applications

Allan Rosencwaig coined the term “Photoacoustic” and showed photoacoustic spectra from various solid materials and biological materials such as cytochrome and hemoglobin. The first “in vivo” application was conducted in 1964 by Amar et al. [10] on the eyes of a living rabbit. The authors used a ruby laser that produced on average 50 mJ pulses of a very short duration in order to detect the response in the form of acoustic wave. In 80’s Olsen obtained the first 2D image of a muscle phantom with a radar transmitter to irradiate with pulsed 5.655 GHz energy, and used a hydrophone to record the acoustic responses. The acoustic waves were found to have pulse durations corresponding to the measured penetration depth of the microwave energy. Bowen in 1981 [11] was among the first to propose imaging of soft tissues using this method. All the cited discoveries, laser’s development and subsequent applications summarized in Fig.1.2 permitted an improvement in photoacoustic imaging and its effective, promising and innovative usage in different cases of diagnostics and medical researches.

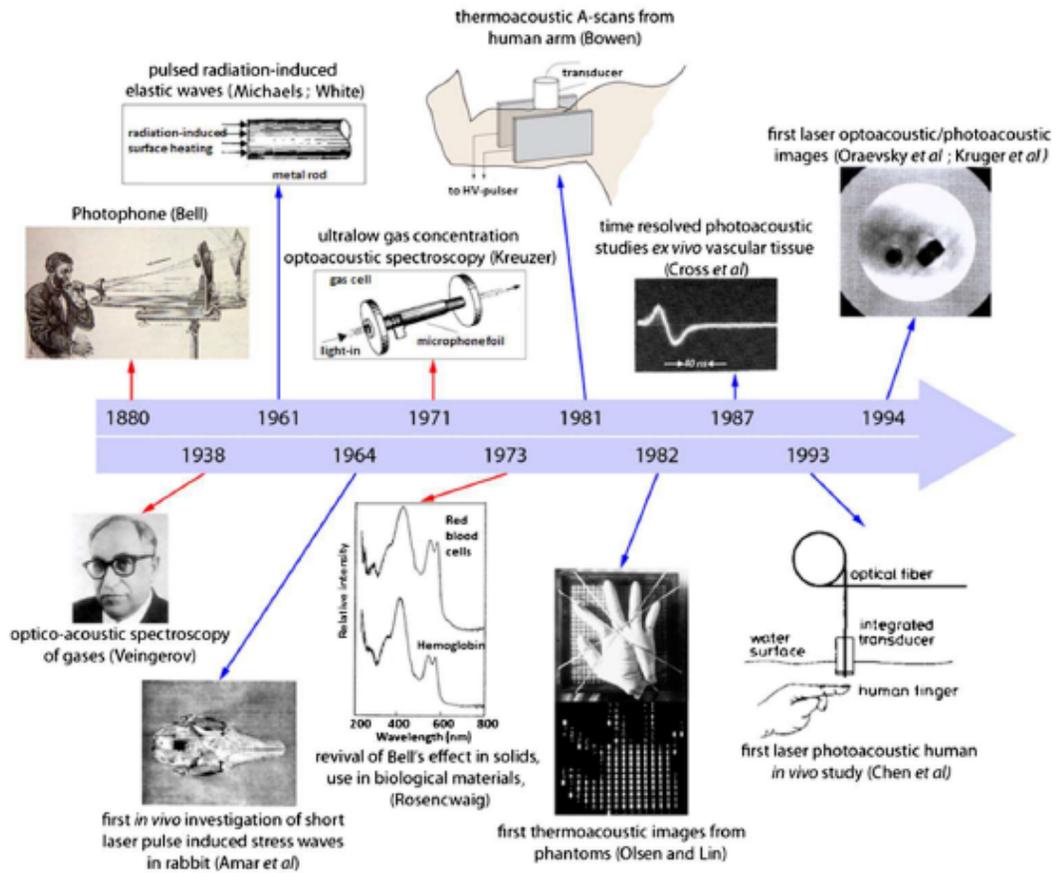


Figure 1.2: Time-line of various significant events in the history of photoacoustics from 1880 toward the modern times. Red arrows indicate the low frequency intensity modulated variant; blue arrows indicate the short pulsed variant. [1]

1.2 Description of the photoacoustic effect

1.2.1 Photoacoustic generation and physical description

In the photoacoustic effect, ultrasound waves are generated by a light-absorbing material following the absorption of modulated light, usually pulsed laser light on a nanosecond timescale. The absorbed light energy causes localized heating, which in turn produces a temperature rise. This increment gives rise to an initial pressure increase due to rapid thermal expansion, which is followed by relaxation, generating broadband low-amplitude ultrasound waves. The photoacoustic imaging is made possible thanks to detection of these ultrasound signals using an ultrasound transducer. This sequence of events is summarized in Fig.1.3.

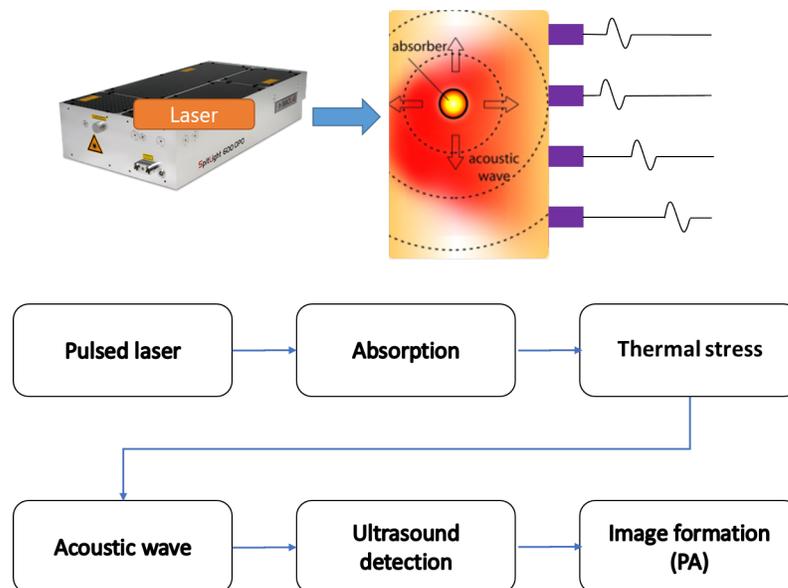


Figure 1.3: The basic principle of photoacoustic imaging.

If impulsive heating is assumed, and in practice, this requires that the acoustic propagation time is small compared with the length scale of the heated volume, then by thermodynamic considerations it can be shown that p_0 at a point r is proportional to the absorbed

optical energy $H(r)$.

$$p_0(r) = \Gamma H(r), \quad (1.1)$$

where:

- $\Gamma = \beta c^2 / C_p$ is the Grüneisen coefficient, a dimensionless thermodynamic constant that gives a measure of the conversion efficiency of heat energy to pressure with:
 - β : volume thermal expansivity
 - c : speed of sound
 - C_p : specific heat capacity at constant pressure
- $H(r)$ is the absorbed energy distribution equal to the product of the local absorption coefficient $\mu_a(r)$ and the optical fluence $\Phi(r, \mu_a, \mu_s, g)$ with:
 - μ_a and μ_s are the absorption and scattering coefficients over the illuminated tissue volume
 - g is the anisotropy factor.

From the explicit equation:

$$p_0(r) = \Gamma \mu_a(r) \Phi(r, \mu_a, \mu_s, g) \quad (1.2)$$

we can note that p_0 depends on mechanical, thermodynamic and optical parameters.

1.2.2 Photoacoustic imaging

In PA imaging we consider the weak mechanical and thermodynamic variations between different tissues and we assume that the contrast of the image is dominated by optical absorption. The components of the tissue, such as water, lipids, collagen, hemoglobin and others, act like biological absorbers and they can be targeted by irradiating tissue at the corresponding dominant absorption wavelength. As such, using tunable lasers, the photoacoustic spectroscopy technique makes it possible to specifically identify chromophores

and quantify their concentration, in order to obtain tissue composition based on endogenous contrast. Figure 1.4 shows how the absorption changes depending on the wavelength.

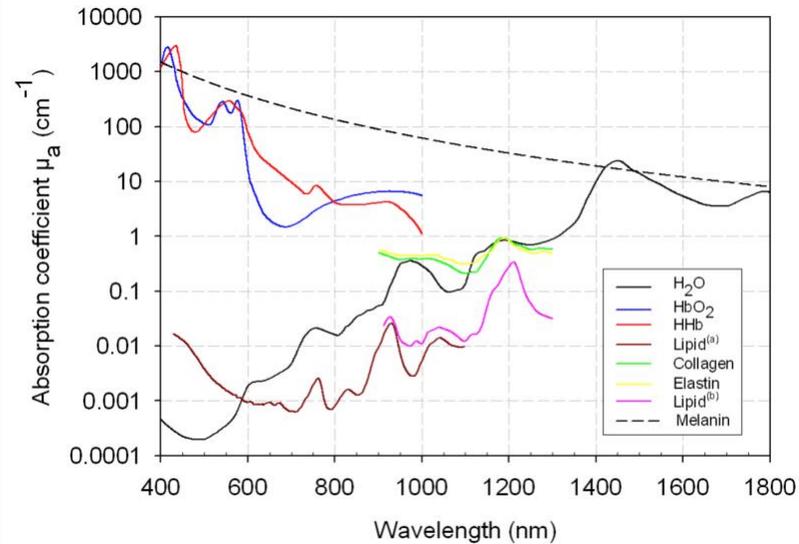


Figure 1.4: Spectrum of the absorption coefficient for different components of the tissue. [12]

In particular, the omnipresence of hemoglobin in living tissues allows the imaging of the microvasculature, which is one of the most important uses of photoacoustic imaging. In addition to visualizing anatomical structures, blood oxygenation measurement is possible by using wavelengths around 650 and 950 nm, around the isobestic point of oxygenated and deoxygenated hemoglobin.

Spatial resolution and penetration depth

All of photoacoustic applications can be achieved over a wide range of length scales from micrometres to centimetres with scalable spatial resolution. The resolution depends on ultrasounds system and on frequency characteristics of generated acoustic waves. The frequency of PA signals typically ranges from 1 to 100 MHz [12]. Thus, by varying the duration and energy of the pulses it is possible to obtain different acoustic frequencies, influencing the final spatial resolution. In addition, the depth to lighting has a great impact

on acoustic waves because of tissue attenuation, limiting the maximum frequency of the photoacoustic waves and therefore giving an upper limit to the spatial resolution obtainable. Therefore, a higher ultrasound frequency of the photoacoustic signal results in better spatial resolution because the relative value decreases. The maximum penetration depth simultaneously decreases as well, because higher ultrasonic frequencies are more attenuated than lower frequencies. Thus, there is a trade-off between resolution and penetration depth, described by the following relationship:

$$\text{Relative Spatial Resolution} = \frac{\text{Penetration Depth}}{\text{Resolution}}$$

In practice, it is found that the ratio between the imaging depth and the corresponding acoustic resolution is of the order of 200 (at 2 cm depth, the resolution of photoacoustic imaging is thus of the order of 100 μm [12]).

1.3 Visibility problems

1.3.1 Origin of visibility problems

In biomedical imaging, randomly distributed subresolution sources or scatterers usually result in speckle artifacts. Improving visibility in photoacoustic imaging using dynamic speckle illumination [13]. Acoustic speckle has a great importance, in ultrasound imaging, to characterise soft tissues and to visualise large or complex-shaped structures. In contrast, photoacoustic imaging is known to be mostly speckle-free because of the emission of strongly coherent acoustic waves interfering constructively in some directions and destructively in others. The absence of speckles in PA imaging is often presented as an advantage of this technique. In fact, the developed speckle is formed by the interference of coherent waves with completely randomized phases but waves from particles close to the boundaries have approximately equal phase delays. As a result, as the initial photoacoustic pressure rises are always positive, these waves add constructively to manifest the boundaries.

On the other hands, the absence of speckle can create some reconstruction artefacts, attributed to the strong initial phase-and-amplitude correlation among the ultrasound waves generated by the individual absorbing molecules or particles after quasi-instantaneous optical excitation. For imaging blood vessels, resulting of the absence of low speckle contribution due to red blood cells, the PA signal is dominated by the shape of the blood vessels and the acoustic wavelength at the mean frequency corresponds to the diameter of the vessel. As figure 1.5 shows, we have two related effects:

1. Limited bandwidth: central part of the reconstructed object are not visible because ultrasound detectors filter out low-frequency components of PA waves emitted by large absorbers (large as compared to the detection wavelength range). For instance, for large blood vessels, only the vessel boundaries may be visible in PA images.
2. Limited view: elongated structures in some directions are not visible on the reconstructed object because the field of view of the probe does not cover all the angles and consequently PA waves escape detection. In other words, the limited view is due to the coherent acoustic waves being directional which can not be measured if they don't reach the probe.

1.3.2 State of the art

Different approaches were proposed in order to solve the problem of PA imaging artefacts with finite-aperture detectors over the course of the time. They are divided into:

- *Methods based on detected acoustic waves emitted in all possible directions.*

One of them consists in enhancing the detection aperture by performing multiple PA acquisitions at different angles. This is possible by rotating a linear transducer around the object [14] or by rotating the imaged object itself[15]. Another technique consists in augmenting the size of detection aperture through additional transducers [16] or acoustic reflectors placed at edges of the imaging zone [17]. Alternatively,

Visibility artefacts

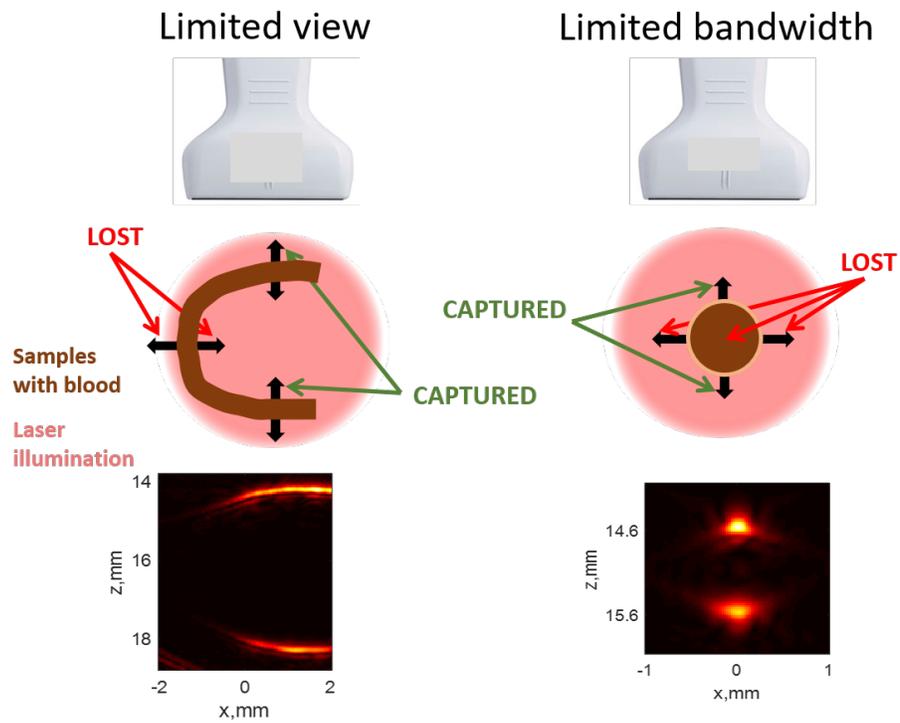


Figure 1.5: The two types of visibility artefacts. We can note that in the reconstructed object we lost (a) the vertical part of the blood vessel because of the limited view, (b) the central part of the body vessel because of the limited bandwidth.

the imaged object and the ultrasound probe can be placed inside a reverberating cavity[18]. In all these cases, it's possible to detect acoustic waves emitted in all possible directions even with a single-element probe but it is required that the whole range of imaging angles is accessible, that is not always achievable in a real clinical environment. Besides, the acquisition time is increased by mechanical scanning of the detectors and the temporal resolution is degraded.

