ORIGINAL ARTICLE

A new generative approach for optical coherence tomography data scarcity: unpaired mutual conversion between scanning presets

Mateo Gende[1,2] · Joaquim de Moura[1,2] · Jorge Novo[1,2] · Manuel G. Penedo[1,2] · Marcos Ortega[1,2]

Abstract

In optical coherence tomography (OCT), there is a trade-off between the scanning time and image quality, leading to a scarcity of high quality data. OCT platforms provide different scanning presets, producing visually distinct images, limiting their compatibility. In this work, a fully automatic methodology for the unpaired visual conversion of the two most prevalent scanning presets is proposed. Using contrastive unpaired translation generative adversarial architectures, low quality images acquired with the faster *Macular Cube* preset can be converted to the visual style of high visibility *Seven Lines* scans and vice-versa. This modifies the visual appearance of the OCT images generated by each preset while preserving natural tissue structure. The quality of original and synthetic generated images was compared using BRISQUE. The synthetic generated images achieved very similar scores to original images of their target preset. The generative models were validated in automatic and expert separability tests. These models demonstrated they were able to replicate the genuine look of the original images. This methodology has the potential to create multi-preset datasets with which to train robust computer-aided diagnosis systems by exposing them to the visual features of different presets they may encounter in real clinical scenarios without having to obtain additional data.

Keywords Optical coherence tomography · Generative adversarial network · Style transfer · Synthetic images

1 Introduction

Optical coherence tomography (OCT) is a non-invasive medical imaging technique that can generate cross-sectional, 3-dimensional images of ocular tissue at a micrometre resolution [1]. Thanks to the advances in signal processing, optics and electronics, the quality and resolution of the images obtained through OCT have steadily improved through the years, leading to its widespread use as a diagnostic tool [2], with around 30 million OCT scanning procedures being performed every year worldwide [3]. For reference, OCT imaging has been used to study relevant ocular pathologies such as diabetic macular edema, the most common cause of blindness in patients of diabetes mellitus [4–6]; glaucoma, the leading cause of irreversible blindness worldwide [7–9] and age-related macular degeneration,

the leading cause of blindness in people over 50 in the developed world [10–12]; as well as to study the vascular structure of the eye [13–16].

In an OCT scanning session, a low coherence optical beam is swept through the retina of the patient, generating two-dimensional images or B-scans. Due to the forward and backward scattering of light waves, these B-scans present speckle noise, the main quality affecting factor in the OCT images [17]. OCT scanners can be configured to combine and average several B-scans over the same location, leading to a reduction in noise and an overall increase in tissue visibility and quality of details. Nonetheless, this method requires that the tissue does not move throughout the scanning process, to be able to average readings taken at the same point. This requirement limits the overall amount of scans that can be performed in a session due to involuntary eye movements. This translates into a constraint between the total surface of tissue that can be analysed in a given time and the quality of the obtained tomograms.

OCT scanner platforms typically provide a series of configuration presets. This way, the specialists can choose which type of scan to be performed according to whether they need to sacrifice quality in order to be able to scan

✉ Joaquim de Moura
  joaquim.demoura@udc.es

Extended author information available on the last page of the article.

a broader surface of tissue, or whether they can afford the extra time required to perform a higher quality scan. While many presets exist, with different sweeping patterns such as radial or annular scanning, the two most widely used by expert clinicians in medical services are volume scan patterns. The first one scans a square-shaped section of the eye fundus, averaging a small amount of B-scans per final tomogram or image slice. This results in a great number of slices over a wide patch of the retina presenting speckle noise. This scanning preset is known as *Fast Scan* in Heidelberg SPECTRALIS® platforms or *Macular Cube* in Carl Zeiss CIRRUS-HD® models, the two most common OCT imaging platforms in clinical settings, and will be referred to as *Macular Cube* in this manuscript. The second one, known as *Seven Lines* in SPECTRALIS® platforms or *Five Line* in CIRRUS-HD®, is a more intensive scan. In this configuration, a thinner band of the retina is scanned averaging many B-scans per slice. While this scan produces only a few slices over a narrow strip of tissue, these provide much higher visibility, texture detail and image resolution, resulting in images that are much clearer and easier to analyse than the aforementioned preset. We will henceforth refer to this type of scanning preset as *Seven Lines*. An example of the sweeping pattern and a resulting slice for each of these two most representative scanning presets is presented in Fig. 1.

While these presets are the two most prevalent in medical services, this trade-off between scanning time and image quality affects any other configuration, constituting a common paradigm in OCT imaging. Consequently, this leads to a forced choice between the quality of the produced images and the amount of tissue that can be analysed, which therefore results in a shortage of high quality OCT images. Furthermore, the images that are produced by using different presets present visual differences that complicate the development of automatic computer-aided diagnosis (CAD) systems.

Recent years have seen an advancement both in medical imaging techniques as well as in computational architectures and algorithms, leading to the development of new and improved CAD systems based on deep learning [18, 19]. These systems can aid in the detection of several relevant pathologies while achieving equal or better results than board-certified specialists [20–23]. Nonetheless, the development of CAD systems based on machine learning requires well-curated data for training. The images in the training datasets should cover all the possible variabilities in imaging platforms, acquisition conditions, presets and configurations that are present in the clinical setting that the system is intended to work in, so that it can perform its diagnosis functions adequately under such conditions. The labour and economical costs associated with the acquisition
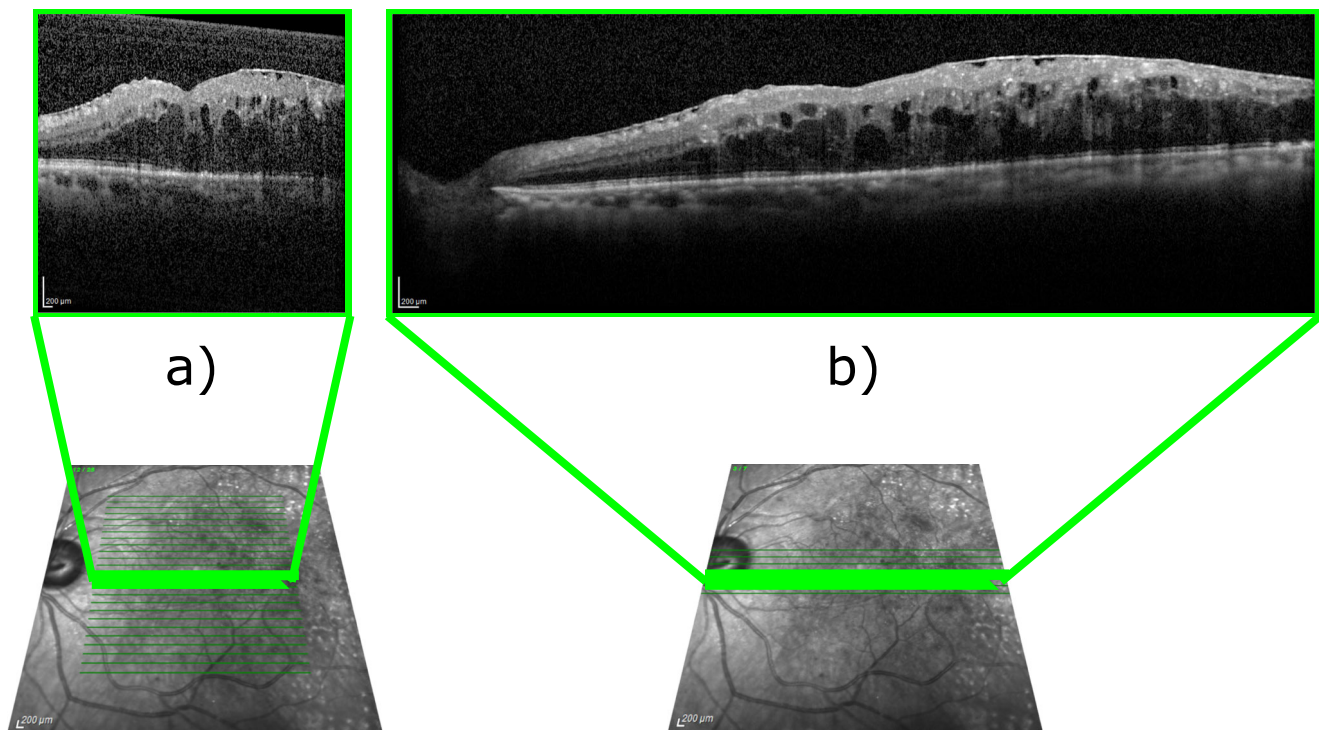

Fig. 1 Representative examples of OCT images taken from the same location using the two most common scanning presets. **a** OCT image acquired with the *Macular Cube* preset showing considerable speckle noise. **b** OCT image from the *Seven Lines* preset displaying lower speckle noise and increased tissue visibility

of these images, combined with their sensitive nature, lead to an aggravation of the problem of data scarcity, which affects all the domains of application of deep learning [24] and is, therefore, especially prevalent in medical imaging [20].

Given the relevance of this issue, some works have addressed the problem of improving the quality of OCT scans to facilitate the visual inspection of the images and their clinical diagnosis. In 2004, Adler et al. [25] first proposed the use of wavelet filters to reduce the speckle noise in the OCT images. Similarly, other authors continued to improve the image denoising by applying different statistical models such as Bayesian inference [26, 27], non-local means [28] or Huber total variation regularisation [29]. Other approaches expanded upon the wavelet filtering method by applying 3-D block matching techniques [30] or dictionary learning [31]. More recently, some works have approached this task by using deep learning. Apostolopoulos et al. [32] presented a study where they employ an artificial neural network to increase the contrast and reduce the noise in OCT images. Similarly, in the work of Xu et al. [33], a non-linear mapping convolutional neural network is used to perceptually enhance the images resulting in a reduction of the speckle noise. In the work of Seeböck et al. [34], the authors use a generative adversarial network (GAN) to reduce the variability between OCT images acquired with two devices, demonstrating performance gains when using the transformed images for the segmentation of retinal fluid. Lastly, the approach proposed by Huang et al. [35] uses a GAN to both remove the speckle noise that is present in the images and increase the image resolution. These works offer promising results concerning the speckle noise reduction, the perceptual quality enhancement or the increase of resolution of the OCT images, concerning visual inspection. Nevertheless, none of them has addressed the visual differences that exist between images acquired with different scanning presets and their relation to the problem of data scarcity for machine learning-based OCT CAD systems. This leaves the problem of training CAD systems with different presets to be addressed, because while speckle noise can be considered the main quality-affecting factor for OCT images, it is not the only visual difference between images acquired with multiple presets.

To mitigate this problem, this work presents a fully automatic methodology for the mutual unpaired conversion of OCT images that were acquired with different scanning presets. To do this, contrastive unpaired translation architectures are employed for the target conversion. The first approach presented in this work consists of training a model to translate the more numerous images acquired with a low-quality extensive scanning preset such as *Macular Cube* into the style of the higher visibility *Seven Lines* preset to

help mitigate the issue of high quality data scarcity in OCT datasets. A second, complementary approach was designed with the intention of performing the opposite translation, transferring the visual features of images obtained with the *Macular Cube* preset to original *Seven Lines* images. The images generated by these two approaches are assessed based on their quality in a complete methodology aimed at determining the optimal point at which the generative models are able to confer the intended visual features of the target preset. This methodology can not only increase the total available number of images by means of over-sampling, helping to mitigate the problem of data scarcity that is so prevalent in medical imaging, but it also has the potential to create multi-preset datasets that can be used to train CAD systems in a robust, variability tolerant manner. While this methodology is exemplified through the use of these two most common scanning presets in this work, it is also extensible to any other scanning preset or OCT imaging device as they are all affected by this compromise between scanned area in a given time and image quality. This way, deep learning models can be trained with the visual features of the various presets and acquisition conditions that it may encounter in its use in a clinical environment without the need to procure the otherwise scarce original images acquired with such presets. Preliminary results of this strategy were obtained in [36], demonstrating that this conversion approach can be suitable to address this problem of data scarcity.

2 Methods

In this section, the materials and resources that were used for the implementation of this work are covered, along with a description of the OCT image translation methodology. Specifically, the reader can find information on the dataset that was used (Section 2.1), the software and hardware resources (Section 2.2), a description of the image translation methodology (Section 2.3) and an explanation of the experiments conducted to validate the synthetic generated images (Section 2.4).

2.1 Dataset

Regarding the dataset, a total of 1034 OCT images, acquired using a Heidelberg SPECTRALIS® platform, were used. These images were obtained from 56 different patients participating in a study of diabetic macular edema in accordance with the Declaration of Helsinki, approved by the local Ethics Committee of Investigation from A Coruña/Ferrol (2014/437) the 24th of November, 2014. OCT image resolutions ranged from 511×495 to 1535×495 pixels. In total, 517 of the images were obtained using the

Fast, or *Macular Cube* preset. In this preset, the scanner sweeps a $20° \times 20°$ patch of the eye fundus, averaging 9 B-scans to form every tomogram, obtaining 25 slices per scan. The remaining 517 slices were acquired with the *Seven Lines* preset. This preset involves scanning a longer and thinner, $30° \times 5°$ eye fundus strip and using the average of 25 B-scans for each of the 7 produced tomograms per session. A representative example of the OCT images produced by these presets can be found in Fig. 1.

2.2 Software and hardware resources

In this work, the PyTorch [37] machine learning library (version 1.7.1) under Python (version 3.7.7) for convolutional neural network training and validation was employed. OpenCV [38] (version 3.4.8) and NumPy (version 1.15.0) were used for all image manipulation and processing requirements. Regarding the hardware, the training and validation process of the models was performed on a computer consisting of an NVIDIA GeForce GTX TITAN X GPU, an Intel Xeon E5-2640 CPU and 64 GB of RAM.

2.3 Methodology

To perform the conversion between OCT images, the process was modeled as a "style transfer" approach, in which a neural network attempts to confer the visual features of a target class to an original image while preserving the original structure. In particular, a contrastive unpaired translation generative adversarial network (CUT-GAN) [39] architecture was used for this purpose.

The typical GAN training method involves the use of a generative network G and an additional discriminative network D whose task is to determine if images belong to the target class from the training set $y \in Y$ or were synthesised by the generative network $\hat{y} = G(x)$, while the generative network trains to maximise the probability that D makes a mistake [40]. This way, the training procedure pushes the generator G to synthesise images that resemble the target class from the training set, by using an adversarial loss:

$$\mathcal{L}_{\text{GAN}}(X, Y) = \mathbb{E}_{y \sim Y} \log D(y) + \mathbb{E}_{x \sim X} \log (1 - D(G(x)))$$
(1)

A CUT-GAN architecture is intended for the unpaired translation of images from one domain to another. As such, the generative network G consists of an encoder G_{enc} and a decoder part G_{dec}, and its task is to transfer the characterising features of the target domain to original images $x \in X$ without modifying that which is common to both domains. This is achieved by adding a patchwise noise contrastive estimation loss [41] that takes advantage

of the ability of the encoder part of the generative network to capture domain-invariant features such as the location of the inner limiting membrane and the choroid, as well as that of the decoder part, which has the means to synthesise domain-specific features like tissue texture as well as speckle noise, or lack thereof.

To calculate this contrastive loss, a set of features is extracted from the output from a series of layers $l \in \{1, 2, ..., L\}$ of the generative encoder. These features are obtained by applying the encoder to patches of both the original $\{\mathbf{z}_l\}_L = \{H_l\left(G_{\text{enc}}^l(x)\right)\}_L$ and the synthesised images $\{\hat{\mathbf{z}}_l\}_L = \{H_l\left(G_{\text{enc}}^l(G(x))\right)\}_L$, passing them through a two-layer MLP network H_l. Specifically, a patch is extracted from a location s in the original image, along with the a patch extracted from the same location in the synthesised image, and a series of patches from other locations $S \setminus s$ in the original image. This patchwise loss is then calculated as a cross-entropy loss between the positive and negative examples:

$$\mathcal{L}_{\text{Patchwise}}(X) = \mathbb{E}_{x \sim X} \sum_{l=1}^{L} \sum_{s=1}^{S_l} \ell\left(\hat{z}_l^s, z_l^s, z_l^{S \setminus s}\right)$$

$$\ell(q, q^+, q^-) = -\log \left[\frac{\exp\left(q \cdot q^+/\tau\right)}{\exp\left(q \cdot q^+/\tau\right) + \sum_{n=1}^{N} \exp\left(q \cdot q_n^-/\tau\right)} \right], \quad (2)$$

where $\tau = 0.07$ serves as a temperature with which the distances between the query and the other examples are scaled, and $\{0, 4, 8, 12, 16\}$ are the layers selected for the contrastive loss. By penalising differences in the inner representation of the same image patch in both images, as well as similarities between patches extracted from different regions, the network is trained to preserve the anatomical structure of the eye while changing the visual appearance of the OCT images according to the target preset, as illustrated in Fig. 2.

In order to address the problem of the visual variability between images acquired with different OCT scanning processes, two complementary approaches were taken, which are detailed below:

1. **First approach: *Macular Cube* to *Seven Lines* Conversion:**
   The first approach was designed to address the issue of high quality data scarcity in OCT by perceptually increasing the quality of the more numerous *Macular Cube* preset scans. A CUT model was trained to transfer the higher visibility style of *Seven Lines* images to samples acquired with the *Macular Cube* preset, effectively conferring them the reduction in noise of the more intensive scanning preset, along with its more precise visual features. Through this approach, high quality data scarcity can be compensated by converting the more readily available images into the higher quality
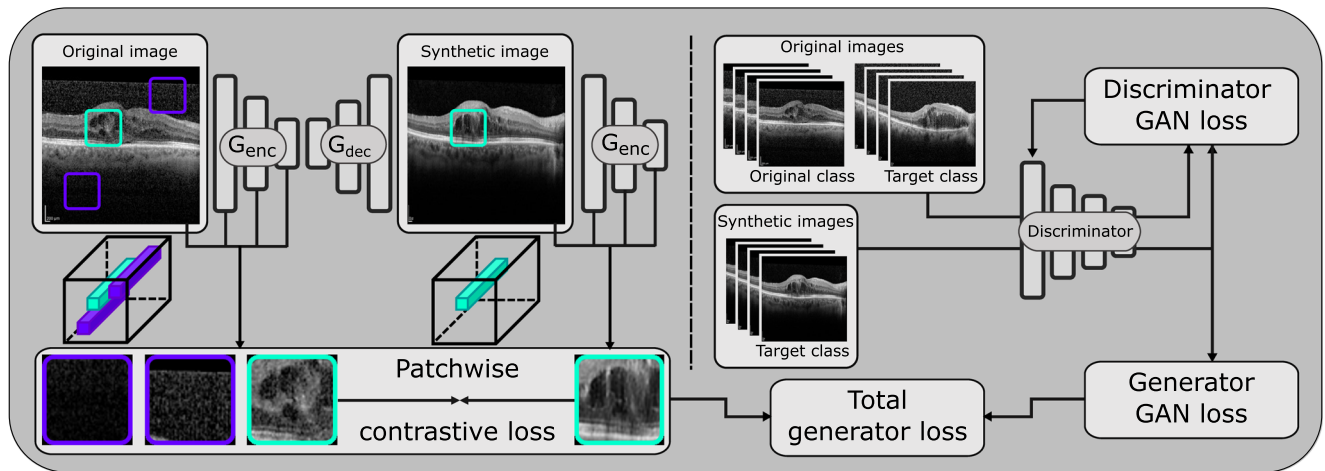
Fig. 2 Generative model training procedure: patchwise contrastive loss is calculated by comparing the inner representation of patches that were extracted from both the original and synthetic images. A discriminator is trained in parallel to discern between examples of the original and the target class, which can be used to compute the generator GAN loss

visual style of *Seven Lines* scans. An example of this conversion is illustrated in Fig. 3.