- *Methods based on sparsely distributed absorbers that act individually as isotropic PA sources.*

Dean-Ben et al. demonstrated that reduction of limited-view artefact can be possible by the localization approach [19] or by using a nonlinear combination of tomographic reconstructions representing sparsely distributed moving particles [20]. However, the need for sparsely distributed absorbers requires the use of contrast agents that extends considerably the acquisition time.

- *Methods based on generation of artificial PA sources inside the imaged object.*

For example, one approach consists in heating tissue locally with a focused ultrasound beam and thus generating artificial PA sources via the temperature dependence of the Gruneisen coefficient [21]. By scanning the focused ultrasound beam across the sample and accumulating the resulting PA images, the whole object is reconstructed but this approach is also time-consuming and not completely safe in medical use.

- *Methods based on fluctuations of optical absorbers.*

An approach based on multiple-speckle illumination was proposed by Gateau et al. [22]. In this work, a random intensity distribution of speckle patterns, which changed from pulse to pulse, induced fluctuations in each pixel of the corresponding series of PA images. It was demonstrated experimentally that a second-order fluctuation image provided a representation of the absorbing distribution free of limited-view and limited-bandwidth artefacts. However, speckle illumination is also

limited. The PA signal at each image pixel is the sum of PA signals from a large number of speckles fluctuating within the acoustic PSF (Point Spread Function) : in the case of a diffuse regime, that means typically a depth of 1 mm in the tissues, there are 10^6 speckles. There are, therefore, too many speckles to see a fluctuation in the real application case. Moreover, the work by Gateau et al. [22] provided no clear theoretical explanation for the visibility enhancement in fluctuation imaging with multiple-speckle illumination. Since that work, there have been other studies based on PA signal fluctuations. In particular, fluctuations were deployed to obtain super resolution in PA imaging [23]. Most generally, PA fluctuation imaging utilizes some randomness in PA generation to provide enhanced images as compared to conventional PA imaging. More details are presented in the following chapter.

2. PHOTOACOUSTIC FLUCTUATION IMAGING

In the previous chapter we have discussed about the photoacoustic imaging and its visibility problems. Here, we turn our attention on the photoacoustic fluctuation imaging, the technique used to correct these problems by exploiting the natural fluctuation of the blood flow. We detail the entire methodology to acquire the data, the algorithms used to post-process them to obtain the PA images, the principles of the fluctuation approach and, finally, a set of results to validate its correctness.

2.1 Methodology

2.1.1 Experimental device - Hardware components

The experimental results presented in this report were obtained from the following set-up:

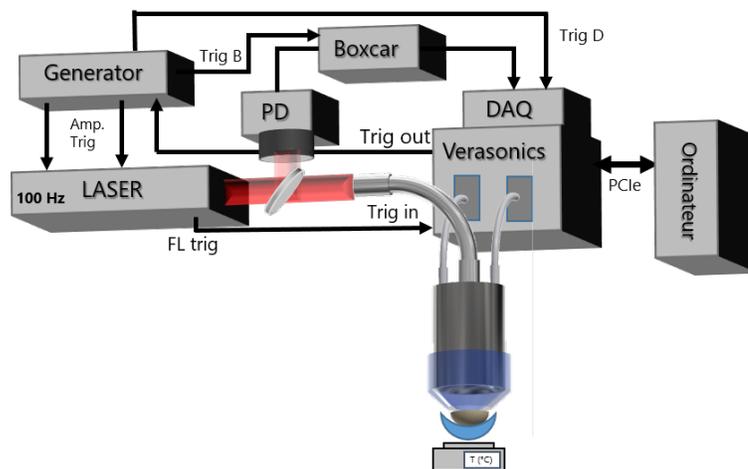


Figure 2.1: Experimental Set-Up

In general, as explained in chapter 1.2.1, a PA acquisition is made from a laser pulse that illuminates a sample coupled to the imaging probe. At each laser pulse, the waves generated are detected by the probe and the signals recorded by the ultrasonic acquisi-

tion system (AS, Verasonics Vantage System). Eventually, the data is transferred to the computer. The laser fluence is recorded by a photodiode (PD). For US acquisitions, the same probe is used in transmission/reception. Here we examine in detail the different components of our setup.

Laser (SpitLight Innolas, Germany)

The light beam is emitted by a Nd-YAG crystal laser pumped by a diode that emits at 1064 nm and that has a second harmonic generator to emit equally at 532 nm. The maximum repetition frequency (PRF) is 100 Hz. The Spitlight laser has an optical parametric oscillator that allows to generate wavelength beams ranging from 680 nm to 980 nm and to change the emission wavelength between these values at the PRF of the laser. For each laser emission, a signal (FL Trig) is transmitted at the "Trig in" input of the Verasonics, which waits for the maximum energy to trigger the Pockels cell (PC, Trig out + Pockels Trig) laser, a sort of optical port that opens during about 5 ns and that allows to release a short light pulse.

Optical fiber (Cereamoptec, Germany)

A fused-silica custom fiber bundle is coupled to the laser and its output is tight to the center of the probe and allows the light to be brought to the center of the field of view and to move the probe without adjusting the beam trajectory.

Probe (Imasonic)

The ultrasound probe is a custom spherical array detector (Imasonic, France) with central frequency $f_0 = 8MHz$ and focal distance = 35mm. It consists of 256 elements distributed over a spherical shell in the form of a Fermat spiral. The imaging area is centered on the center of the 35mm radius sphere and measures about 8x8x8 mm. An 8mm diameter hole was created to insert the optical fiber and to bring light to the imaging area.

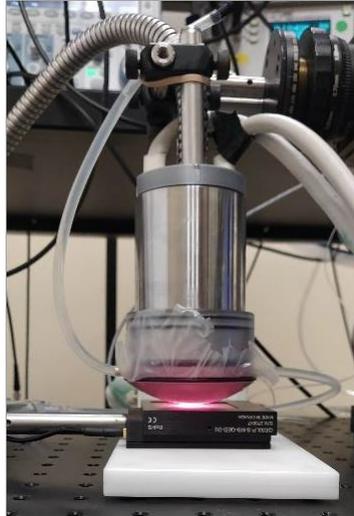


Figure 2.2: 3D imaging probe.

Two connectors are connected to the Verasonics acquisition system and allow the 256 elements to be used simultaneously. For our *in vivo* experiences, a coupling cone was used, allowing to fill the gap between the transducer elements and the sample by degassed deionized water. The contact with the sample was insured by a latex membrane and it was therefore possible to modify the depth of the sample surface by a few mm, inflating or deflating the latex membrane, without losing acoustic contact.

Photodiode and Boxcar

Within the framework of future quantitative experiments, it is necessary to know the laser fluence at the surface of the sample to obtain the optical absorption, as explained in the theory of photoacoustic imaging (Eq. 1.2.2). So, the laser output fluence was measured from a rapid photodiode. The signal of the photodiode was integrated by a gated integrator (Boxcar, Stanford Research System, SR 25) on a time window centered on the pulse, whose size was determined by the use of an oscilloscope and digitized by the DAQ, the Verasonics integrated acquisition card. This stage of integration of the PD signal allows to quantify the energy of each impulse, because the length of the impulse and therefore of the signal to be sampled produced by the photodiode is about $20ns$, and a consequent

sampling rate of at least 4 times lower than $1/20ns$ is required. The integrated signal for a dozen pulses is visible in Fig.2.3, where we can see the variation of energy produced by the laser between the shots.

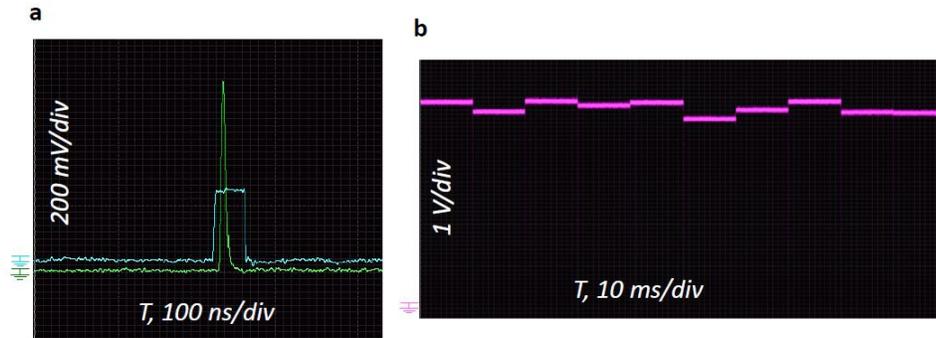


Figure 2.3: Signals measured with oscillator. a : Photodiode signal for a shot (green) and gated integrator (blue). b : Result of integration for about ten impulses. This signal is then converted by the DAQ.

It is the Trig D (in Fig.2.1) supplied by the generator and triggered by the Verasonics acquisition system (Trig Out) that allows the DAQ to start converting the PD signal. In order to validate the synchronisation of the photodiode and the laser, before each series of measurements we made a visual comparison between the signal coming from the photodiode and the signal coming from a quantitative calibrated pyroelectric detector that allows to quantify the energy of each pulse (Gentec QE50, France), located in the imaging area. The synchronization verification is shown in Fig.2.4. The photodiode doesn't give the fluence in mJ, so the pyroelectric detector is used for a calibration.

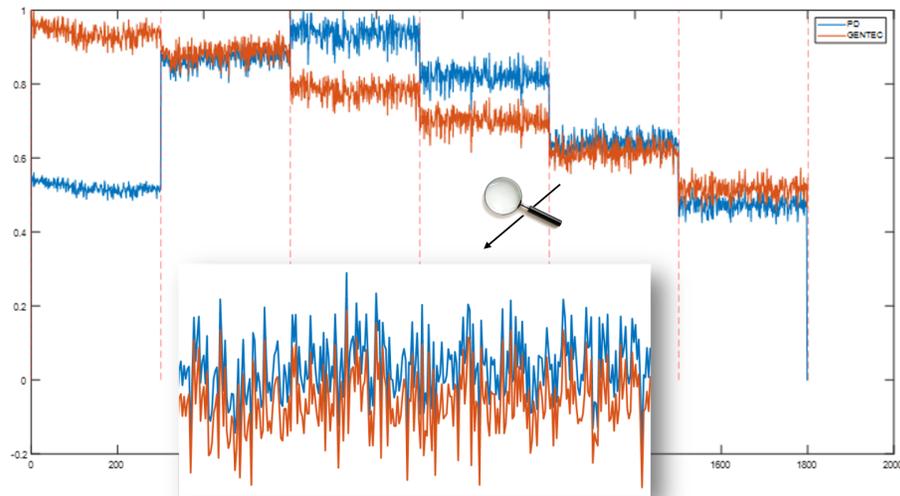


Figure 2.4: Plot of the normalized signals coming from the PD (in blue) and the pyroelectric detector (in orange). The zoom on the plot at each of the six wavelengths (x-axis: 300 laser pulses for [700 730 760 790 820 850] nm) allows to compare them and validate the synchronisation.

2.1.2 Vasculature model

The imaged sample was a chicken embryo whose chorioallantoic membrane (CAM) is an optimal *in vivo* model to study blood vessels thanks to its visibility, accessibility and rapid developmental growth for experimental manipulation. The CAM consists of two epithelial sheets that limit a thin layer of stroma. It is within the stroma that the blood vasculature and lymphatics reside. It is important not to mistake the CAM with the yolk sac membrane (YSM), which is also highly vascularized, but it's associated with the yolk and it has distinct properties. Fig. 2.5 shows the difference of the two membranes at day 4 post-fertilization: we can clearly distinguish the vascularized CAM that expands from the hind-gut and YSM in the background. By contrast, in an older embryo, their difference is not so noticeable because the CAM exceeds the area covered by the yolk. The yolk is associated with the YSM that is always present but under the embryo [24].

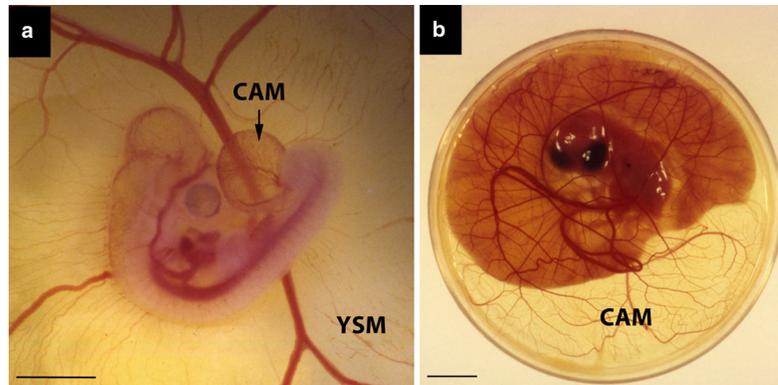


Figure 2.5: Chicken embryo and associated structures. a) At day 4 post-fertilization CAM and YSM are highly vascularized but distinguishable (YSM is in the background). b) At day 12 post-fertilization in a petri dish: YSM is under the embryo and it is only associated with the yolk. [24]

For our experiments, we stored the non-fertilized eggs obtained from a local farm in a cellar at 13°C and we moved them to an incubator at 37°C and 66% of humidity in order to image them after 9-10 days, when they are sufficiently developed as shown in Fig.2.6 and as seen during our optimization tests.



Figure 2.6: Embryo vasculature at day 10 after incubation.

After making a small hole through the shell using a scalpel, 1.5 mL of egg white was removed with a syringe, and the top part of the shell was cut with scissors. Warm phosphate-buffered saline (PBS) solution was poured into the shell to ensure acoustic

contact with the latex membrane. The sample was placed in a plastic container with a central hole containing water maintained at 37°C using a hot plate (Fig.2.7).

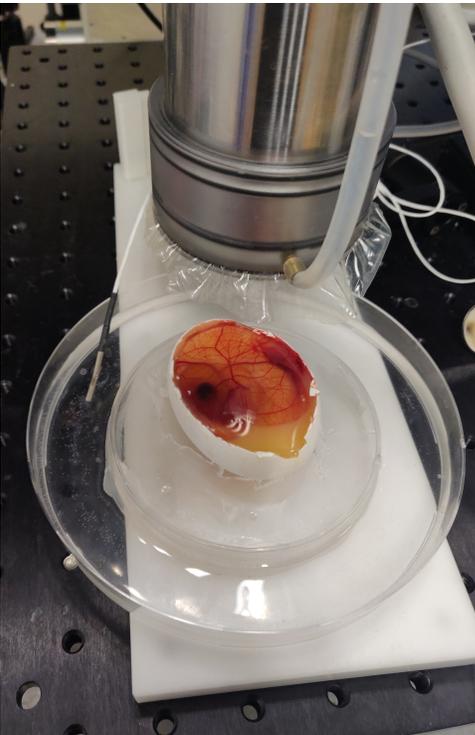


Figure 2.7: Imaged embryo placed in a plastic container with a central hole containing water. The container was filled with water maintained at 37°C thanks to a hot plate.