2. **Second approach: *Seven Lines* to *Macular Cube* Conversion:**

The second approach has the complementary purpose of converting higher quality *Seven Lines* scans into the style of the extensive *Macular Cube* preset. Therefore, a second CUT model was trained on the same data to perform the opposite translation. This results in an overall increase in the amount of available images. Moreover, these additional images can be used to train machine learning-based CAD systems in a preset variability-tolerant fashion by exposing them to the different presets that the system may find in a real clinical setting. An example of the visual features of this conversion can be found in Fig. 4.

For both of these approaches, a residual network [42] backbone with nine residual blocks was used as a base architecture for the CUT generator. The original images were resized to 286×286 pixels, with random crops of 256×256 being used as training inputs. During the training, the models were optimised using Adam [43] with $\beta_1 = 0.5$, $\beta_2 = 0.999$ and a learning rate of 2e−4. The training process lasted for a maximum of 400 epochs, linearly decaying the learning rate for the last 200. Finally, both of these models were used to generate the synthetic counterpart for every original image, as illustrated in Fig. 5.
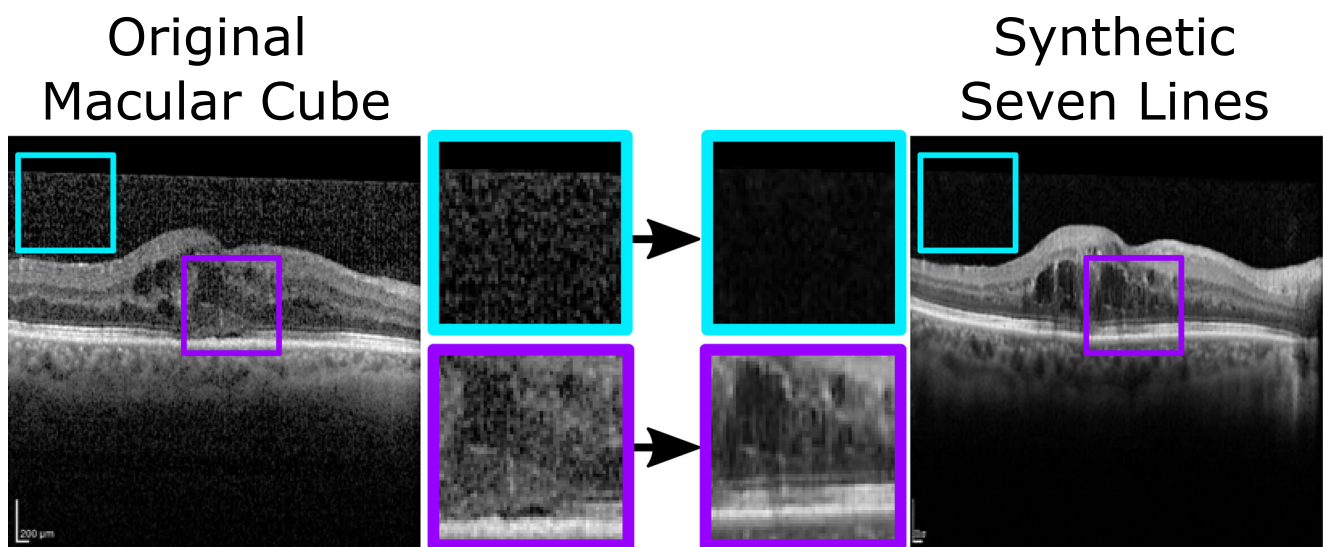

Fig. 3 First approach: *Macular Cube* to *Seven Lines* conversion. This conversion produces an overall reduction in the amount of speckle noise over the background and the retina, as well as enhanced tissue texture details in the synthetic generated image
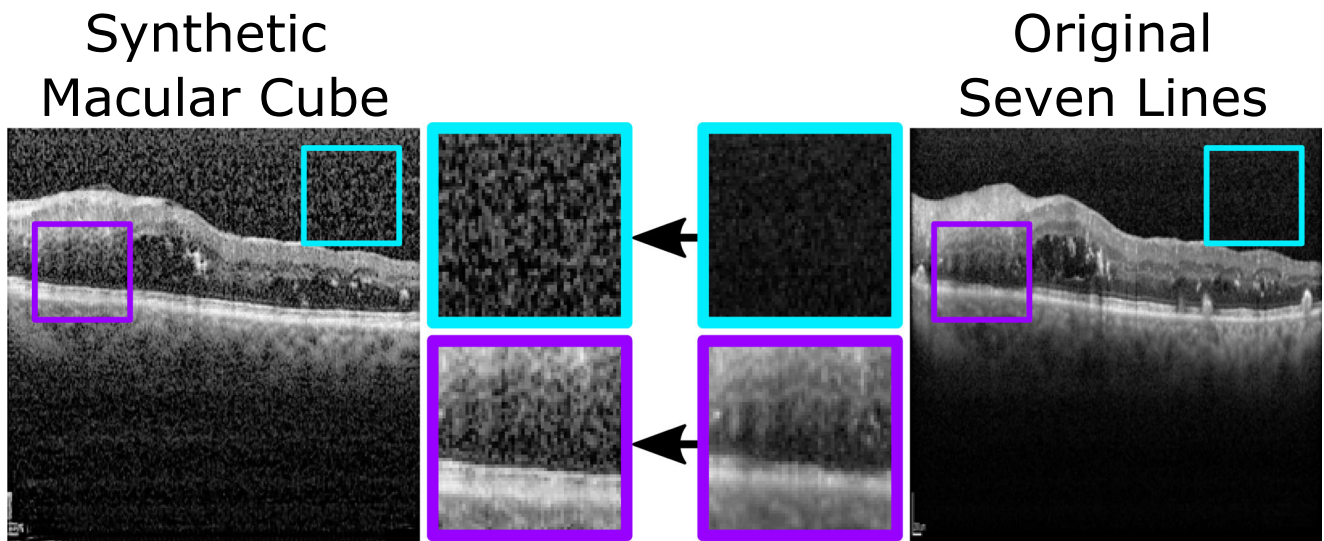
Fig. 4 Second approach: *Seven Lines* to *Macular Cube* conversion. In this translation, speckle noise visibility is increased and overall contrast is reduced while preserving the original tissue, producing synthetic generated images that visually resemble original *Macular Cube* scans

2.4 Evaluation

A series of experiments were conducted to evaluate the images that were generated by both of the approaches described above. This subsection describes how these experiments were carried out.

In order to validate the perceptual quality of the synthetic images, a qualitative experiment using the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [44] was performed. BRISQUE is an image quality evaluator which, unlike other measures such as Peak Signal to Noise Ratio or Structural Similarity Index

Measure [45, 46], does not require a reference image to compare. Instead, it returns a score indicative of the perceptual quality of the processed image, with lower BRISQUE scores indicating a higher image quality. To do this, a series of luminance coefficients are used to measure distortions and their orientations in the image. These are used to compute a series of features at multiple scales. Then, these features are classified and quantified by support vector machines. This way, the different distortions and their effect on image quality and perception can be measured. In this line, BRISQUE has been previously used to assess the quality of medical images with favourable results [47–49]. In this experiment,
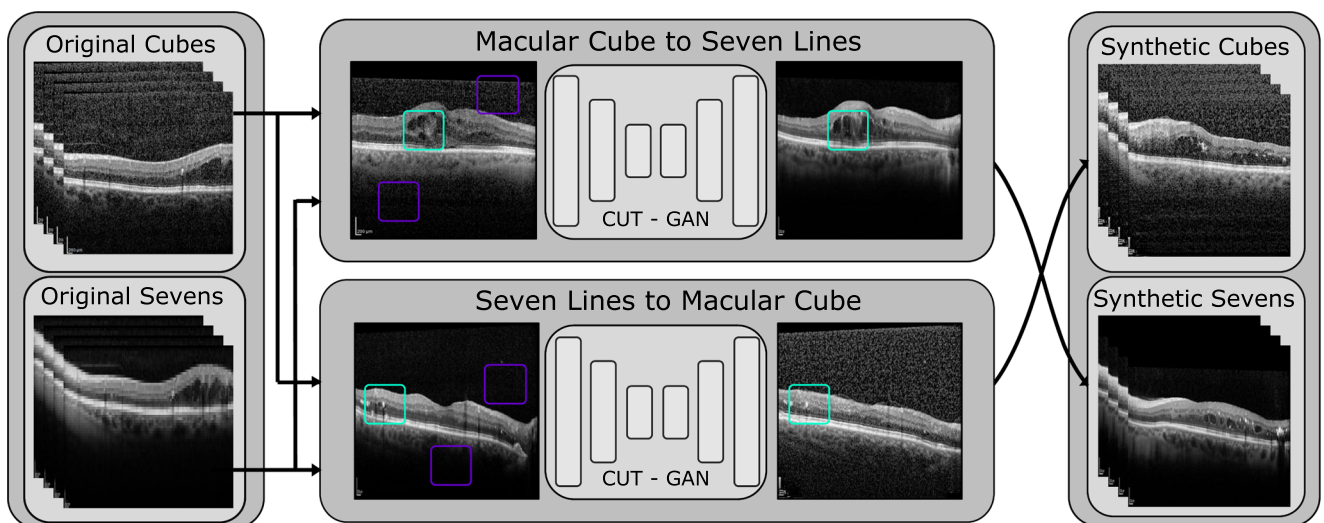

Fig. 5 Summary of the two computational approaches presented in this work. The original OCT images are used to train both conversion models. These models are then used to convert each original OCT image into a synthetic sample of the opposite class

the BRISQUE score of each set of the original *Macular Cube* and *Seven Lines* images was calculated and compared to those corresponding to the synthetic images.

Complementarily, a set of experiments aimed at measuring the perceptual quality of the synthetic generated images was conducted. The equivalent number of looks (ENL) and the contrast-to-noise ratio (CNR), calculated as $\text{ENL} = \frac{\bar{x}_{\text{BG}}^2}{s_{\text{BG}}^2}$ $\text{CNR} = \frac{\bar{x}_{\text{ROI}} - \bar{x}_{\text{BG}}}{s_{\text{BG}}}$, where $\bar{x}$ denotes the arithmetic mean and s denotes the standard deviation of the intensity values in the images, were used as referenceless image quality estimators. A random representative subset of 100 images of each class was annotated with the location of a homogeneous region of interest (ROI) and the background (BG) containing no tissue. This subset was employed to calculate these estimators. Furthermore, the referenceless Blind Image Quality Index (BIQI) [50] and the Natural Image Quality Evaluator (NIQE) [51] were measured for both the original and synthetic generated images. Section 3.2 covers the results of these experiments.

Subsequently, the separability of the generated images was assessed. While the training process of a GAN uses a discriminator network D to enforce the similarity between original and synthetic images of the target class, this is an intentionally simple architecture in order to avoid overpowering the generator, and it tends to be biased due to the GAN training process. To validate whether the images that were converted between the two presets display the visual features of their target presets, an automatic

separability experiment was conducted. In this experiment, an external classifier model was trained to classify between images acquired with the *Macular Cube* preset and images acquired with *Seven Lines*, using a subset of the original images. This network was then tested separately with the remaining original images and with the synthetic generated images. Afterwards, the results produced by the original images were compared with those of the synthetic images. The aim of this experiment is to determine whether the synthetic generated images are classified as their original class or their target class. A densely connected convolutional network [52] architecture was chosen to serve as the external classifier model. This architecture has seen extensive use in medical image classification and screening, surpassing other convolutional neural network architectures [53–56]. Figure 6 illustrates the structure of this model.

Since two separate generative models were trained, one for each approach, each class was tested separately. The Accuracy, calculated as $\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}}$, was used to evaluate this classification. When testing the two classes separately, positives are considered to be those images of the target class. Due to the absence of true negatives and false positives, the specificity cannot be computed for this test. The results of this experiment can be found in Section 3.3.

Finally, with the purpose to further assess the perceptual validity of these images, a test was conducted to ascertain whether the specialist clinicians are able to discern between the synthetic and original images of both classes. The
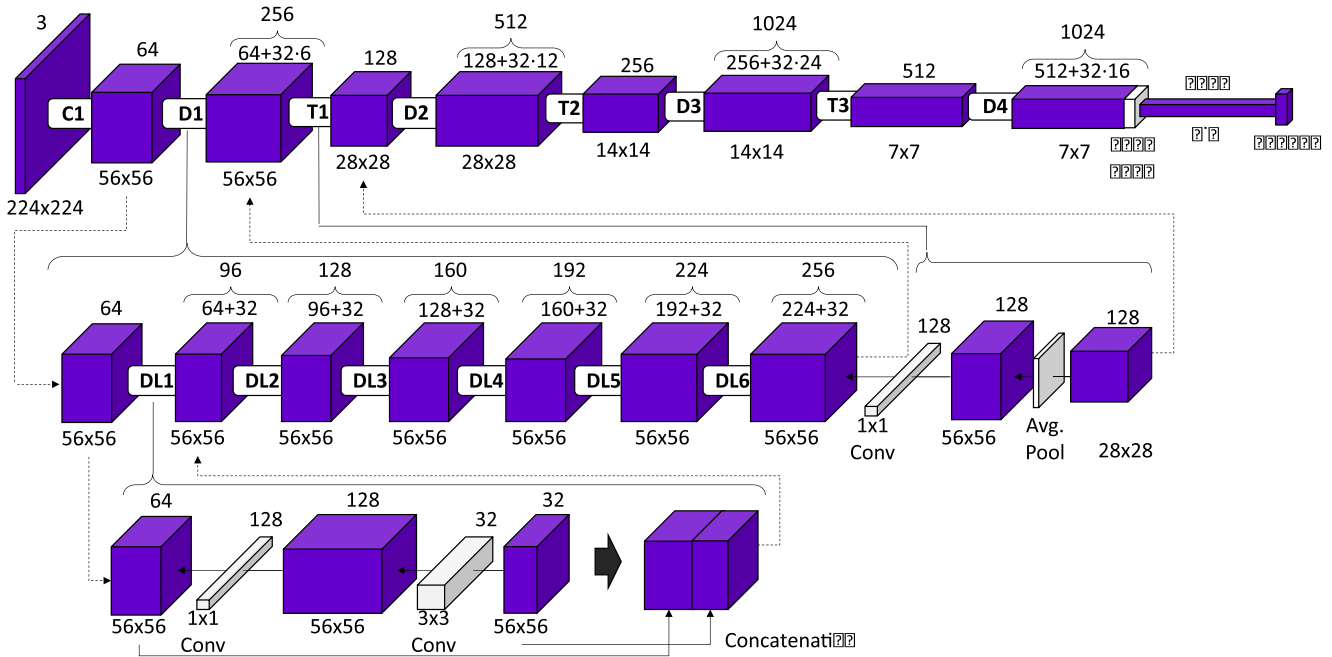

Fig. 6 DenseNet-121 model summary. First row shows the complete model. Second row illustrates a dense block structure with its corresponding transition layer. Third row details the composition of every dense layer and the concatenation of inputs and results

motivation behind this test was to assess whether these models can preserve the natural tissue structure of the eye, as well as to verify that artificial artefacts are not introduced in the synthetic generated images. The clinicians were asked to classify a random set of 200 images into whether they were acquired with the *Macular Cube* or the *Seven Lines* preset and if they were original images obtained with an OCT platform or they were generated by the network. The results of this experiment are discussed in Section 3.4.

3 Results

In this section, the results produced by the aforementioned synthetic image generation methodology are presented, along with those of the tests that were previously described.

3.1 Generative model training

The curves for both of the GAN losses, along with the contrastive losses for the training of the models are displayed in Fig. 7. These show the loss pattern that is often apparent in GAN training where both discriminator and generator losses tend to converge to a relatively stable value as the training progresses. This mutually dependent stability, however, complicates the task of determining the optimal stopping point of the training process. To work around this problem, both generative models were trained for up to 400 epochs, which was found to be a sufficient length for them to produce satisfactory results. During the training, a checkpoint copy of the generative network state was saved every 20 epochs, for a total of 20 checkpoints per model. At each of these checkpoints, the complete set of images was generated. Inference time at this stage
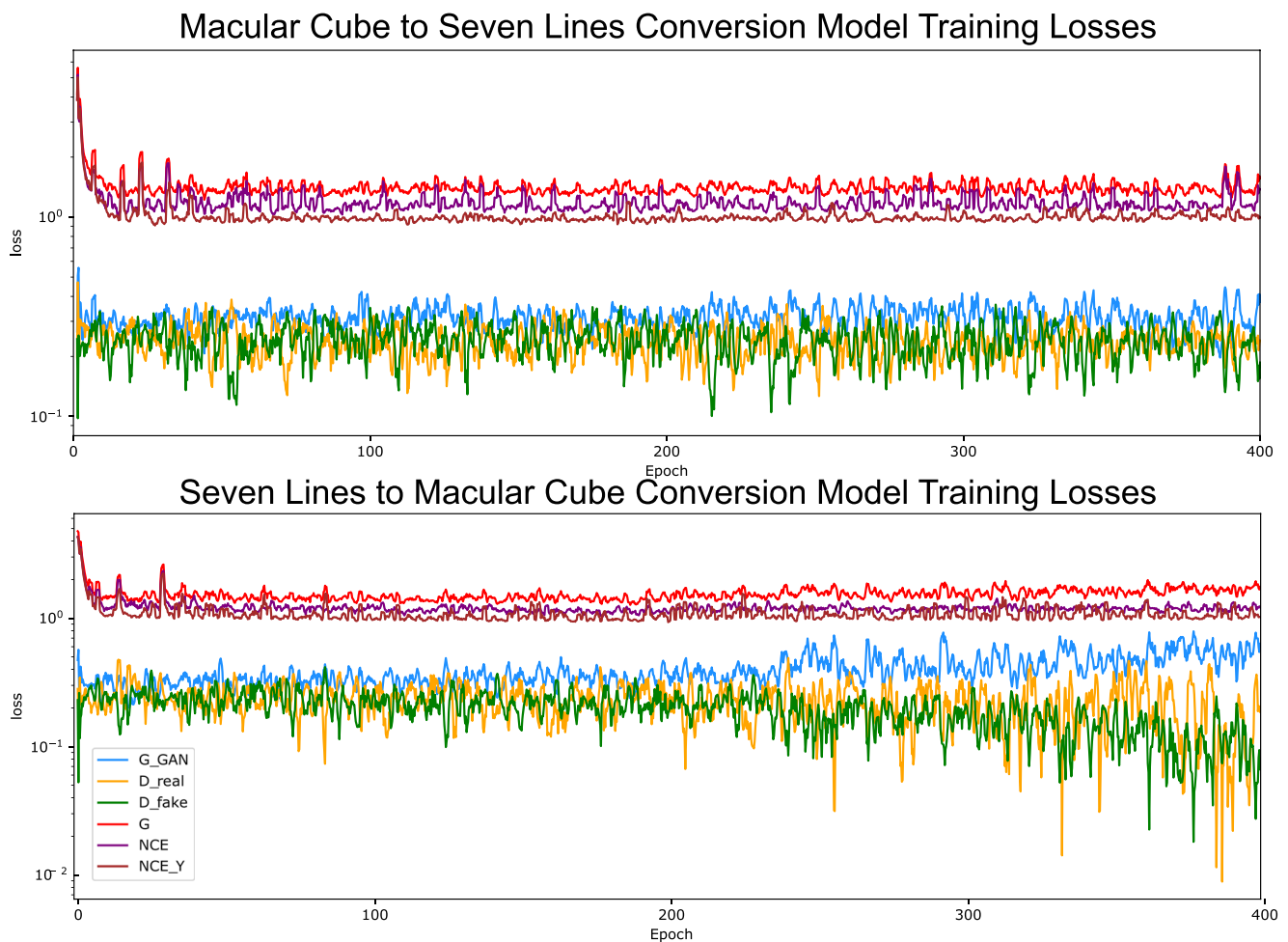

Fig. 7 Training losses for both approaches. G_GAN, generator GAN loss; D_real, discriminator loss for real images; D_fake, discriminator loss for synthetic images; G, generator loss; NCE, patchwise loss for images of the original class; NCE_Y, patchwise loss for images of the target class

was measured at 250 milliseconds per generated image. Then, the images that were generated by each checkpoint were evaluated using BRISQUE to determine the training checkpoint which produced the best results for each model This process is explained in the next subsection.