2.1.3 Software

The software is included in the acquisition system Verasonics Vantage 256. The Verasonics Vantage Research Systems is a programmable electronic platform for transmitting, receiving and processing ultrasound information. Essentially all aspects of this ultrasound system are able to be controlled by the user, and it is possible to create new and novel methods of acquisition and processing. The System is provided with a folder with scripts that can be modified according to the user's needs. The scripts supplied to program the system are written in the MatlabTM language, which provides an environment for defining the various programming objects contained in a sequence of events. The scripts define the order of actions to be carried out in the Vantage unit, as well as actions that take place in the software environment of the host computer. In our case, we have developed an interface that allows to manage in a simple and effective way the various stages of the acquisition (Fig.2.8).

Our system consisted of two modes of use:

- **Live Mode:** it allowed the display of the volume in the three planes XY,XZ,YZ and the adjustment of different parameters in real time; that enables the movement on the sample to image, in order to find the most interesting structures.
- **Save Mode:** it allowed to save a desired number of RF (Radio-Frequency) data.

Before focusing on each of the two modes, we describe in detail the part of the code common to both. First of all, the code allowed to load the positions of the probe transducers used (as in chapter 2.1.1), the number of flash angles to use and to set some specific parameters of the system, like the number of transmit and receive channels, the speed of sound and the transducer connectors to use (both of them in our case). It was also necessary to increase the gain of the amplification stages located downstream of the active elements of the receiver.

- *Pgagain*: amplifier output buffer placed before the A/D converter (en dB).

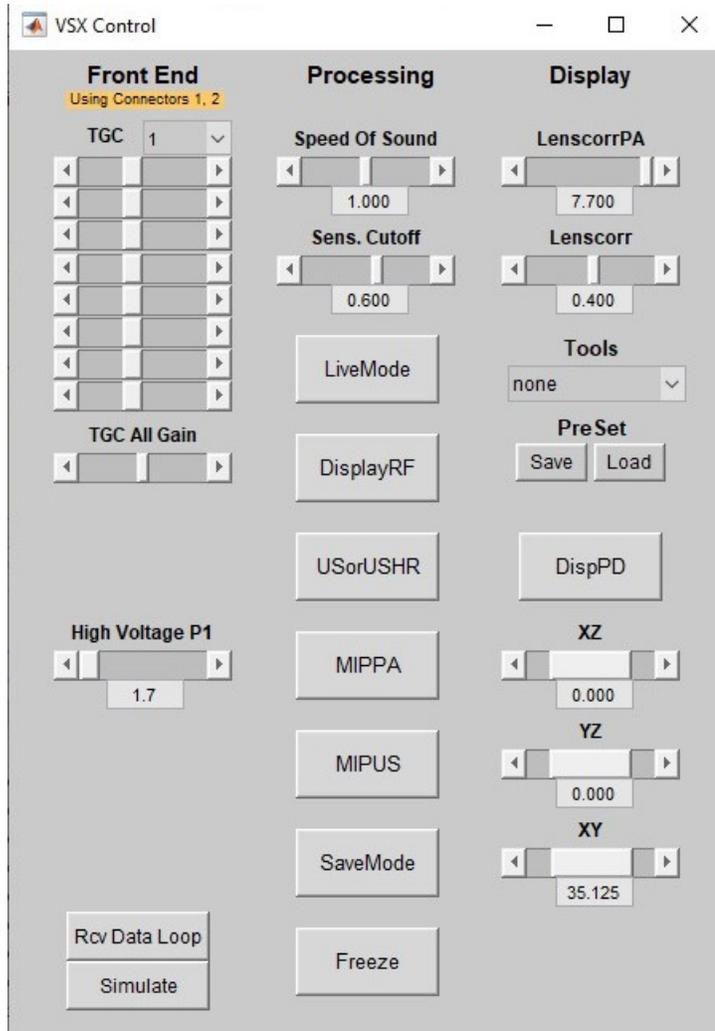


Figure 2.8: Main interface of our acquisition software where the buttons LiveMode, DisplayRF, USorUSHR, MIPPA, MIPUS, Save, DispPD and the sliders XZ, YZ, XY and LenscorrPA represent the custom functions that we added.

- *Lnagain*: low noise input amplifier stage (en dB).

Now we can describe in detail the codes corresponding to the two modalities.

Live Mode

As said before, the 'Live mode' allowed, in real time, the visualization of the volume imaged through its projection on the three planes XY, XZ and YZ. Thanks to the button "USHR/US" we could switch from the image USHR (ultrasound, high resolution) to the US (ultrasound) reconstruction, while the other three projections always corresponded to the PA (photoacoustic) reconstruction. We firstly needed to define some storage space for the RF data that we wanted to acquire. This required defining a Receive Buffer in the host computer (the Receive Buffer is the only buffer that can receive RF data transferred from the Vantage local memory on the acquisition modules). It is a Matlab cell array $RcvData\{n\}(i,j,k)$, where n is the buffer number, i is the row index, j is the column index, and k is the frame number. In LiveMode, we used a first buffer composed as follows:

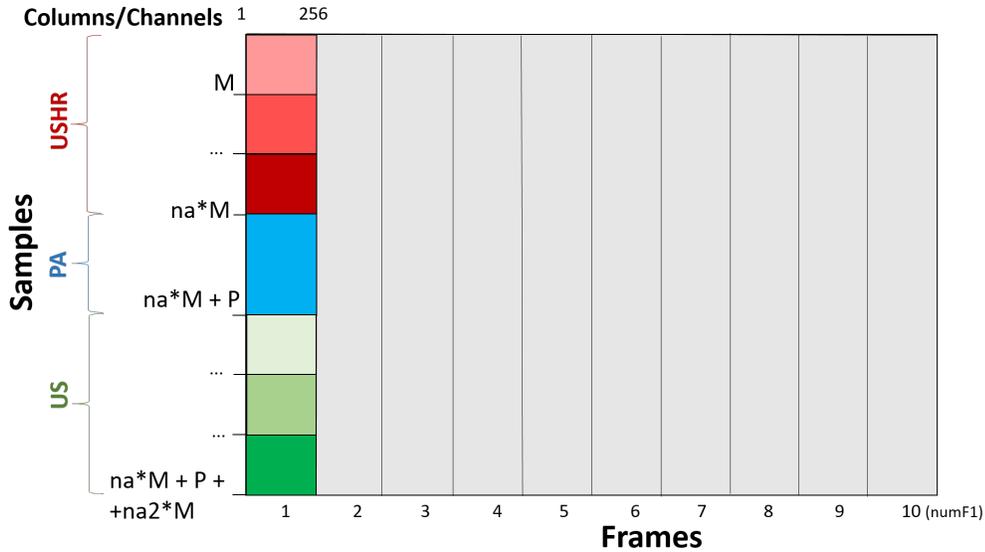


Figure 2.9: ReceiveBuffer composition for LiveMode. M is equal to the number of samples for US and USHR acquisitions and P is equal to the number of samples for PA acquisition. na and $na2$ are the numbers of flash angles for USHR and US acquisitions respectively.

The number of row for each frame were set according to the number of samples of RF data per channel that will be acquired. Its calculation was based on the depth of the acquisition, the sample rate and the number of acquisitions per frame. In this case, the number of acquisition per frame was set to 1 because in this part we just needed to reconstruct and show the volume, not save it. *na* and *na2* are the the number of emissions used for reconstructions for USHR and US acquisition respectively. The number of columns per frame was set equivalent to the number of receive channels available in the system. When data are transferred from the Vantage Unit's local memories to the host, the data from all receive channels are transferred together, with each channel transferred to its corresponding column of the Receive Buffer.

The TX object definition

The TX structure was defined for each unique transmit action in our sequence. It is important to specify that, in order to deactivate the ultrasonographic probe in the case of photoacoustic acquisition, the vector 'Apod' was placed at 0 (mode reception only).

The TGC object definition

The TGC is the time gain control function. It allows to have a time-varying gain during RF data processing: it increases through time to compensate acoustic attenuation. In this code, the TGC for the photoacoustic signal was set to the maximum value (1023).

The Receive object definition

Receive operations specify other attributes of the input signal processing. The Receive structure specifies many parameters, including how long the receive period should run, how the received data are to be sampled and filtered, which receive channels are active, and where to store the RF data.

Reconstruction Objects

They describe the reconstruction methods required to form pixel information from the acquired receive data, and together with the ReconInfo objects, specified all the attributes needed to reconstruct a full frame. Normally, the reconstruction software automatically calculated Look-Up-Tables to specify the reconstruction processing. We decided to specify custom LUTs for reconstruction in our setup script to solve an internal problem of Verasonics system that did not allow a receive only reconstruction and to be able to tune the origin of time used in the reconstruction which is due to the laser latency of $400ns$ after trigger.

Sequence Control Objects and Event object definition

After the choice of the sequence Control objects, used to associate special function actions with a sequence event, we modified the event list in order to guide the acquisition process. As written previously, in LiveMode we gave the possibility to:

- Display the three projections on the axes XZ, XY and YZ of the reconstructed image;

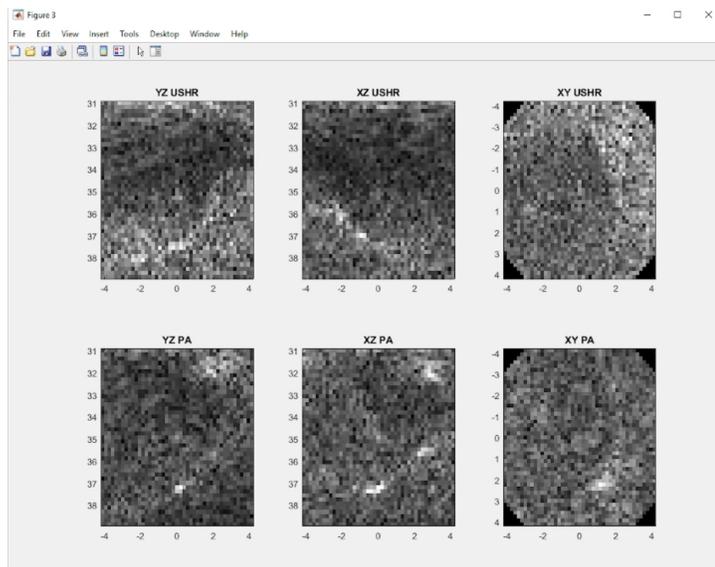


Figure 2.10: Real time reconstruction of the imaged object.

- Activate volumetric reconstructions and display the MIPs (maximum intensity projections) by clicking on "MIPPA" or "MIPUS" buttons;

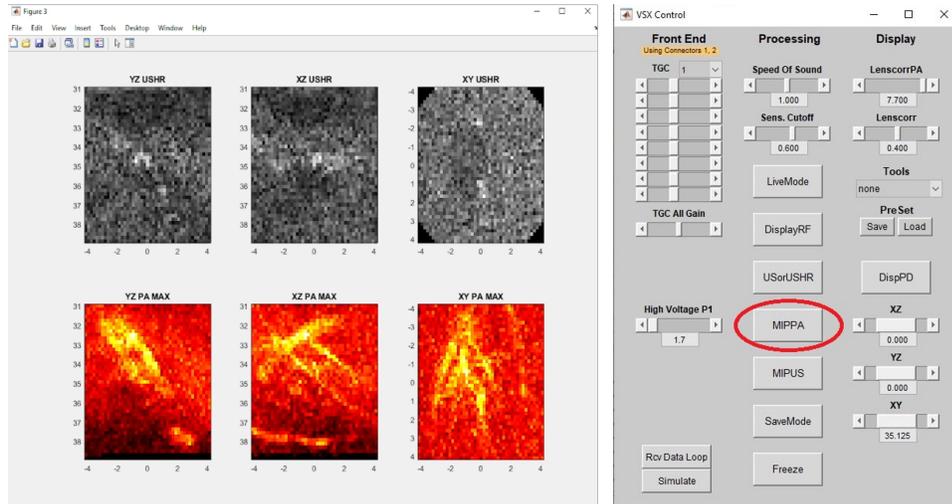


Figure 2.11: Interface of the volumetric reconstruction in Live Mode. The red circle underlines the button used to display the MIPs.

- Display the RF signals before saving to control signal intensity and avoid saturation;

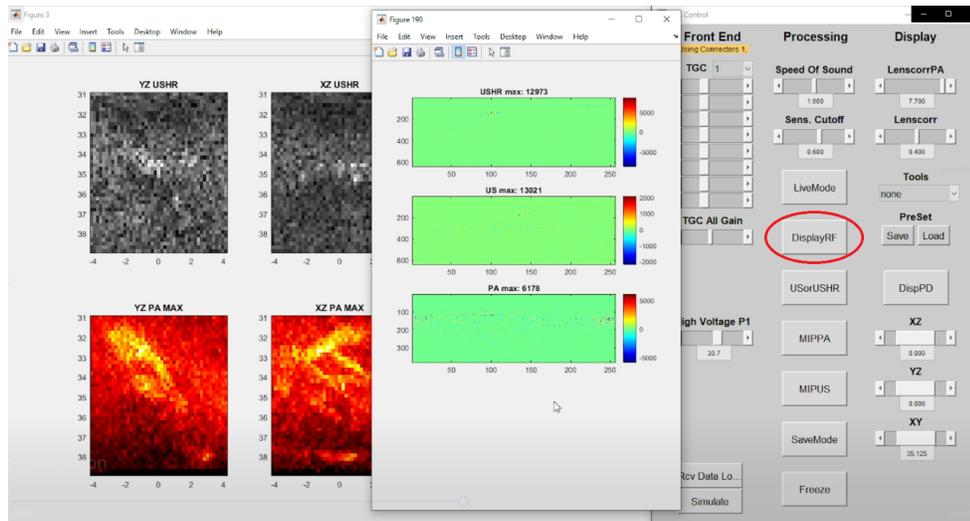


Figure 2.12: Interface of radio-frequency signals control. The button DisplayRF permits to show the max. of RF signals to avoid saturation. It can be decreased by modifying the TGC value.

- Switch from USHR reconstruction to US reconstruction through the "USorUSHR" button;
- Set independently several parameters such as voltage, t_0 and TGC. Being able to change the TGC values in real time was useful to avoid saturation problems. The t_0 is a value that can be varied for the improvement of the reconstruction; it is set to account for the laser latency after trigger (400 ns) and also the potential travel time in the protective layer between water and the transducer.

All of these possibilities were implemented thanks to the process objects called during the events and the reload of the events sequence. In LiveMode version, the sequence of events provided, in order, the acquisition of USHR, PA and US signals. For the PA signals, in particular, the Seqcontrol of the event required the sending of a TriggerOut, which triggered the laser to emit at the wavelength defined at 730 nm. Below, we have inserted a Verasonics function to transfer acquired data to the host buffer. Finally, we had the signal reconstruction events and its display. The Livemode acquisition was concluded with a "JUMP" event that allowed, in the absence of other orders, to return to the first event of the succession in order to continue to obtain images of the sample in real time.

Save Mode

Once we found the desired position on the sample, we saved the data thanks to the 'Save mode', by pressing on the SaveMode button. Knowing that the laser worked at a PRF_{PA} (maximum repetition frequency) equal to 100 Hz and having placed the PRF_{US} at 200 Hz, the diagram of our acquisition system was:

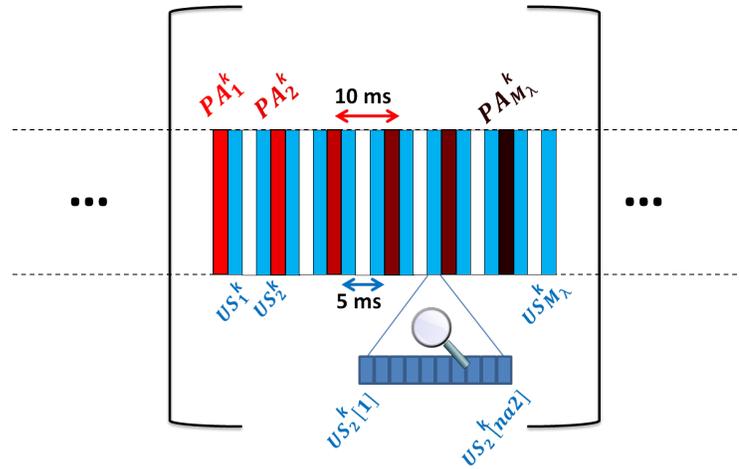


Figure 2.13: Sequence of an acquisition US-PA with $M_\lambda = 6$ wavelengths, $PRF_{US} = 200$ Hz, $PRF_{PA} = 100$ Hz and $na2 = 10$ (number of emissions to form an US image).