3.2 Image quality assessment

A test was conducted using the BRISQUE score to validate the perceptual quality distribution of the synthetic images. The aim of this experiment was to determine whether the quality distributions of the synthetic generated images are similar to those of the original images. In this experiment, every OCT image for each of the classes in the original dataset was evaluated using BRISQUE. Then, the BRISQUE score for each image generated by every checkpoint of the generative models was calculated and compared to the original images. Figure 8 shows the evolution of the BRISQUE score for both computational approaches for different epochs.

These graphs are indicative of the ability of the generative models to approach the image quality of the target classes. The highest BRISQUE score (corresponding to the lowest perceptual quality) is achieved by the original *Macular Cube* images, represented in red. These images also show the highest variability, with their standard deviation being indicated in red dashed lines. Original *Seven Lines* images, represented in green, show an overall lower BRISQUE score, also having a lower variability than their *Macular Cube* counterparts. Overall, all the synthetic images show a reduced variability in image quality, with synthetic *Seven Lines* images displaying a considerably higher quality than their original *Macular Cube* counterparts and stabilising around the lower fringes of the original *Seven Lines* distribution. In accordance with these results, the lowest scoring of the *Seven Lines*-generating model checkpoints and the highest of the *Macular Cube*-generating model checkpoints were selected, and the BRISQUE score distributions of their generated images were studied. The histograms of the indicated distributions of the original and the synthetic images can be found in Fig. 9.

Complementarily, a representative subset of the original *Macular Cube* images achieved an ENL and a CNR of 94.34 ± 44.41 and 143.35 ± 15.38 respectively, while the synthetic generated macular cube images were measured at 64.81 ± 27.73 and 144.26 ± 10.88. On the other hand, the higher-quality original *Seven Lines* images achieved an ENL and CNR of 152.56 ± 81.08 and 146.88 ± 25.59, while their synthetic counterparts were rated at 188.28 ± 111.08 and 147.68 ± 23.32. In terms of automated image quality scores BIQI and NIQE, the original *Macular Cube* images achieved scores of 25.79 ± 4.20 and 6.42 ± 2.53 respectively,

while synthetic generated *Macular Cube* images were rated at 35.62 ± 4.52 and 8.72 ± 1.11. Regarding the *Seven Lines* images, the originals were rated at scores of 24.76 ± 1.84 and 6.05 ± 0.80, while their synthetic counterparts achieved a very similar 24.71 ± 1.84 and 6.12 ± 1.90.

3.3 Separability test

As previously mentioned, a test was performed to validate the separability of these synthetic images by training a densely connected convolutional network to classify original images between those obtained with the *Macular Cube* preset and those acquired with *Seven Lines*. The original dataset was randomly partitioned in balanced sets, with 60% (622 images) forming a training set, 20% (206 images) making up a validation set to prevent overfitting and the remaining 206 images being used to test the network. Training and validation losses and accuracy values for this model are presented in Fig. 10.

The synthetic images converted by both of the generative models at the checkpoints that were selected in the previous subsection based on their BRISQUE score were then tested with this network. The accuracy obtained for the original and synthetic dataset is represented in Table 1. It should be clarified that, for the *Macular Cube* to *Seven Lines* model, the positives are the images of the *Seven Lines* class, while the opposite is true for the inverse model.

3.4 Validation by clinical specialists

Complementarily, a final test was conducted in order to assess whether medical specialists are able to detect the synthetic images. A random subset of 200 images which are representative of the four classes was created, with 72 of them being original *Macular Cube*, 31 synthetic *Macular Cube*, 64 original *Seven Lines* and 33 synthetic *Seven Lines*. The synthetic images were generated by the models that were selected as described in Section 3.2. Two ophthalmologists from the Hospital Clínico San Carlos in Madrid were asked to determine whether each of the OCT images was of the *Macular Cube* or *Seven Lines* type, and whether they were original or synthetic. One of the clinicians is a medical resident, while the other is an expert specialist with extensive medical experience. The two confusion matrices representing the final results of the test are displayed in Fig. 11.

4 Discussion

In this section, the results obtained from the generative methodology as well as those of the tests that were performed to evaluate the synthetic generated images are discussed.
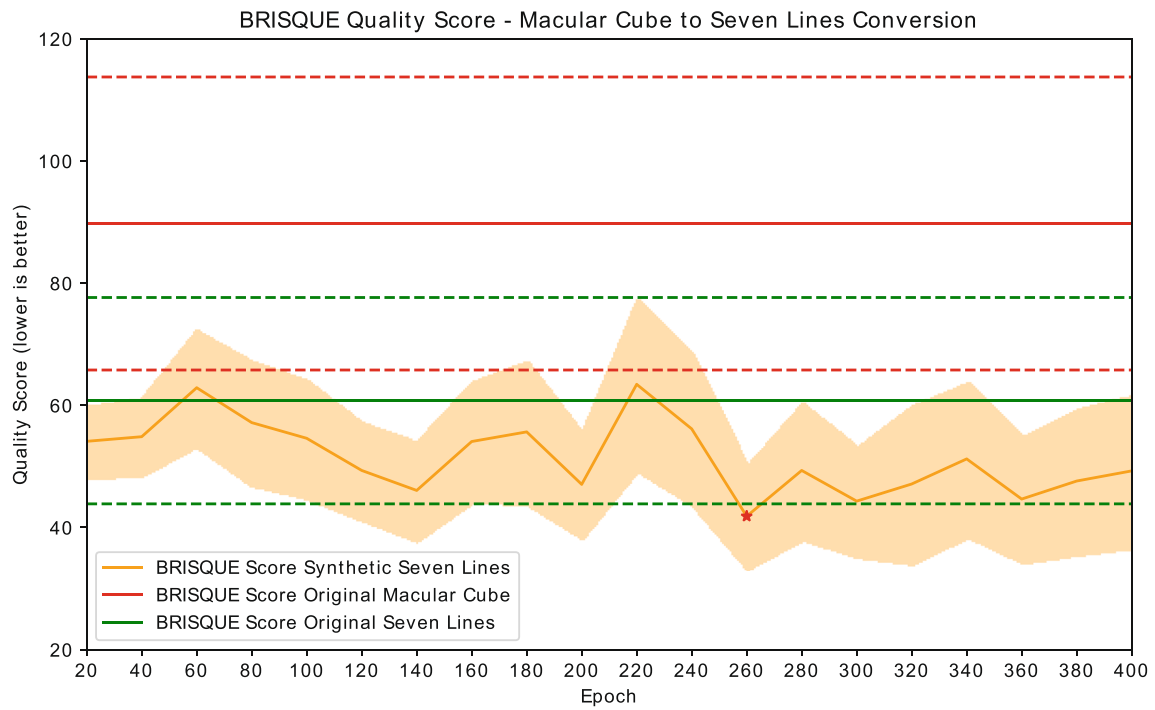
(a)
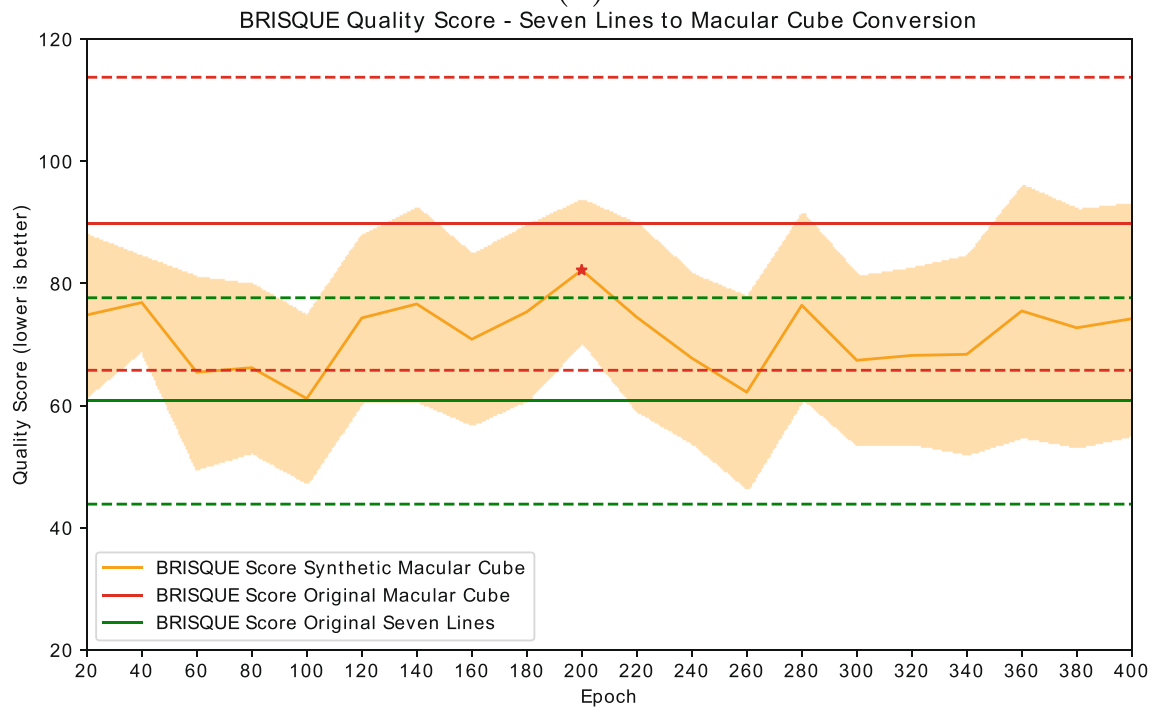
(b)

Fig. 8 BRISQUE score evolution for each computational approach. **a** 1st approach: Macular Cube to Seven Lines conversion model. **b** 2nd approach: Seven Lines to Macular Cube Conversion model. Red and green solid lines indicate, respectively, the average BRISQUE score for original Macular Cube and Seven Lines images. Dashed lines indicate standard deviation of these sets. A red star indicates the chosen checkpoint for each model
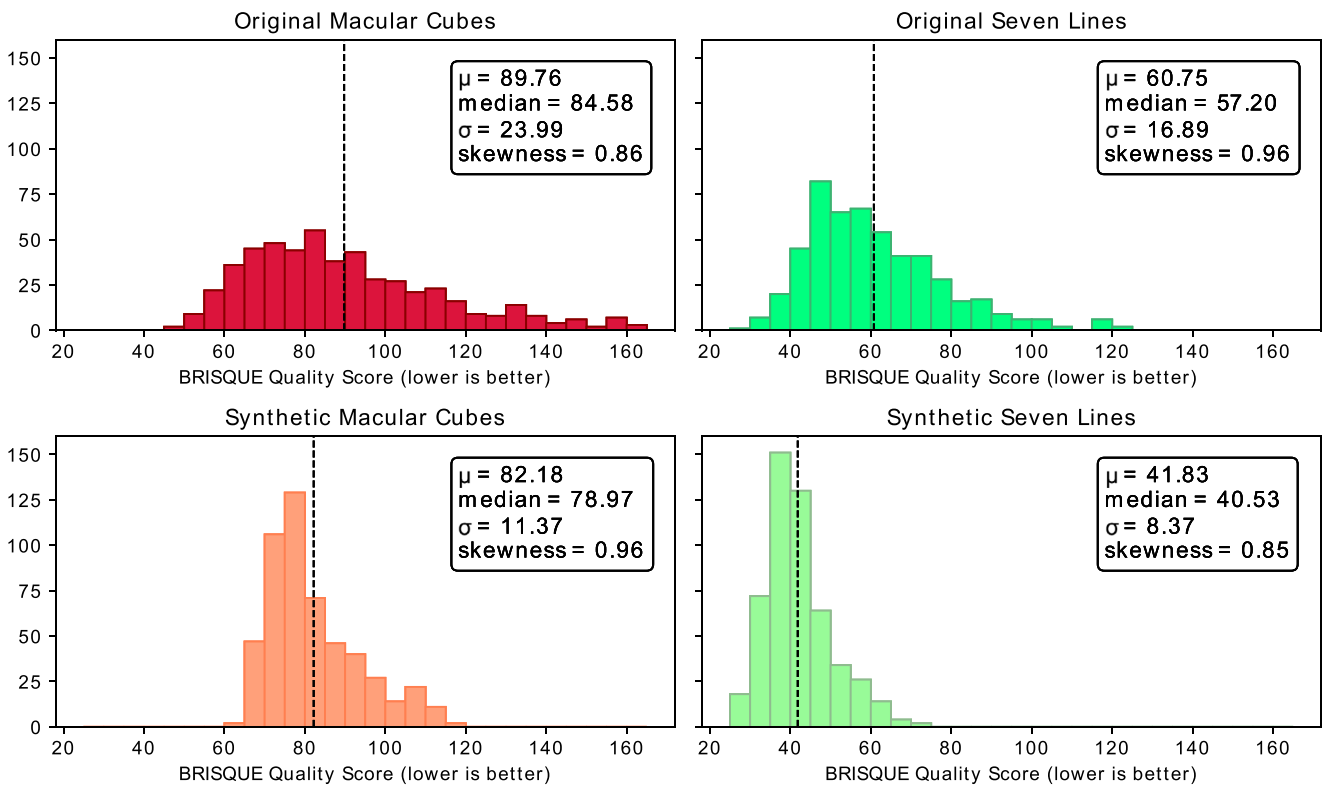
Fig. 9 BRISQUE score distribution for each of the data sets, including those generated by the chosen checkpoints of the generative models. *Top left*: Original *Macular Cube* images. *Top right*: Original *Seven Lines* images. *Bottom left*: *Seven Lines* images converted to synthetic *Macular Cube*. *Bottom right*: *Macular Cube* images converted to synthetic *Seven Lines*. Dashed lines indicate the mean value of the distribution

4.1 Image quality assessment

The image quality assessment results obtained from evaluating the synthetic generated images using BRISQUE (Fig. 9) show that all the sets are positively skewed, with most images having a lower BRISQUE score than the mean and a long tail of samples with increasingly higher scores, formed by images with progressively lower quality. This is especially apparent for the original *Macular Cube* samples, which is to be expected of the set with the highest variability in noise and tissue visibility, being the fastest scanning preset considered. Conversely, while the synthetic
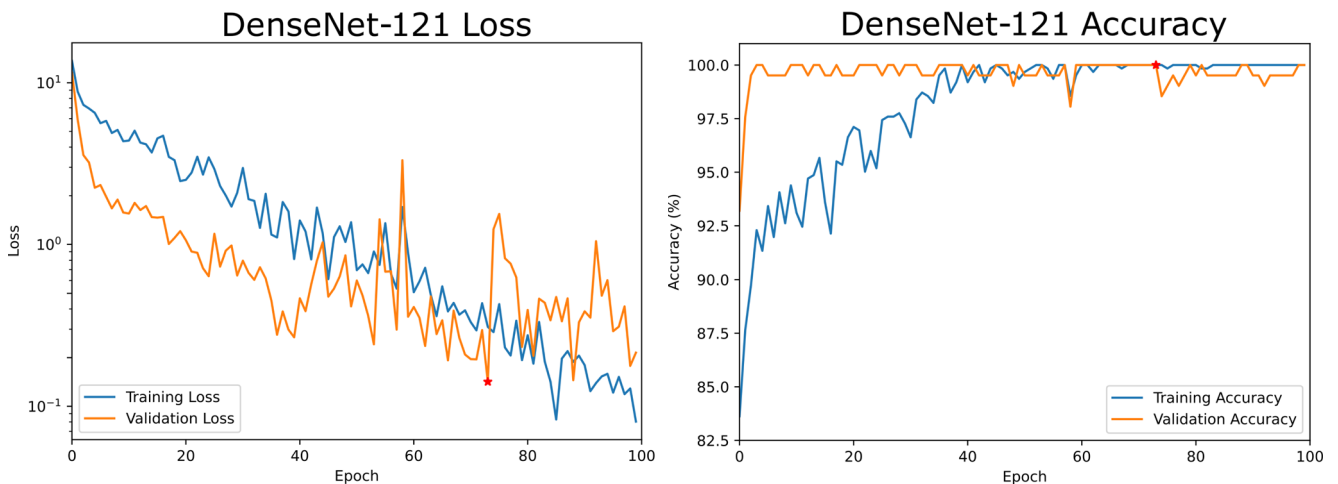

Fig. 10 Training and validation losses and accuracies for the DenseNet-121 classifier used to test synthetic image separability. A red star indicates the epoch at which the model produced the lowest validation loss and that was chosen for testing

Table 1 Accuracy obtained with the DenseNet-121 model for each image class, both for original and synthetic images

		N	Accuracy	
			Macular Cube	Seven Lines
Original images	Training	622	100.00%	100.00%
	Validation	206	100.00%	100.00%
	Test	206	100.00%	100.00%
Synthetic images	Test	1.034	99.42%	100.00%

images also present a similar positive skewness, there is a significant reduction in the amount of unusually highly scored images. These histograms also show the previously mentioned decrease in variability for the synthetic generated images. Overall, the generative models show a great consistency and stability at generating the synthetic images, producing images that show score distributions coherently formed around their respective target quality.

The images were also visually inspected to ensure that the measured changes in BRISQUE score correspond to changes in actual perceptual quality. The synthetic *Seven Lines* images show a significant reduction in speckle noise and tissue visibility in the retina and choroid (Fig. 12). Conversely, regarding synthetic *Macular Cube*, it is apparent that generated images representing the original tissue with an addition of speckle noise and visual features bear resemblance to the original *Macular Cube* samples (see Fig. 13). In both figures, the values of the BRISQUE

score correlate with perceptual changes in image visibility, with noisier images being rated higher scores, and images showing greater retinal and choroid visibility achieving lower scores.

Aside from this, the results for the GAN evolution (Fig. 8) show that, for some epochs, synthetic *Macular Cube* images seem to be rated at a similar or higher quality than their original *Seven Lines* counterparts. While this behaviour is not necessarily unusual when training a GAN due to the oscillatory nature of both the discriminator and generative networks, these images were also inspected. A sample of images from training epoch 260 of the *Seven Lines* to *Macular Cube* translating model, in which the network perceptually increased the quality of the *Seven Lines* images, can be found in Fig. 14. This behaviour, combined with the absence of an absolute indicator of GAN training progression, is what motivates the use of an external quality evaluator such as BRISQUE to assess image quality and determine a satisfactory epoch to stop the training process at a point where the generated images present the desired visual features.