Compared to the LiveMode case, each frame had a number of acquisitions other than 1. In particular, the second buffer was then composed as follows:

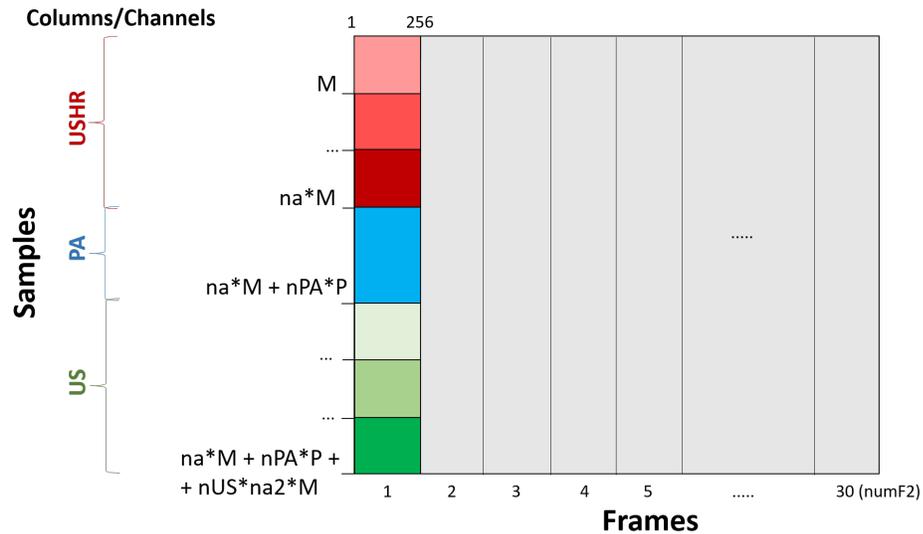


Figure 2.14: ReceiveBuffer composition for SaveMode. M is equal to the number of samples for US and USHR acquisitions and P is equal to the number of samples for PA acquisition. na and $na2$ are the numbers of flash angles for USHR and US acquisitions respectively. nPA and nUS are the number of acquisitions for PA and US ($nPA = 50$, $nUS = 100$). The number of frames used was 30.

In this case, the number of frames used was automatically calculated based on the number of PA images desired. In our experiments, we chose to acquire 1500 PA images at a time, so the number of frames per save was 30, each containing the 50 PA acquisitions to reconstruct.

The event sequence in SaveMode started with a number of acquisition not saved in order to warmup the laser. Compared to the LiveMode, after PA acquisition we added an event called “readAI” which, thanks to a Verasonics function, was capable to command the DAQ to read the photodiode signal and store it in its own buffer memory. This signal is displayed at the end of saving after it got transmitted on the host RAM memory, as showed in Fig.2.15 and the system is automatically frozen.

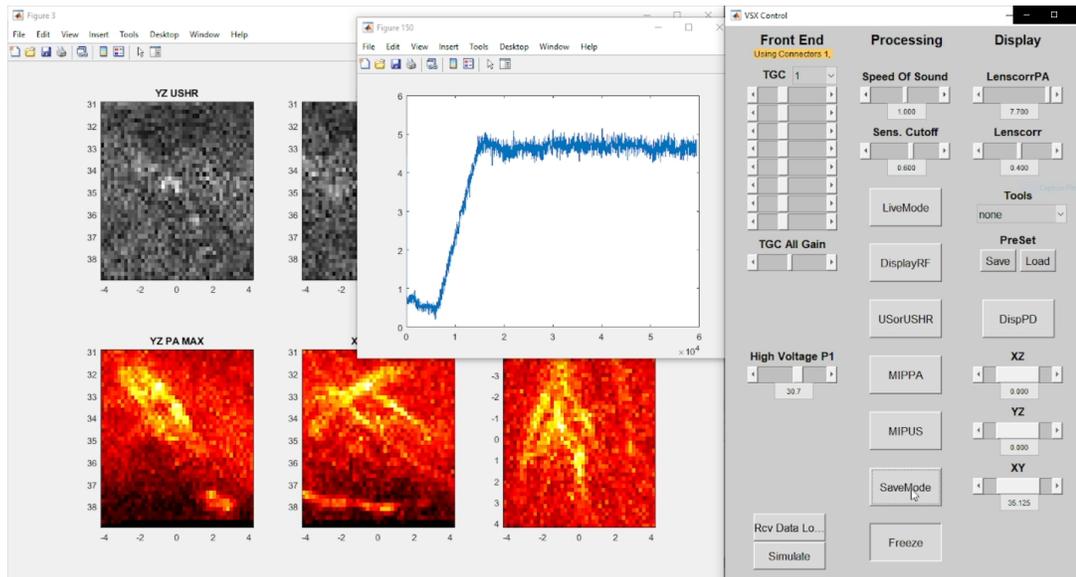


Figure 2.15: Display of photodiode signal trend after acquisition.

At that point, we had two choices:

- Return to LiveMode to change the position;
- Change the laser wavelength and make an acquisition at a different wavelength or in multi- wavelength mode to obtain a spectroscopy image by changing laser settings.

We also implemented another feature to make our system as safe and practical as possible: an automatic resolution of memory problems. Using the values of Fig.2.14 for US and USHR images, we found ourselves able to save up to a maximum of 1200 PA images at any position we wanted to image, without running into memory problems. Therefore, in our code, if the number of PA images to save was greater than 1200, our software automatically saved the US images for the half of the total number of frames and did not allow them to be saved in the other half part. In our case, with 1500 desired images, we finally acquired 15 frames with Buffer2 dimension and 15 frames with Buffer3 dimension. The third buffer was composed as follows:

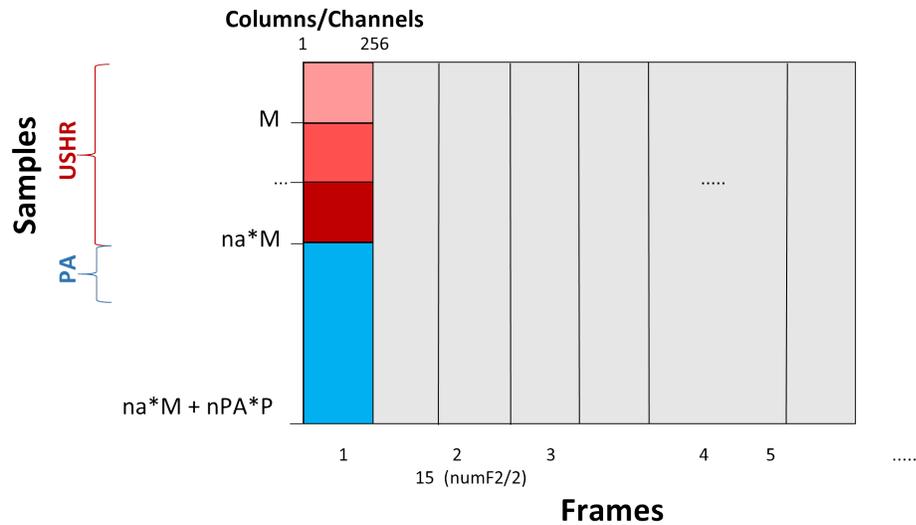


Figure 2.16: ReceiveBuffer composition for SaveMode without US storage.

Finally, our software allowed in post processing, thanks to the Timetags function of Verasonics, to keep track of the time t to which each frame was beginning to be acquired, to validate the continuity of the acquisition. From a graphical point of view, we calculated the time difference $t_{i+1} - t_i$, with the number of frames from 1 to the total number of frames and if there were no delays we had a constant output equal to the frame acquisition time.

2.2 Photoacoustic image reconstruction

The output of the acquisition machine as a result of a PA acquisition with a multi-element array is a radio-frequency (RF) signal. In this project, we used the algorithm called "*Delay and sum beamforming*" to provide images from a set of PA signals.

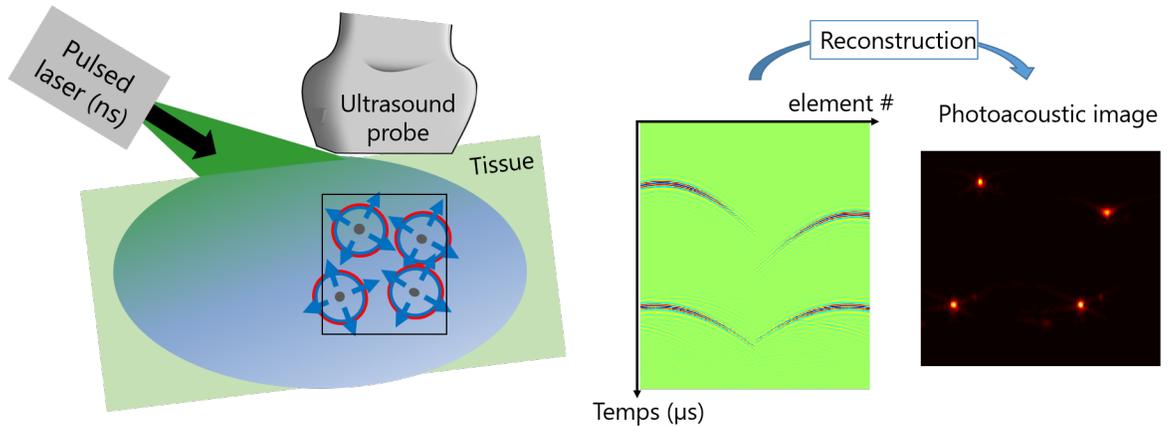


Figure 2.17: Summary of the different steps to obtain the final photoacoustic image.

2.2.1 Delay and sum beamforming

In this section, the beamforming algorithm is described in detail. We used it to demonstrate the results of standard photoacoustic (PA) reconstruction. The idea is very simple: add a delay to each signal point such that the signals from a particular location are aligned before they are summed.

In fact, the signals captured by the transducer are similar in wave form, but show different delays and phases. The time of flights are proportional to the covered distances. The signal of each element is shifted by a corresponding time of flight difference depending on the reconstructed point. As a result, all the signals received from the transducers are aligned in time. The signals of all elements are summed up. Fig.2.18 shows a typical radio-frequency (RF) frame that is available at the output of the acquisition machine as a result of PA acquisition with a multielement array. Each cell of this frame contains the

signal $S(t_i, k)$ which is equal to the quantized value of voltage on transducer element k registered at time t_i . The interval between successive time values t_i and $t_{(i+1)}$ is equal to $\Delta t_s = 1/f_s$, where f_s is the sampling frequency of the acquisition machine.

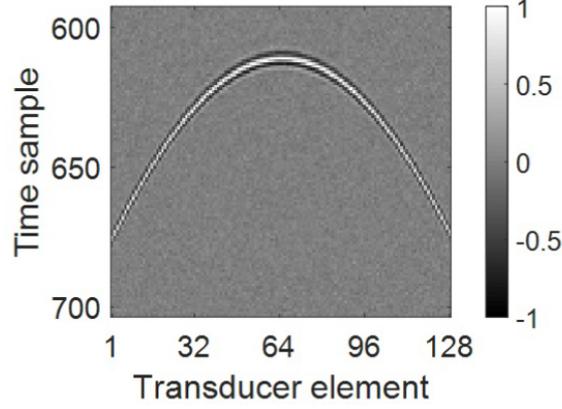


Figure 2.18: A typical radio-frequency (RF) frame available at the output of the acquisition machine as a result of PA acquisition with multielement linear array. [25]

In delay-and-sum beamforming each point of the reconstruction image $A(x, y, z)$ is obtained with the following equation:

$$A(x, y, z) = \sum_{k=1}^{N_{el}} S(t(k, x, y, z), k), \quad (2.1)$$

where:

- N_{el} is the total number of transducer elements;
- $S(t(k, x, y, z), k)$ is the signal on element k at the time moment $t(k, x, y, z)$;
- $t(k, x, y, z)$ corresponds to the arrival of the signal from the source placed at x, y, z on the element k .

If we assumed that PA waves propagate in a homogeneous isotropic medium, the arrival time $t(k, x, y, z)$ is given by

$$t(k, x, y, z) = \frac{1}{v_s} \sqrt{(x_k - x)^2 + (y_k - y)^2 + (z_k - z)^2} \quad (2.2)$$

where:

- v_s is the speed of sound in the medium;
- x_k, y_k and z_k are the coordinates of the transducer element k .

In our project, we used RF data acquired through the "Save Mode" functionality explained in section 2.1.3.

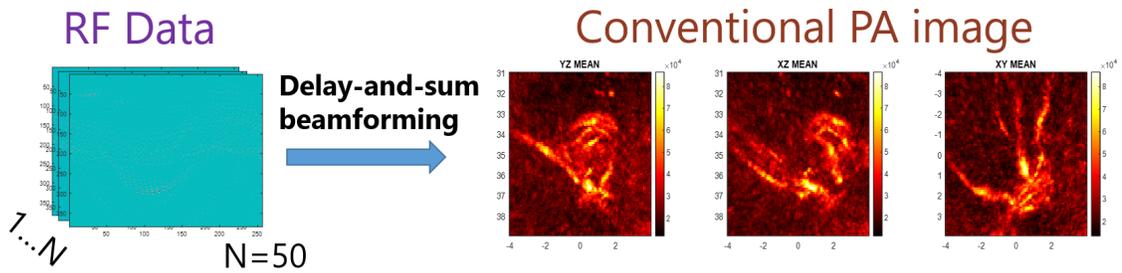


Figure 2.19: Delay-and-sum beamforming applied to a N=50 RF data to obtain the conventional PA image (displayed as MIPs).

2.3 Photoacoustic fluctuation imaging

While it has been successfully demonstrated experimentally for super-resolution imaging (induced either from multiple-speckle illumination [26] or from random distributions of absorbing particles [27]) and for visibility enhancement (with multiple-speckle illumination). S. Vilov et al. [25] proposed a theoretical framework relevant to both fluctuations induced by random illumination patterns and fluctuations induced by random distributions of absorbing particles. They provided a physical insight into why visibility artefacts are absent from second-order fluctuation images and they demonstrated experimentally that harnessing randomness induced by the flow of red blood cells produce photoacoustic fluctuation images free of visibility artefacts. In fact, red blood cells have a bigger size than speckles and there are around 10^{-3} of them within the acoustic PSF, so it's possible to see the fluctuation due to the blood flow.

2.3.1 Principles and theory

S. Vilov et al. [25] demonstrated that if we assume that we want to image vessels in the $f(\vec{r})$ support, where $f(\vec{r})$ represents the vessels, the absorption at each point of the medium depends on the presence or not of an optical absorber:

$$\mu_k(\vec{r}) = \mu_0 g_k(\vec{r}) f(\vec{r}) \quad (2.3)$$

where $g_k(\vec{r})$ is a fluctuating function which represents the presence or not of a red blood cell. Its average value $\langle A_k(r) \rangle = \eta$ is the volume fraction of red blood cells. With a single photoacoustic image the obtained formula is:

$$A_k = \Gamma \mu_0 F_0 g_k(r) f(r) * h(r) \quad (2.4)$$

and its module in order to remove the oscillations:

$$|A_k| = \Gamma \mu_0 F_0 \eta |f(r) * h(r)| \quad (2.5)$$

In the formula (2.5), $|f(r) * h(r)|$ is the convolution between the imaged object and the PSF of the system and it is important to note that there is a coherent summation. When we represent the PSF with the oscillations (just the real part of the PSF), there are the oscillations at the center frequency of the transducer (along Z) and after the Fourier transform of the PSF the central part of the spectrum is lost. Therefore, the convolution between the PSF and the vessel in the form of "C", vertical structures in the image are lost too.

Finally, as expressed in 1.3, conventional photoacoustic imaging based on the acoustic coherence of received signals is limited. By contrast, with several images, it is feasible to calculate the variance:

$$\sigma^2[A](r) = \Gamma^2 \mu_0^2 F_0^2 \sigma_g^2 V_g f^2(r) * |h(r)|^2 \quad (2.6)$$

based on a incoherent summation since the modulus of the PSF is calculated before the convolution with the object. The equation (2.6) is valid for both speckle illumination imaging and fluctuation imaging due to red blood cell flow.

2.3.2 Noise sensitivity

However, in the real application, the noise is added to the calculation of variances: on the fluctuation image, the fluctuation is increased by the noise level.

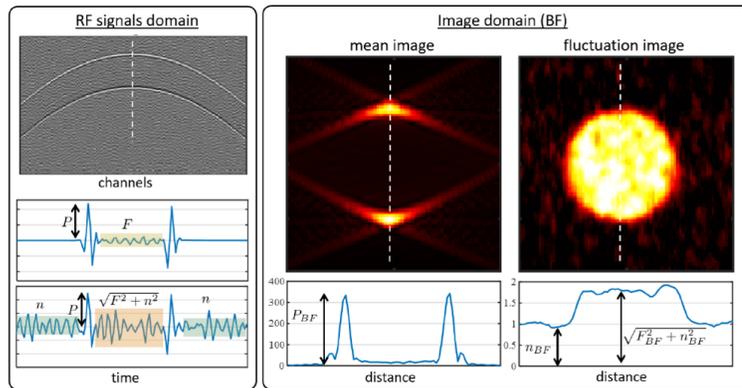


Figure 2.20: In the left section (RF signals domain) a peak P of the interface at the entrance and exit of the vessel and a fluctuation level F can be defined. When a noise n is added, it affects the entire signal, so the fluctuation F must be extrapolated from $\sqrt{F^2 + n^2}$. In the right part of the figure (Image domain), on the average image there are no fluctuations while the fluctuation image present the level of fluctuation increased by the noise level. [29]

For which fluctuation to noise ratio is it possible to form a fluctuation image? It has been demonstrated by the team that with a sufficient number of acquired images it can be obtained a good image even with a small ratio (Fig. 2.21). Consequently, the number of acquired images must be bigger than 500.

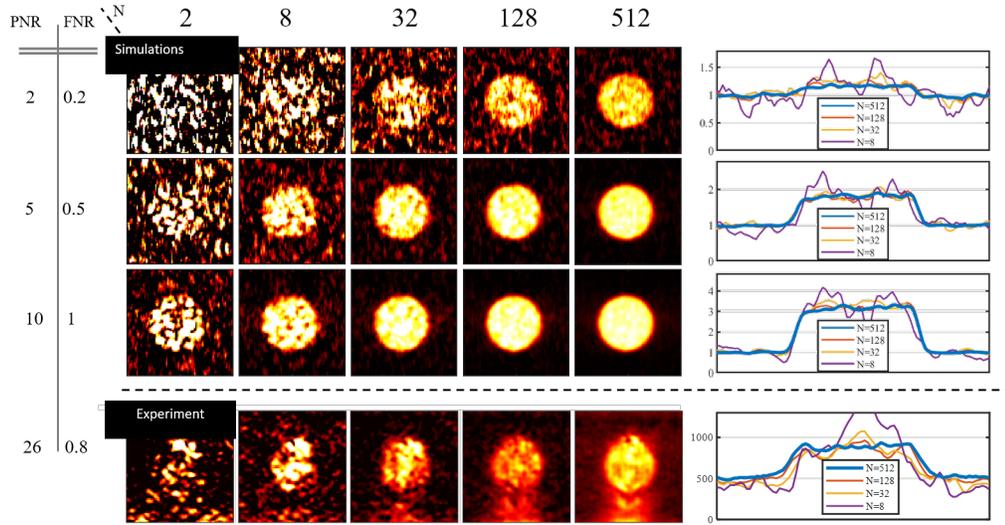


Figure 2.21: The upper section of the figure was obtained thanks to a simulation set. The first line shows a fluctuation to noise ratio (FNR) equal to 0.2 (noise louder than the fluctuation). With enough images, a good reconstructed image is obtained. When the FNR increases, fewer images needed to obtain a comparable image quality. In experience a FNR of 0.8 has been founded. [29]

2.3.3 Parasitic fluctuations and SVD

In addition to the noise problem, there is also a problem due to the existence of parasitic fluctuations, especially of the laser, whose energy fluctuates from pulse to pulse affecting the variance calculation.

In Fig.2.22 we can note that the fluctuation image $\sigma[A](r)$ is dominated by the average image $|\langle A_k(r) \rangle|$.

The proposed theory to solve this problem is the SVD (Singular Value Decomposition) filtering [28], which is based on the following steps:

- The data are represented under the form of a variable $a(x, y, z, t)$ of size (nx, ny, nz, nt) . Its sampling is rearranged into a 2D space-time Casorati matrix A of dimension $(nx \times ny \times nz, nt)$.

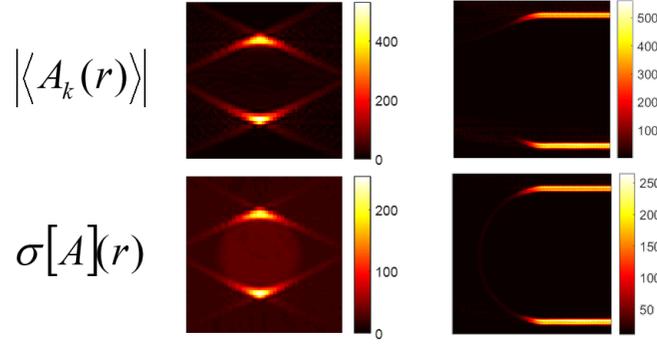


Figure 2.22: Fluctuation image $\sigma[A](r)$ dominated by the average image $|\langle A_k(r) \rangle|$, before SVD.

- The Casorati matrix A is decomposed in three matrices:

$$A = U\Delta V^* \quad (2.7)$$

Where Δ is a non-square $(nx \times nz, nt)$ diagonal matrix, U and V are orthonormal matrices with respective dimensions $(nx \times nz, nx \times nz)$ and (nt, nt) and * stands for the conjugate transpose. Columns of U and V matrices correspond respectively to the spatial and temporal singular vectors of A. The singular value decomposition (SVD) can be thought of as decomposing a matrix S into a weighted, ordered sum of separable matrices A_i . By separable, we mean that the matrix A can be written as an outer product of two vectors $A_i = U_i * V_i$. Specifically, the matrix A can be decomposed as:

$$A = \sum_i S_i A_i = \sum_i S_i U_i * V_i \quad (2.8)$$

Thus, the SVD can be used to find the decomposition of an ultrafast ultrasonic dataset into separable space and time filters. Here U_i and V_i are the i^{th} columns of the corresponding SVD matrices and S_i are the ordered singular values.

- The limits A and B are chosen to complete the filtering by SVD:

$$A_{svd} = \sum_{i=a}^b S_i A_i = \sum_{i=a}^b S_i U_i * V_i \quad (2.9)$$

The advantage of SVD is that the singular vectors are ordered from the most energetic to the least energetic. The most energetic one is the average object because by definition it has a nonzero temporal average and therefore a strong energy. It has been shown that the time trace of $|U_1(\vec{r})|$ matches exactly the laser fluctuation, while on the others U_i the mean is zero and the fluctuations describe the fluctuations due to blood flow. The choice of the limits a and b serves to eliminate a certain range of singular values that correspond to the average fluctuating object with the parasitic sources, in order to keep only the fluctuating object due to the blood flow. As result, we theoretically have the correct recostruction of the vessel:

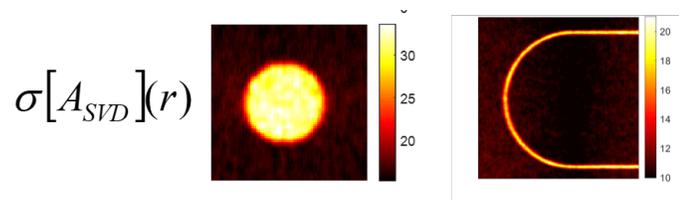


Figure 2.23: Fluctuation image after SVD, free of artefacts.

2.4 Results

In this section, we show the results obtained in post-processing. These results will form the data-set used in our deep-learning approach, that we will explain in the next chapter. We imaged around 40 embryos to get a data-set of 140 conventional images with their respective fluctuation PA images. In each example, the three plots at the top of the graph represent the MIPs of the conventional PA image, affected by visibility problems. The plots in the lower part represent the MIPs of the fluctuation PA image, the corrected one, obtained thanks to SVD. If we focus on the visibility problems described in section 1.3, Fig.2.25 and Fig.2.26 show that the "Limited bandwidth" problem has been properly corrected. The "Limited view" problem has been corrected too, as in Fig.2.27 and Fig.2.28. In addition, we also found some error in our processing. In fact, we had sometimes some absorbing structures that disappeared in fluctuation PA image (Fig.2.29). A reason could be the presence of haemorrhages (where there is blood but no fluctuation of red blood cells) or the presence of pieces of eggshell egg shell, provoked by the cutting of the shell itself during the egg opening. We also noted on the final image the presence of bright points that absorbed more than the other structures (Fig.2.30). Although we tried to well manage the acquisition, in some acquisitions there were probably some pieces of eggshell, bubbles on the probe surface (chapter 2.1.1), on the latex membrane or on the embryo chorioalantoic membrane (chapter 2.1.2). For the last example, the 3D reconstruction is showed. (Fig.2.27 and 2.27).

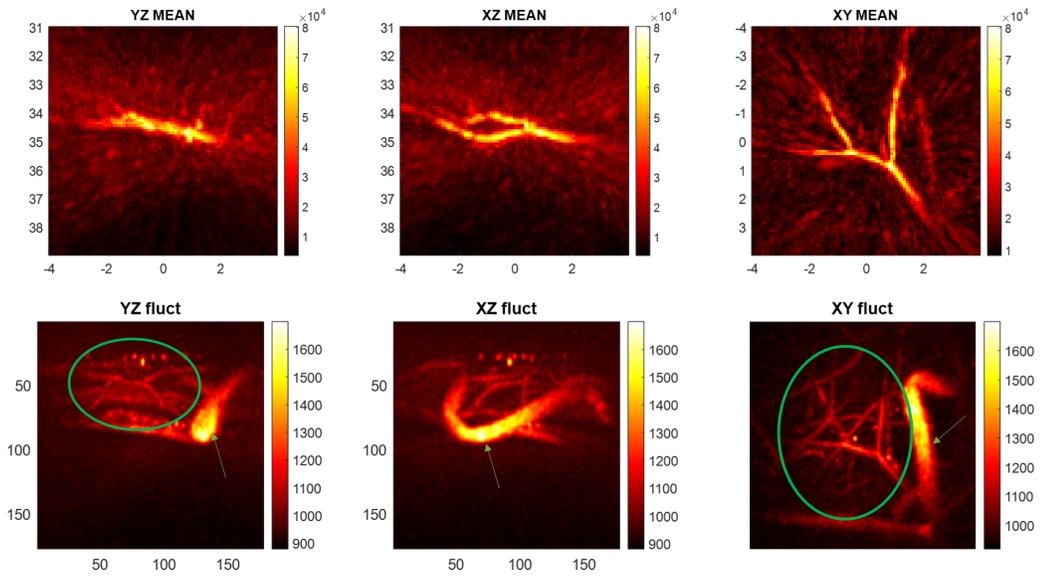


Figure 2.24: Result 1. The fluctuation PA image points out a big vertical vessel and a really complex vascular net that were absolutely not visible in the conventional PA image.

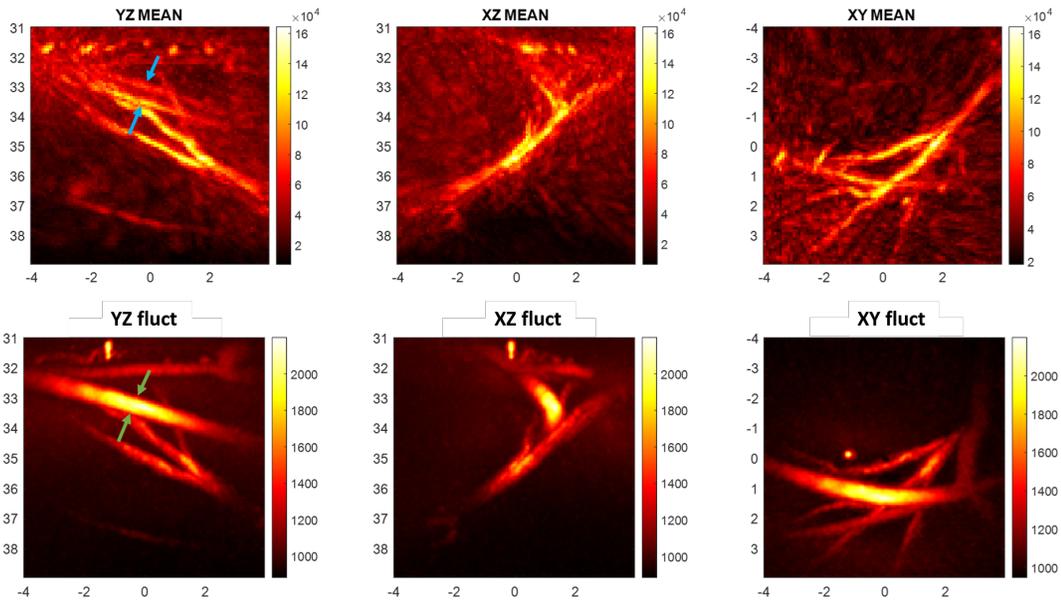


Figure 2.25: Result 2. The "Limited bandwidth" problem has been properly corrected. Blue flashes reveal a vessel whose just upper and lower edges are reconstructed, while green flashes show the vessel successfully reconstructed in its inner part.

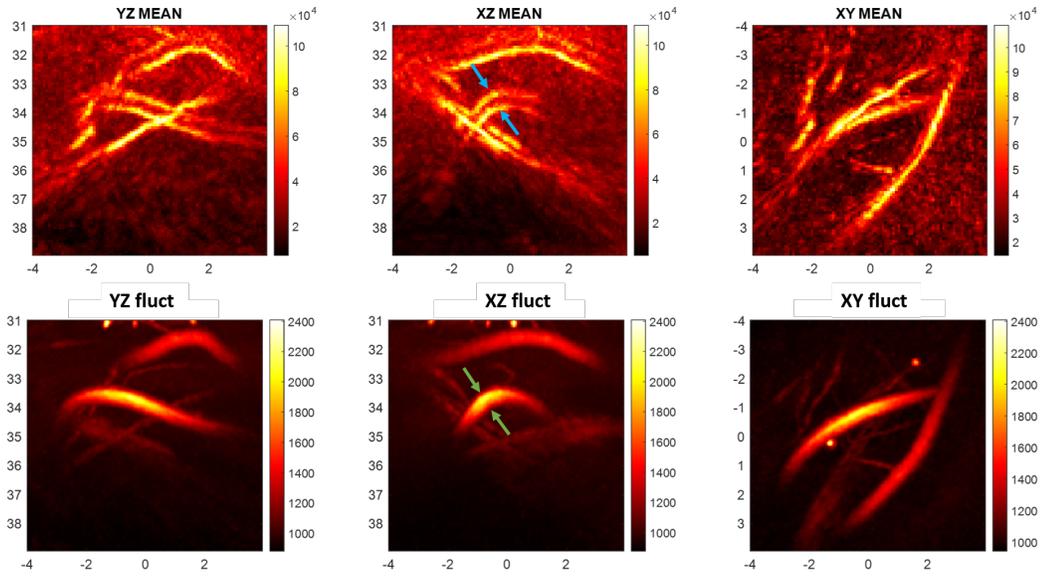


Figure 2.26: Result 3. As in Fig.2.25, the internal part of the object is well reconstructed.