4.2 Separability test

The results obtained from the automatic separability test (Table 1) show that the synthetic images are able to mimic the visual features of their target class, with every original *Macular Cube* converted to *Seven Lines* being correctly recognised as a *Seven Lines* image and only 6 synthetic *Macular Cube* images being confused as *Seven Lines*. The
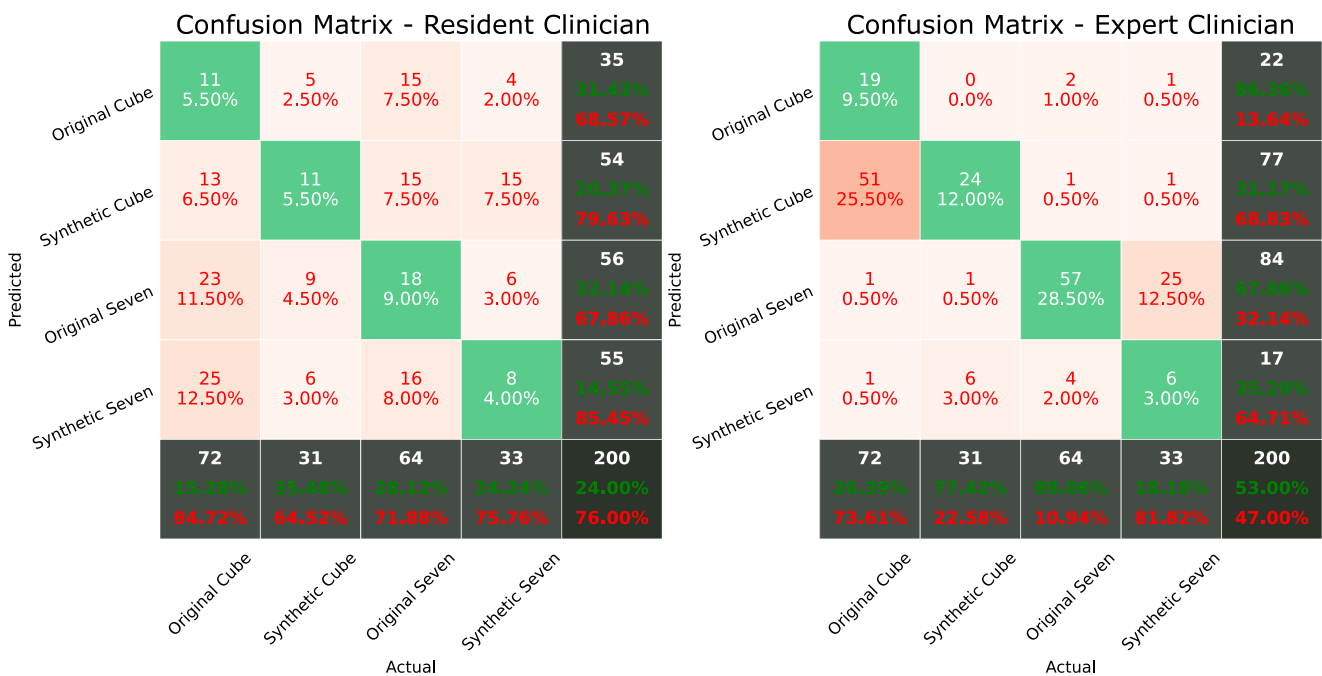

Fig. 11 Confusion matrices for the classification results of both specialists
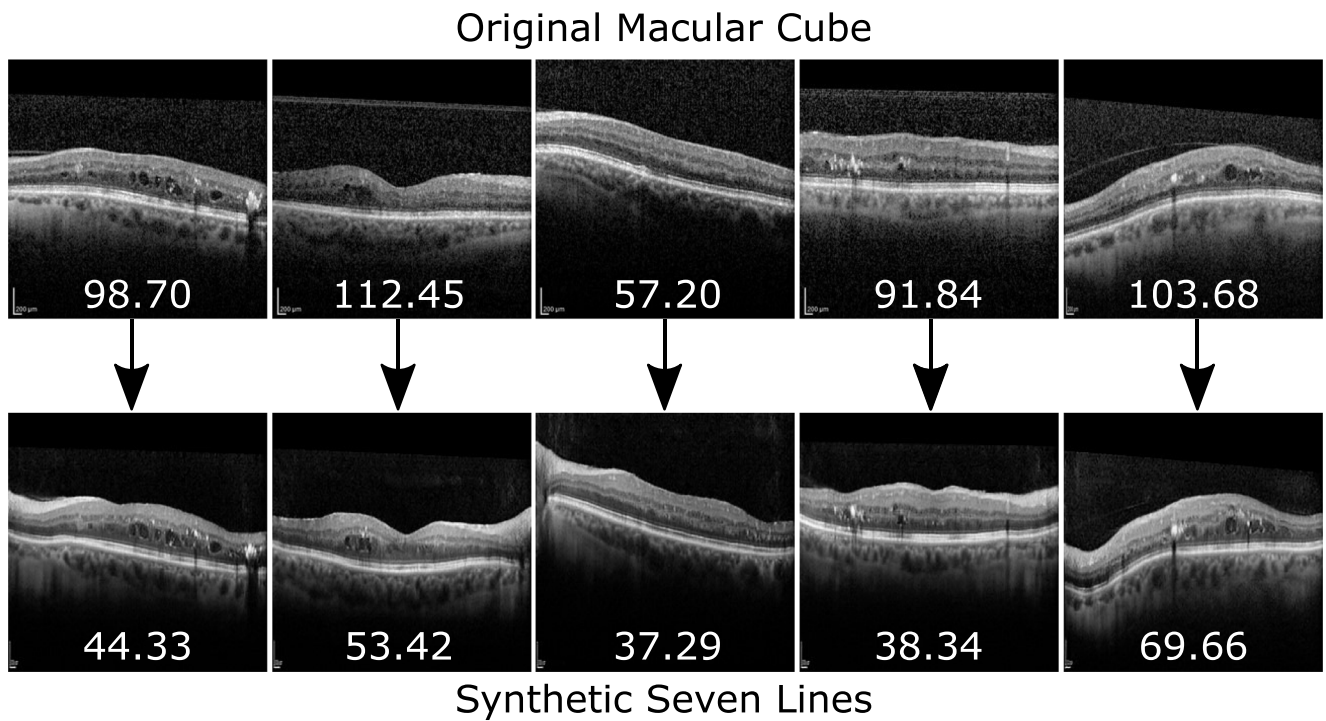
Original Macular Cube

Fig. 12 Sample of original *Macular Cube* images and their converted *Seven Lines* synthetic counterparts. Values represent BRISQUE score for each image

synthetic test set contains the images which the network recognises as their original class, but converted to the opposite class. These synthetic generated images were classified by the model as their target class instead of the original one, indicating that they display the intended visual features.
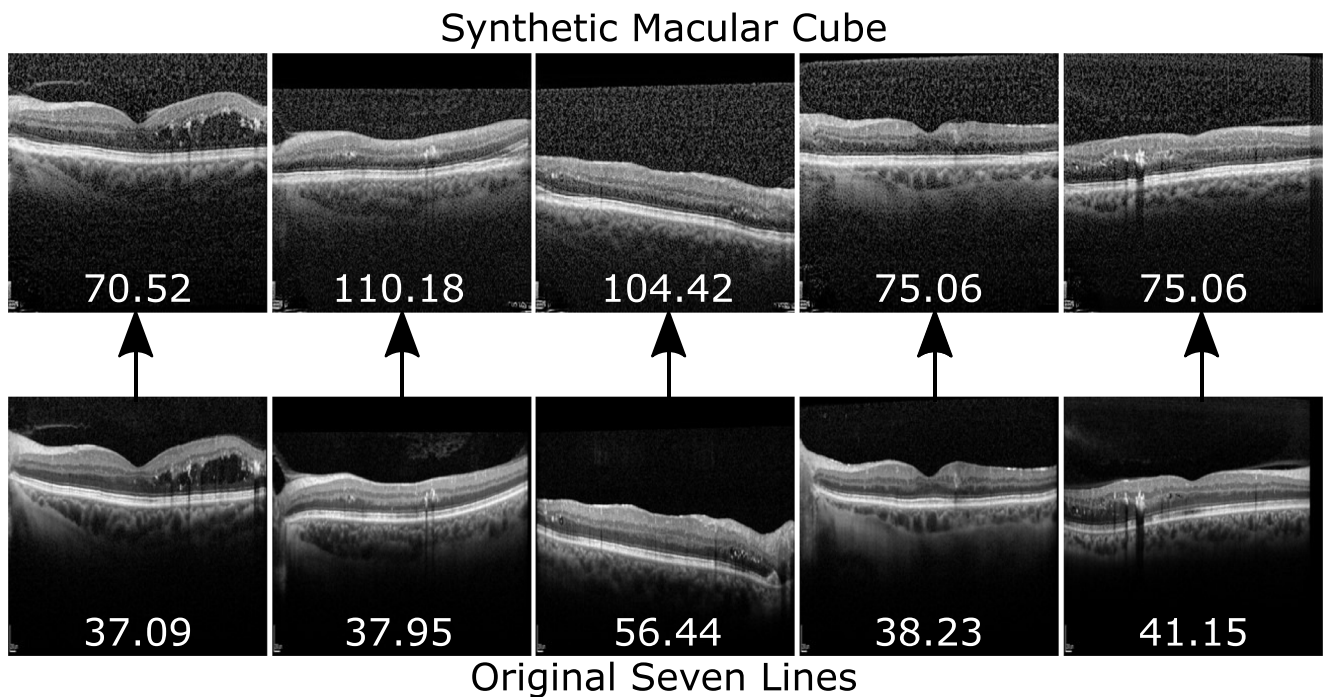
Synthetic Macular Cube