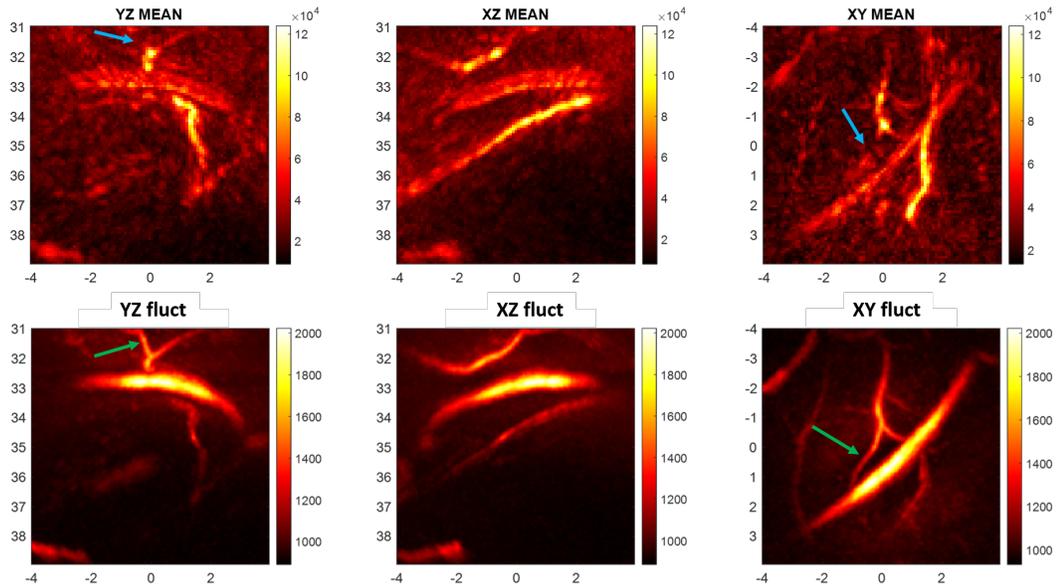


Figure 2.27: Result 4. The "Limited view" problem has been solved too. Blue flashes point the absent vertical structure which became visible and well defined after SVD.

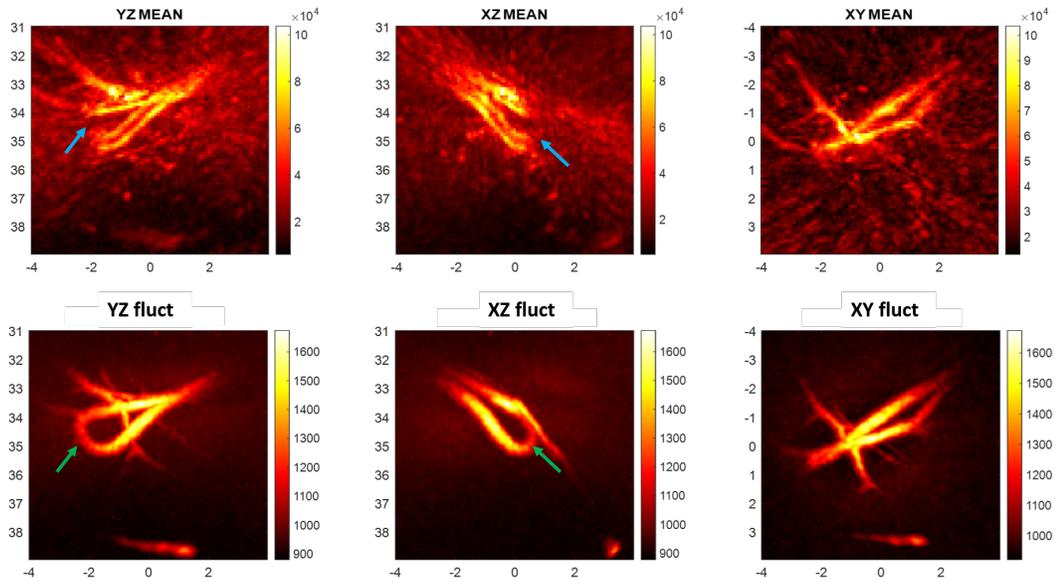


Figure 2.28: Result 5. As in Fig.2.27, green flashes show the correction of the visibility problem that did not allowed to see the entire structure.

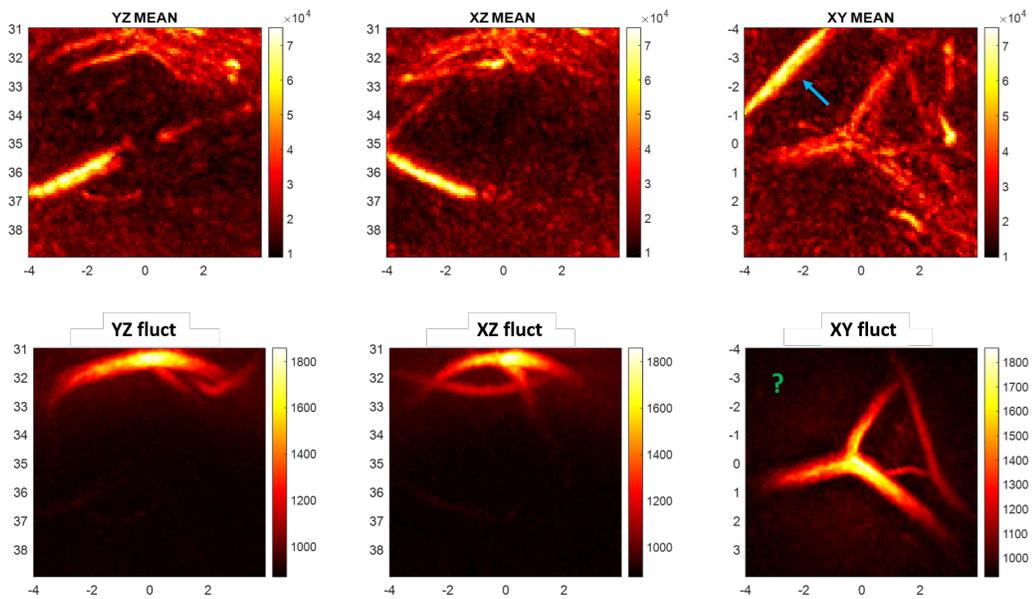


Figure 2.29: Result 6. Error in reconstruction: absorbing structure pointed by the blue flash that disappeared in fluctuation PA image.

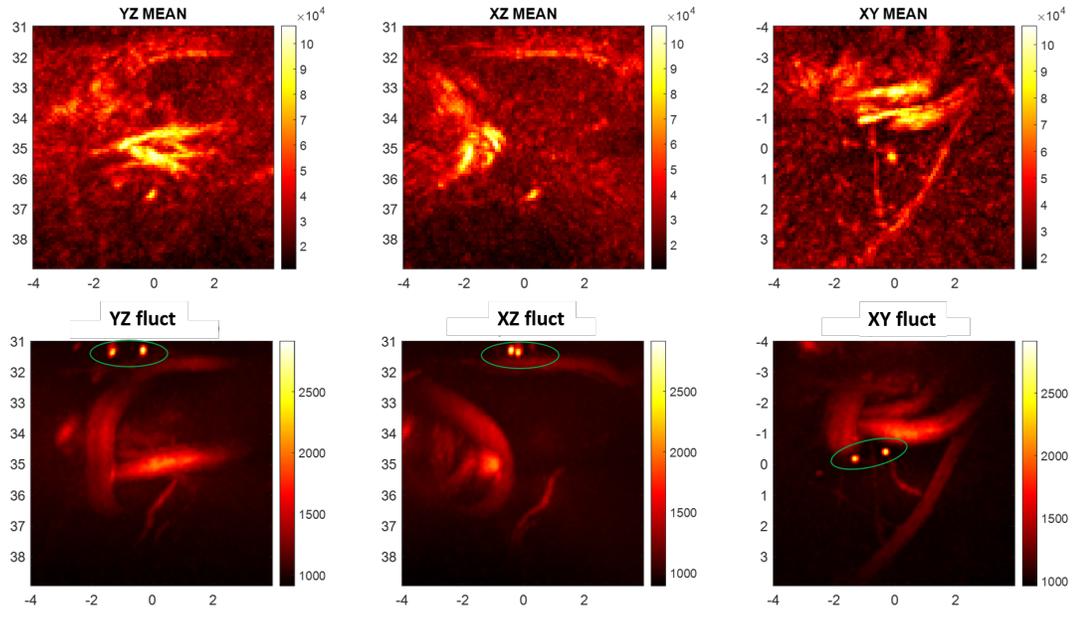
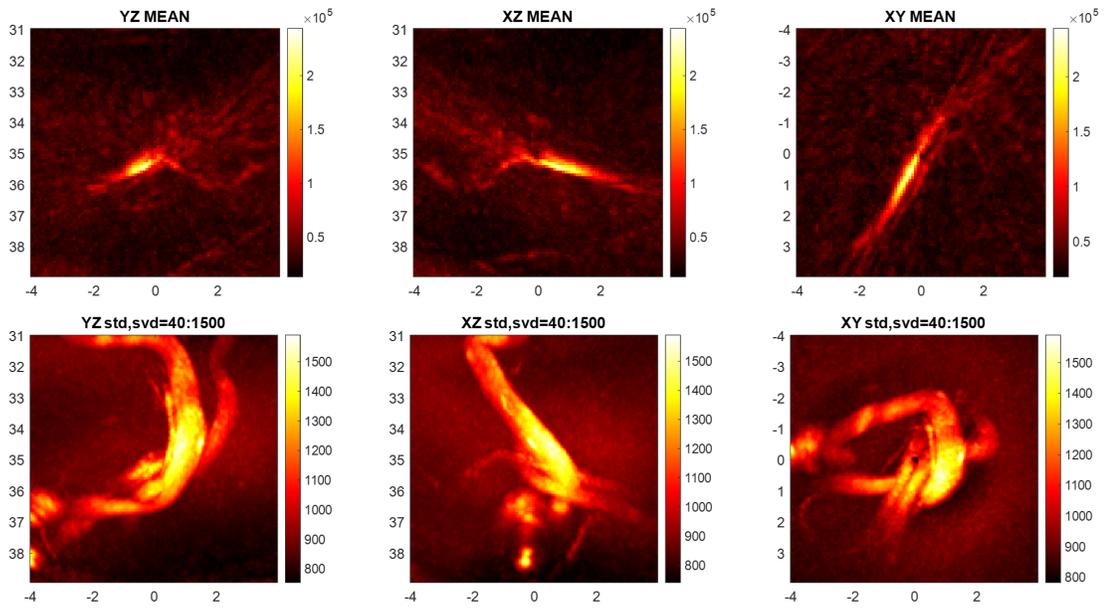


Figure 2.30: Result 7. Presence of bright point that absorb more than the other important structures and do not allow to have a good visibility of them.



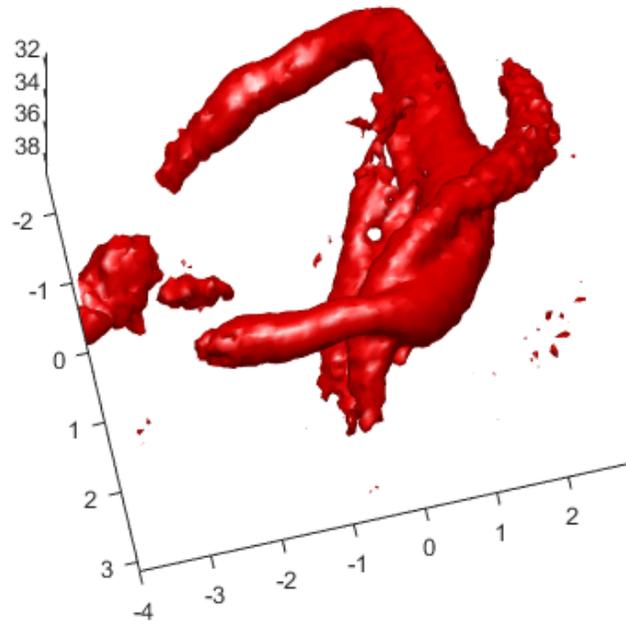


Figure 2.31: Result 8. 3D Visualisation

3. DEEP LEARNING APPROACH

3.1 Introduction

In the previous chapter we showed the first approach used to correct artifacts of visibility, namely photoacoustic fluctuation imaging. Here, we present the second approach used, which responds to the limit of the fluctuation imaging : the acquisition time. As shown, in fact, it needs a large number of images to obtain a correct variance estimator which forcibly limits the temporal resolution. On the contrary, the deep learning approach allows to correct artefacts from a single acquisition. First of all, machine learning is the use and development of computer systems to simulate human learning activities and to continuously improve the performance by studying self-improvement methods to obtain new knowledge. Deep learning is a type of machine learning based on artificial neural networks in which multiple layers of processing are used to extract progressively higher level features from data. The basic unit of a neural network is the neuron and the simplest type of neural network, composed by a single layer is the perceptron. It consists of four main parts including input values, weights, sum, and an activation function. The process begins by taking all the input values and multiplying them by their weights. Then, all of these multiplied values are added together to create the weighted sum. The weighted sum is then applied to the activation function, producing the perceptron's output. As a simplified form of a neural network, perceptrons allows to model only linear phenomena, playing an important role in binary classification.

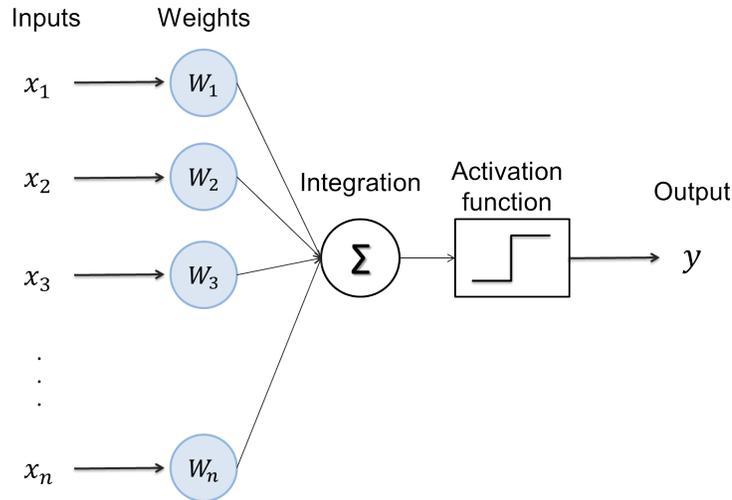


Figure 3.1: Perceptron architecture and process. X is the input vector, the W values represent the weights. An integration and a linear activate function are applied on the product between X and W to obtain the output value.

A succession of neurons bound together in series and parallel through their inputs and their outputs create a neural network. For example, in a neural network of three layers we see that the neurons of the second layer are connected to those of the first layer, that is, the inputs of the neurons of the second layer correspond to the outputs of the neurons of the first layer. Provided that we use non-linear activation functions (such as a sigmoid) we are able to model all nonlinear phenomena. The goal of the training is to determine the weights W to obtain the final output (Y_{final}). In our case, we use the so-called "supervised learning", where we have a reference data (Y_{true}). The training is divided into two phases:

- **Forward propagation:** with each iteration of the training we provide the input data that propagate in the network and we get Y_{final} .
- **Backward propagation:** the error realized by the network (difference between the output obtained Y_{final} and the desired output Y_{true} , called "cost function") is calculated. The purpose of this training phase is to calculate how this error varies according to network parameters. A gradient chain is then calculated and from these and

the parameters W are update to minimize the cost function, thanks to the following gradient descent formula:

$$W = W - \alpha \frac{\partial \text{Error}}{\partial W} \quad (3.1)$$

The whole procedure is shown in Fig.3.2 and 3.3.

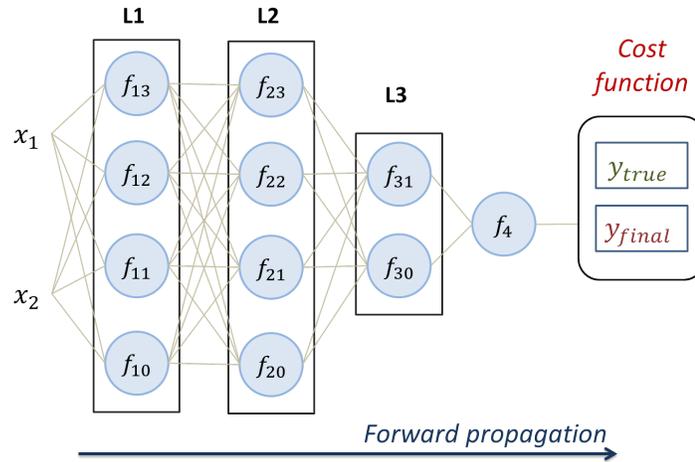


Figure 3.2: Forward propagation process.

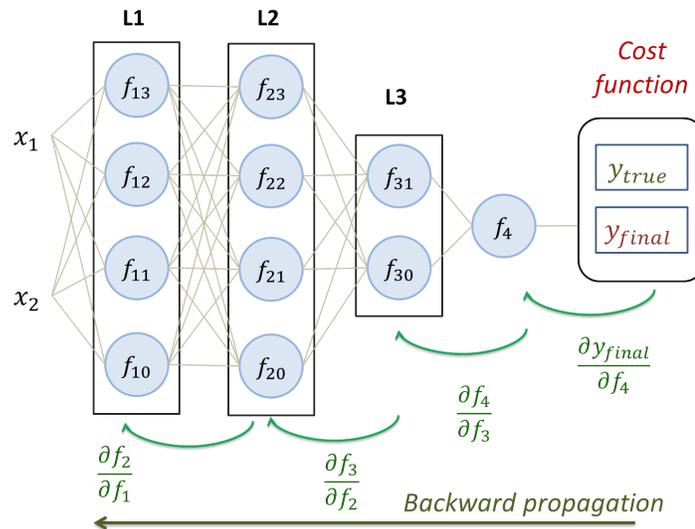


Figure 3.3: Backward propagation process.

This process is repeated as many times as necessary until the network converges to the desired output. Once the network converges, the training phase is over. The next one is the test phase during whose goal is to test the network with data that was not provided during the training.

3.1.1 State of the art - 2D results

With regard to the use of deep learning in photoacoustics, the first publication dates back to 2017 and most of the studies were done on a simulation dataset and then tested on experimental data. In general, there was no ground truth present so it was impossible to know whether the network reconstructed the image well or not. In 2020, Guillaume Godefroy et al. [29] proposed a concrete experimental demonstration of the deep learning approach to the 2D case, starting from an in vitro dataset for a perfectly known object and showing the uncertainty of prediction with an approach called Monte Carlo dropout. Skeletons of tree leaves whose capillaries were filled with black ink and limbo removed with a chemical treatment, were imaged with a linear probe. As in our 3D study, the images were reconstructed with the delay-and-sum beamforming algorithm and a photograph of the leaves was used as ground truth. Thus, the training set consisted of the conventional PA image (input) and the photograph (ground truth). Using a convolution network based on the UNET architecture, it was found that the deep learning approach provided a strongly improved image, with reconstructed vertical structures and correct vessel thicknesses, as opposed to the conventional image. To give an idea of this improvement, a similarity index was calculated on the whole of the test data: the deep learning approach increased by a factor of 2 such similarity, even if errors remained, such as invented or poorly reconstructed structures.

3.2 Methodology

3.2.1 Network architecture

In image processing, the type of network normally used is the Convolutional neural network (CNN). The inputs are the pixels of the image and instead of vectoring the image and then multiplying each pixel by the parameter W , the idea is to use a convolution, preserving the spatial dependencies between the pixels that we would have lost by vectoring. Each neuron of the CNN corresponds to a convolution operation. Feature detectors or filters help to identify different features present in an image like edges, vertical lines, horizontal lines, bends, etc. such as the edges of the image, a variation in coloring etc. These feature maps will be the input data of the neuron of the next layer. Pooling is then applied over the feature maps: the features are divided into sub-zones of 2×2 pixels and in general the pixel with the maximum value is kept to decrease the number of data and accelerate the calculations and extract the most important information.

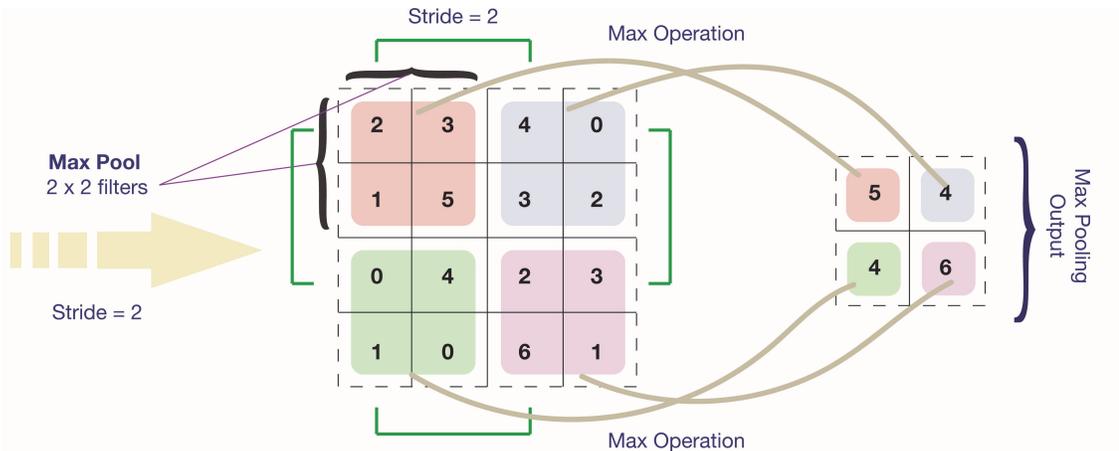


Figure 3.4: Max pooling operation applied on feature maps to decrease the number of data and accelerate the computation time. [30]

Among the types of CNN networks, the architecture that we used was a slightly modified Unet, a well known network architecture first developed for segmentation tasks. It is

a convolutional neural network composed of two paths:

- The contracting path, called the *the encoder*, a traditional stack of convolutional and pooling layers where the network extracts more and more complex features.
- The expanding path, called the *the decoder*, a symmetric expanding path where pooling operations are replaced by upsampling operators to recover at the output the resolution of the input.

Dropout layers are added to this architecture. Dropout is a regularization technique to limit overfitting, implemented by randomly disabling some neurons of the network at each training step. This prevents units from co-adapting too much and forces the network to learn more robust features. [32] Similarly as dropout, batch normalization [31] normalizes the output of the previous activation layer by subtracting the batch mean and dividing by the batch standard deviation. It helps to speed up the learning and also reduces overfitting by adding noise. The last layer contained only one filter instead of two in the original version, as the expected output is a single image. The last activation function was also suppressed as the prediction is no longer a binary image. The cost function was the classical mean squared error, and an Adam optimizer was used with a learning rate of and momentum of 8 [33] with batch sizes of 8 images. An early stopping approach based on the validation loss was chosen to limit under- and over-fitting [34].

3.2.2 Network training

The training phase of our neural network was carried out thanks to the dataset of 140 acquisitions that includes, for each of them:

- conventional single shot PA image obtained by delay-and-sum beamforming algorithm from time signals (RF data) of chicken embryos;
- the post-SVD fluctuation image, used as ground-truth.

The total dataset was subsequently reduced to 105 examples (90 for the training-set, 10 for the validation-set and 5 for the test-set), eliminating those where the images were too noisy, corrupted by artifacts such as bubbles and other bright spots or where hemorrhages were evident that altered the correctness of the fluctuation image, as in the case of the examples in Fig. 2.29 and 2.30. Before the training, the average of the fluctuation image was subtracted to eliminate the background noise and it was chosen to use the real part of the conventional single shot image, which gave better results than the absolute value.

While at first the network was only trained on experimental data, at a later time it has been trained using a network already trained on simulated data (about 400 examples in input) that could help to decrease the size of the training set and the time of the same training of the network.

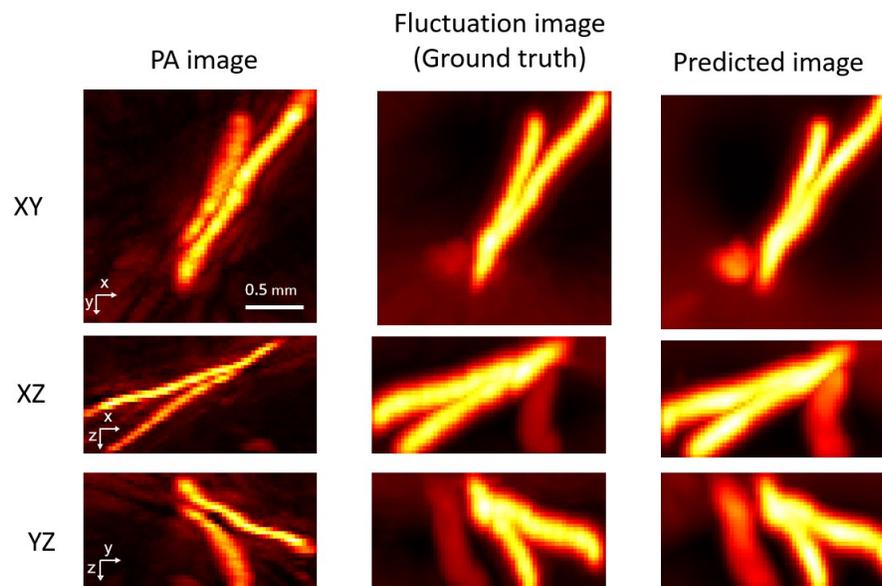


Figure 3.5: Example of deep learning application on simulated data: the first column represent the MIPs of the conventional image are showed, the second column the MIPs of the simulated ground truth and the third one shows the results of the accurate deep learning prediction. [G.Godefroy]

3.3 Network performances - Results

First of all, results obtained with the pretraining of the neural network with the simulation data made the process faster and the metric score value (similarity between output and desired output, normalized value) greater, as shown in the following table:

Training	Number of epochs	Metric Score
Without pre-training	300	0.69
With pre-training	215	0.81

After training the network under these conditions, to evaluate the accuracy of a trained neural network, a normalized cross-correlation (NCC) was computed between each output and the ground truth. It uses local sum to normalize the cross-correlation for feature matching. At first, we focused on the training part, to make sure that the network could really learn the information passed in input and could well predict the output from the conventional image that enormously differed from the fluctuation image given in input as ground truth in some cases.

Testing on the training-set

Below, there are four examples of output predicted by testing the network on some already trained inputs. We can see in all of them that we found a very high index of correlation (greater than 80%); the model learned to correct the artifacts and to extrapolate some structures not visible in conventional PA images. In addition, the predicted images present good contrast and total lack of noise.

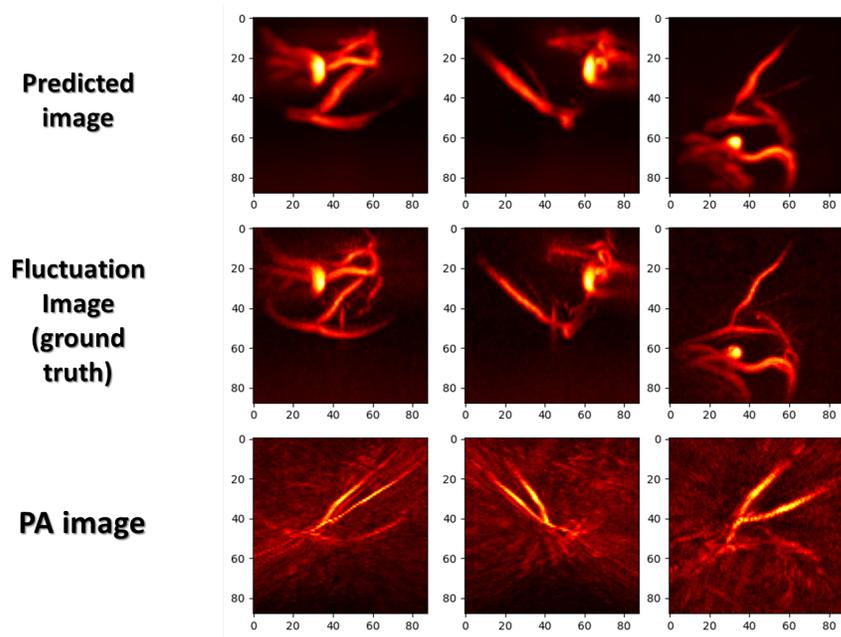


Figure 3.6: Result 1. MIPs of the input, ground truth et output on the three axes. Correlation between ground truth and output: 86%.

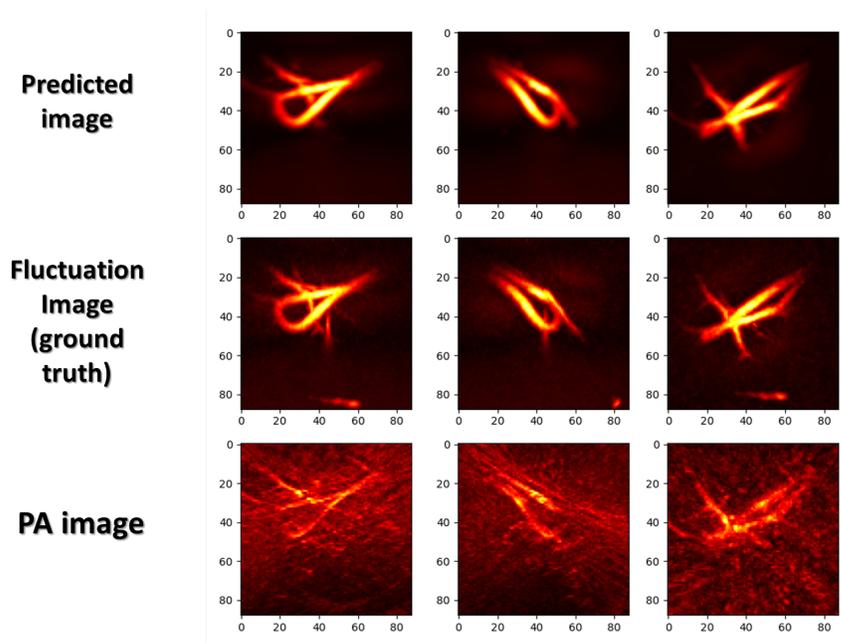


Figure 3.7: Result 2. Correlation between ground truth and output: 89%.

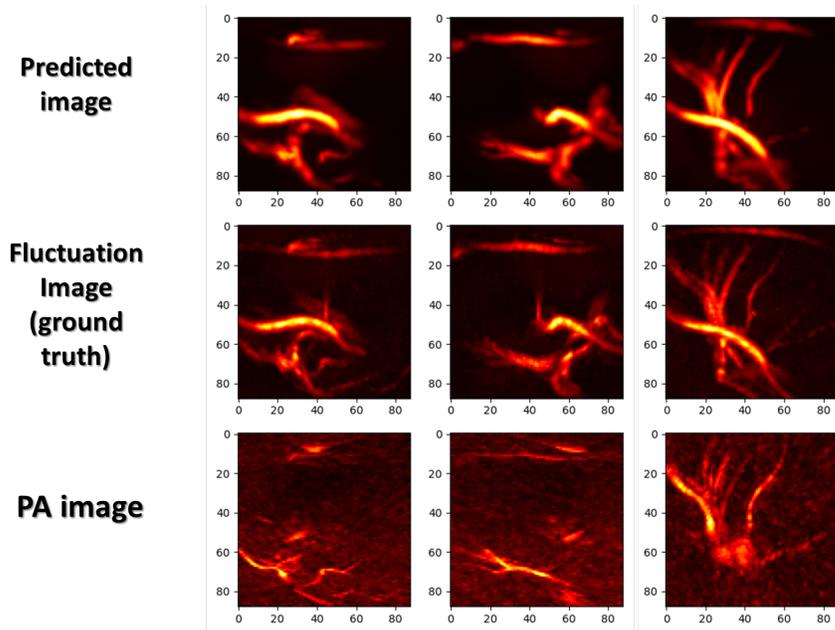


Figure 3.8: Result 3. Correlation between ground truth and output: 87%.

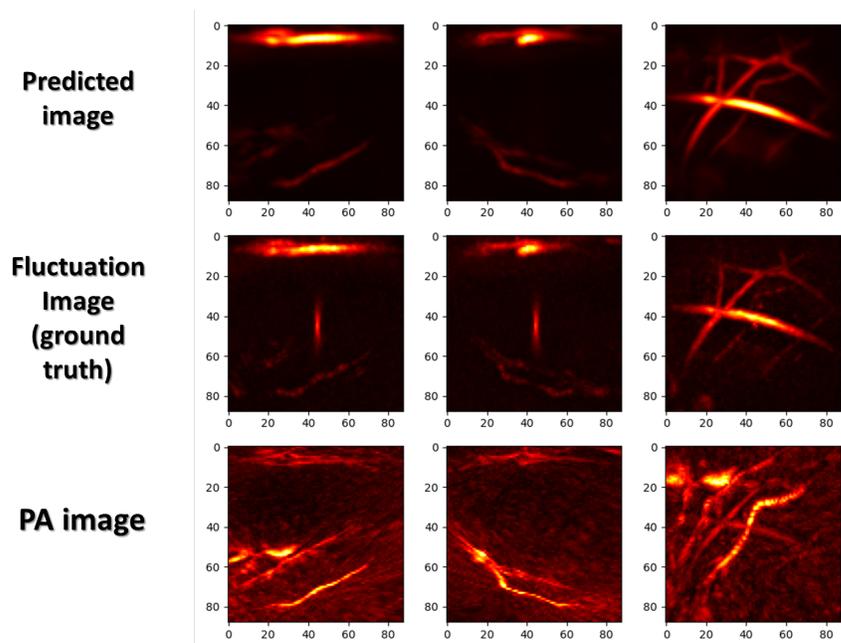


Figure 3.9: Result 4. Correlation between ground truth and output: 84 %.

Testing on the test-set

Subsequently, network performances were evaluated by testing data that were not part of the training set, therefore totally new. The test-set was composed by 5 examples. In this case, it is evident that the network is able to completely eliminate the noise of the conventional image, to greatly increase the contrast and to reconstruct the volume of the vessels despite the obvious errors, as explained in detail in the following results.

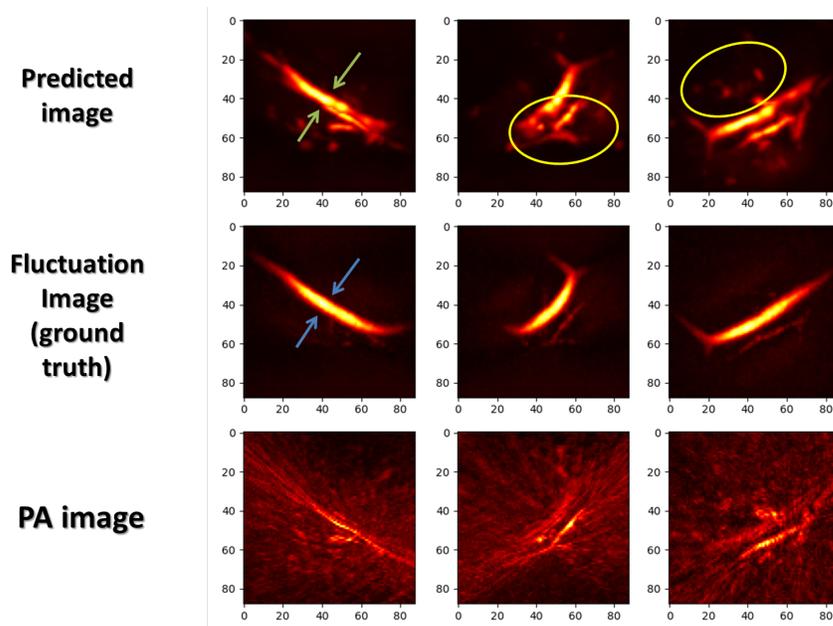


Figure 3.10: Result 1. Prediction of the network tested on an example not known (out of the training set). We note a contrast enhancement and a good reconstruction of the internal part of the vessel (blues fleshes in fluctuation image and green fleshes in predicted image). Therefore, compared to the ground truth, there are some not existent structures that appear on the output (surrounded by the yellow circle).

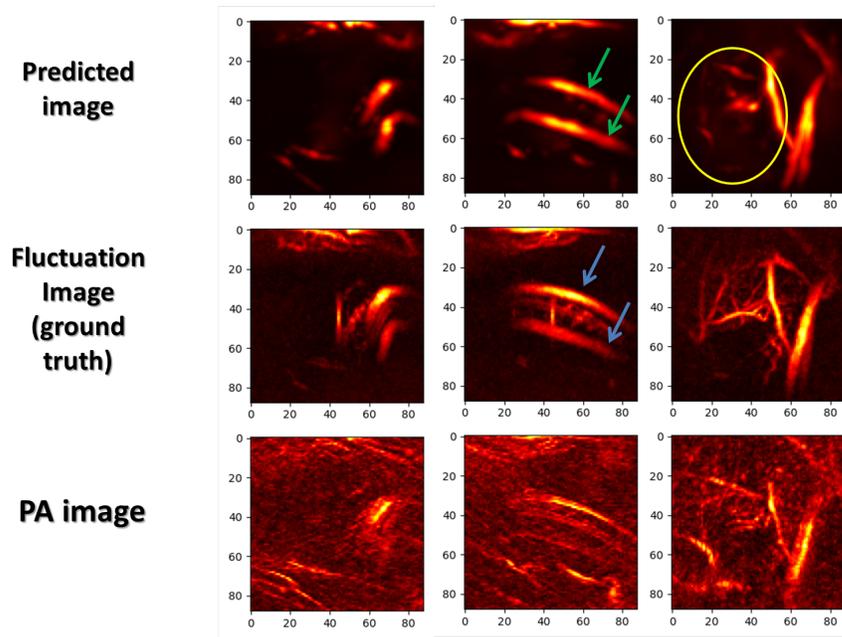


Figure 3.11: Result 2. As in the previous example, vessels are well reconstructed (blue and green fleshes) but the predicted image presents a poorly defined microvasculature that was well resolved in the fluctuation image (yellow circle).

4. CONCLUSIONS AND FUTURE EXPERIMENTS

The possibility of removing visibility artefacts with a neural network has been demonstrated in experiments on a class of complex 3D objects. Vertical parts of objects and the inside of large structures, missing on conventional reconstruction approaches, are recovered. However, some errors are still present in reconstructed images, such as invented or poorly reconstructed structures as well as missing structures. To avoid these problems, we have multiple ideas to test in the future. First of all, it might be useful to improve the quality of fluctuation image to get an optimized ground truth. In fact, the fluctuation image of the chicken embryo, being experimental compared to the 2D case where a photograph was used, presents some artefacts and a background noise that make the learning more difficult. It will be even harder having a good quality of the ground truth in the case of more complex in-vivo structures. For example, in the last part of the project, we experimented our acquisition system and processing on a mouse brain. The mouse was anesthetized and deposited on a hot plate, its snout inserted into a tube connected to the gas, and a rectal thermometer was used to monitor its temperature. Ultrasound gel was deposited on the area to be imaged, then contact with the probe was performed. In that case, because of tissue attenuation and movement artefacts, structures in fluctuation image had a lower intensity and the noise did not permit to have a clear reconstruction to use as input in our network. An idea could be, for example, to make an estimation of the noise value to eliminate from the image. By contrast, in our tests, only its average was removed. Another improvement could be to increase the number of simulated and experimental data, as well as to implement network architectures other than CNN.

However, with all necessary improvements, this project lays the bases for the joint use of PAFI (Photoacoustic fluctuation imaging) and deep learning algorithms for in-vivo 3D quantitative imaging of vasculature and blood oxygenation.

Bibliography

- [1] Srirang Manohar and Daniel Razansky, "Photoacoustics: a historical review," *Adv. Opt. Photon.* 8, 586-617 (2016)
- [2] Rayleigh, "The photophone," *Nature* 23, 274–275 (1881).
- [3] W. H. Preece, "On the conversion of radiant energy into sonorous vibrations," *Proc. R. Soc. London* 31, 506–520 (1880).
- [4] W. C. Röntgen, "Ueber Töne, welche durch intermittirende Bestrahlung eines Gases entstehen," *Ann. Phys.* 248, 155–159 (1881).
- [5] J. Tyndall, "Action of an intermittent beam of radiant heat upon gaseous matter," *Proc. R. Soc. London* 31, 307–317 (1880).
- [6] A. Rosencwaig, "Photoacoustics and Photoacoustic Spectroscopy" (Wiley, 1980).
- [7] M. L. Veingerov, "A method of gas analysis based on the Tyndall-Röntgen optico-acoustic effect," *Dokl. Akad. Nauk SSSR* 19, 687–688 (1938).
- [8] A. H. Pfund, "Atmospheric contamination," *Science* 90, 326–327 (1939)
- [9] K. F. Luft, "Infrared techniques for the measurement of carbon monoxide," *Ann. Occup. Hyg.* 18, 45–51 (1975)
- [10] L. Amar, M. Bruma, P. Desvignes, M. Leblanc, G. Perdriel, and M. Velghe, "Detection, d'ondes élastiques (ultrasonores) sur l'os occipital, induites par impulsions laser dans l'oeil d'un lapin," *C. R. Acad. Sci. Paris* 259, 3653–3655 (1964)
- [11] T. Bowen, "Radiation-induced thermoacoustic soft tissue imaging," in *Proceedings 1981 Ultrasonics Symposium (IEEE, 1981)*
- [12] P. Beard, "Biomedical photoacoustic imaging," *Interf. Focus* 1(4), 602–631 (2011)

- [13] Gateau, Jérôme, et al. "Improving visibility in photoacoustic imaging using dynamic speckle illumination." *Optics letters* 38.23 (2013): 5188-5191
- [14] D. Yang, D. Xing, S. Yang, and L. Xiang, "Fast full-view photoacoustic imaging by combined scanning with a linear transducer array," *Opt. Express* 15, 15566–15575 (2007)
- [15] R. A. Kruger, W. L. Kiser, Jr., D. R. Reinecke, and G. A. Kruger, "Thermoacoustic computed tomography using a conventional linear transducer array," *Med. Phys.* 30, 856–860 (2003).
- [16] B. Huang, J. Xia, K. Maslov, I, and L. V. Wang, "Improving limited-view photoacoustic tomography with an acoustic reflector," *J. Biomed. Opt.* 18, 110505 (2013)
- [17] G. Li, J. Xia, K. Wang, K. Maslov, M. A. Anastasio, and L. V. Wang, "Tripling the detection view of high-frequency linear-array-based photoacoustic computed tomography by using two planar acoustic reflectors," *Quantum Imaging Med. Surg.* 5, 57–62 (2015).
- [18] B. Cox and P. Beard, "Photoacoustic tomography with a single detector in a reverberant cavity," *J. Acoust. Soc. Am.* 125, 1426–1436 (2009)
- [19] X. L. Dean-Ben and D. Razansky, "Localization optoacoustic tomography," *Light. Sci. Appl.* 7, 18004 (2018)
- [20] X. L. Dean-Ben, L. Ding, and D. Razansky, "Dynamic particle enhancement in limited-view optoacoustic tomography," *Opt. Lett.* 42, 827–830 (2017)
- [21] L. Wang, G. Li, J. Xia, and L. V. Wang, "Ultrasonic-heating-encoded photoacoustic tomography with virtually augmented detection view," *Optica* 2, 307–312 (2015)
- [22] J. Gateau, T. Chaigne, O. Katz, S. Gigan, and E. Bossy, "Improving visibility in photoacoustic imaging using dynamic speckle illumination," *Opt. Lett.* 38, 5188–5191 (2013)

- [23] Vilov, S., Godefroy, G., Arnal, B., and Bossy, E. (2020). Photoacoustic fluctuation imaging: theory and application to blood flow imaging. *Optica*, 7(11), 1495-1505
- [24] Nowak-Sliwinska, P., Segura, T., and Iruela-Arispe, M. L. (2014). The chicken chorioallantoic membrane model in biology, medicine and bioengineering. *Angiogenesis*, 17(4), 779-804
- [25] Sergey Vilov. "Photoacoustic imaging beyond the acoustic diffraction limit". Instrumentation and Detectors [physics.ins-det]. Université Grenoble Alpes, 2019
- [26] T. Chaigne, J. Gateau, M. Allain, O. Katz, S. Gigan, A. Sentenac, and E. Bossy, "Super-resolution photoacoustic fluctuation imaging with multiple speckle illumination," *Optica* 3, 54–57 (2016)
- [27] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)," *Proc. Natl. Acad. Sci. USA* 106, 22287–22292 (2009)
- [28] Wall, Michael E., Andreas Rechtsteiner, and Luis M. Rocha. "Singular value decomposition and principal component analysis." A practical approach to microarray data analysis. Springer, Boston, MA, 2003. 91-109
- [29] Godefroy, Guillaume, Bastien Arnal, and Emmanuel Bossy. "Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties." *Photoacoustics* 21 (2021): 100218
- [30] Wall, M. E., Rechtsteiner, A., and Rocha, L. M. (2003). Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis* (pp. 91-109). Springer, Boston, MA
- [31] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." International conference on machine learning. PMLR, 2015

- [32] Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." *The journal of machine learning research* 15.1 (2014): 1929-1958
- [33] Kingma D.P., Ba J. Adam: "A method for stochastic optimization" (2014)
- [34] Prechelt L. "Early stopping-but when?" *Neural Networks: Tricks of the Trade*, Springer (1998), pp. 55-69