Fig. 13 Sample of original *Seven Lines* images and their converted *Macular Cube* synthetic counterparts. Values represent the BRISQUE score for each image
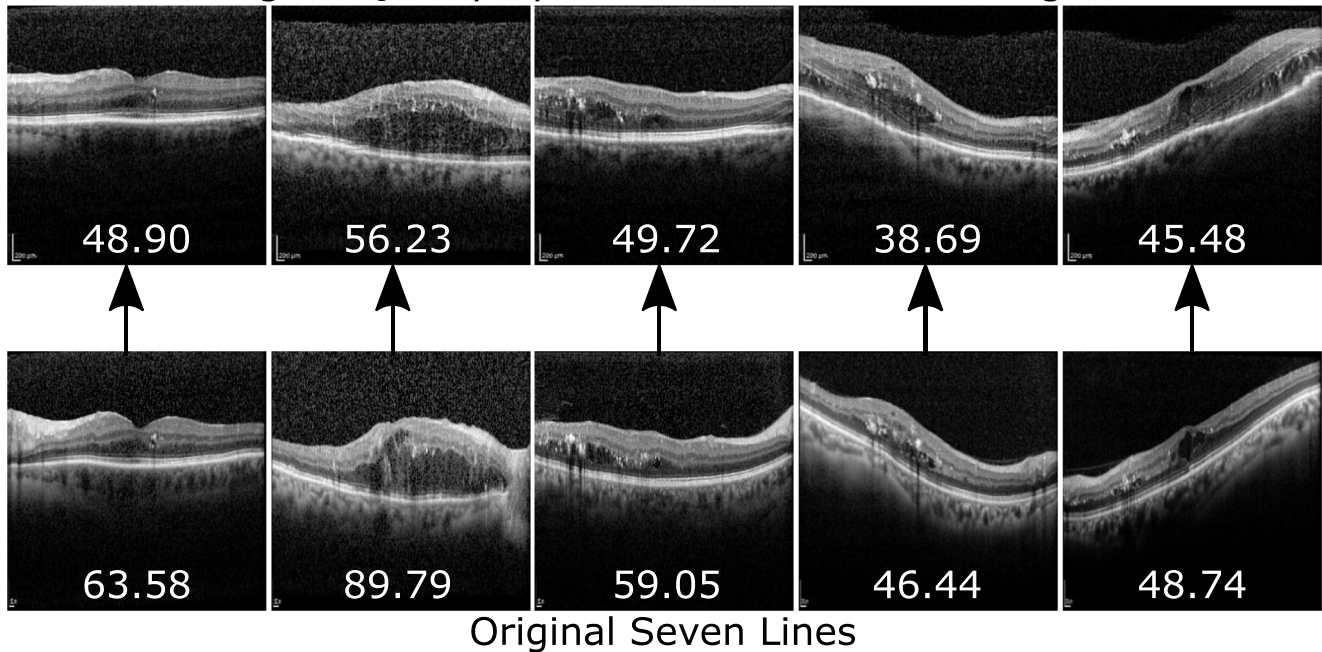
Fig. 14 Sample of images generated by a *Seven Lines* to *Macular Cube* model at a point in training in which it actually increases perceptual quality. Values represent the BRISQUE score for each image

4.3 Validation by clinical specialists

Regarding the ability of the ophthalmology specialists to discern the original and synthetic images of the two presets, the test results (Fig. 11) highlight the difference in experience level between both specialists and its relevance for a complex problem such as this, with the resident achieving an overall accuracy of 24% and the expert correctly guessing more than half of the samples. This complexity is even more evident when evaluating *Macular Cube* and *Seven Lines* separability, with the resident attaining an accuracy of 44% while the expert reached 93%. This indicates that while the OCT images are clearly visually separable according to their acquisition preset, this is by no means a trivial problem, and the ability to do so is acquired with experience. Absence or presence of speckle noise is not enough to differentiate between the images, with other visual features being necessary to distinguish them. When taking into account original and synthetic separability, the results show that the synthetic images are able to deceive even the expert specialist. The expert was not able to determine whether images were originally acquired with a scanner or converted by the generative networks while correctly identifying the visual features that characterise both *Macular Cube* and *Seven Lines* images. Most of the synthetic *Seven Lines* samples, presenting a clearer visibility, were incorrectly classified

as original images while, conversely, most of the original *Macular Cube* images, which display more noise and a perceptually worse appearance, were mistakenly identified as synthetic. These results show that the synthetic generated images are effectively indiscernible from the original ones while at the same time preserving the distinctive visual features of their target classes. The obtained results are also indicative of the absence of visual artefacts that could be introduced in the synthetic generated images. Therefore, the tests that were conducted with the specialists demonstrate the substantial performance of the generative models, showing that they are able to generate images that are interchangeable with the original ones even to the expert eye. All of the results obtained indicate that these models are suitable for the purpose of supplying datasets with images converted to the style of different configurations that can be used just as if they were acquired with their target presets.

It should also be highlighted that while other approaches exist in the literature for the denoising or resolution enhancement of low quality images, this proposal is the first to address the problem of data scarcity in OCT through the mutual conversion of images between scanning presets. Due to this focus shift from the enhancement of low quality images to the mutual conversion between visual features within a domain, no comparison with the currently existing methods can be drawn.

5 Conclusions

OCT is a relevant medical imaging technique that can be used in conjunction with CAD systems to diagnose relevant ocular pathologies and to study the eye tissue. The overall quality and visibility of the OCT images is considerably affected by light scattering. To overcome this, the OCT scanning platforms typically sample each point multiple times and average the signals to obtain a clearer image. Due to involuntary eye movements, there is a limitation to the amount of samples that can be taken in a scanning session. OCT platforms provide a number of scanning presets that determine the number of scans that are averaged per OCT image, balancing the amount of tissue that is scanned and the quality of the produced images. This compromise between sampled area and image quality leads to a scarcity of high quality data. Moreover, the visual differences that exist between images obtained with different presets limits the potential of datasets based on the scanning preset that was used to acquire their images.

In this work, a complete methodology for the automatic mutual conversion of OCT images has been presented. These OCT images were acquired with the two scanning presets most representative of those used by clinical specialists in medical services: *Macular Cube*, a fast scanning preset which produces 25 eye slices over a square patch of the retina with considerable speckle noise, and *Seven Lines*, an intensive preset that can create cleaner images at the cost of only producing 7 B-scans over a narrow band per session; representing a context of image quality versus quantity compromise that is so widespread in several medical imaging areas. This mutual conversion is achieved by training a contrastive unpaired translation GAN model to translate the more numerous *Macular Cube* images into the higher-visibility style of the intensive *Seven Lines* preset and a second model to perform the complementary conversion. The quality of the synthetic images generated by these models is assessed and compared to the originals in order to determine the optimal training model checkpoint, with an aim to validate the quality of these images and to solve the problem of when to stop the GAN training process.

The experiments that were conducted to validate the synthetic generated images show that these are able to display the visual features of those acquired with their target preset. Qualitatively, the BRISQUE score of original and synthetic images of each preset are very similar, with synthetic images presenting a consistent stability around their target quality distributions. In a validation experiment using a dense convolutional network trained to classify the original images based on their acquisition preset, the synthetic generated images demonstrated a clear separability, being classified as if they were originals of their target preset.

Complementarily, as a way to assess the visual and perceptual qualities of the synthetic images, two ophthalmology specialists with different levels of experience were tasked with classifying images according to whether they are original or synthetic, and according to their acquisition presets. In this experiment, the clinicians were unable to discern between the original and synthetic images, while the expert was clearly able to correctly identify the presets of the originals and the intended target ones of the synthetic generated images. Overall, the generative models demonstrated their ability to provide synthetic generated images that are exceptionally similar to the original ones of their target classes, even to the expert eye.

From the obtained experimental results, it can be concluded that this methodology is able to replicate the visual features of each of the presets in images acquired with another. The synthetic images were validated in terms of perceptual quality, automatic separability and expert separability, with results showing that they resemble their target presets in each of these terms. These generative models can be used to supply OCT datasets limited by their acquisition presets with quality synthetic generated images that display the visual features of any other preset.

Plans for future work include assessing the possible benefits that may be obtained from the paired translation of images in terms of tissue preservation, as well as the possible application of this methodology to produce multi-preset datasets that can be used to train CAD systems in a more robust manner, allowing them to train with all the possible presets it may encounter in a real setting without the need to procure these images. Furthermore, a more elaborate analysis and evaluation of these models and how they perform when trained with images belonging to patients of different ages, sexes and affected by different pathologies is considered for future work. Lastly, this methodology should be considered for the exploration of this context of data scarcity related to image quality and acquisition conditions in other fields of medical imaging where it is so widespread, constituting a paradigm in itself.

Xeral de Universidades (20%). Funding for open access charge: Universidade da Coruña/CISUG.

Declarations

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Universidade da Coruña/Ferrol (2014/437) the 24th of November, 2014.

Conflict of interest The authors declare no competing interests.

Disclaimer The funding sources had no role in the development of this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Huang D, Swanson E, Lin C, Schuman J, Stinson W, Chang W, Hee M, Flotte T, Gregory K, Puliafito C, Fujimoto J. (1991) Optical coherence tomography. Science 254(5035):1178–1181. https://doi.org/10.1126/science.1957169, https://science.sciencemag.org/content/254/5035/1178, https://science.sciencemag.org/content/254/5035/1178.full.pdf
2. Drexler W, Fujimoto JG (2008) State-of-the-art retinal optical coherence tomography. Prog Retin Eye Res 27(1):45–88. https://doi.org/10.1016/j.preteyeres.2007.07.005, https://www.sciencedirect.com/science/article/pii/S1350946207000444
3. Swanson EA, Fujimoto JG (2017) The ecosystem that powered the translation of OCT from fundamental research to clinical and commercial impact Invited. Biomed Opt Express 8(3):1638–1664. https://doi.org/10.1364/BOE.8.001638, http://www.osapublishing.org/boe/abstract.cfm?URI=boe-8-3-1638
4. Hee M (1995) Quantitative assessment of macular edema with optical coherence tomography. Arch Ophthalmol 113(8):1019. https://doi.org/10.1001/archopht.1995.01100080071031
5. de Moura J, Samagaio G, Novo J, Almuina P, Fernández MI, Ortega M (2020) Joint diabetic macular edema segmentation and characterization in OCT images. J Digit Imaging 33(5):1335–1351. https://doi.org/10.1007/s10278-020-00360-y
6. Mookiah MRK, Acharya UR, Chua CK, Lim CM, Ng E, Laude A (2013) Computer-aided diagnosis of diabetic retinopathy: a review. Comput Biol Med 43(12):2136–2155. https://doi.org/10.1016/j.compbiomed.2013.10.007
7. Jaffe GJ, Caprioli J (2004) Optical coherence tomography to detect and manage retinal disease and glaucoma. Am J Ophthalmol 137(1):156–169. https://doi.org/10.1016/s0002-9394(03)00792-x
8. Tan O, Chopra V, Lu ATH, Schuman J, Ishikawa H, Wollstein G, Varma R, Huang D (2009) Detection of macular ganglion cell loss in glaucoma by fourier-domain optical coherence tomography. Ophthalmology 116(12):2305–2314.e2. https://doi.org/10.1016/j.ophtha.2009.05.025
9. Hood DC (2017) Improving our understanding and detection, of glaucomatous damage: an approach based upon optical coherence tomography (OCT). Prog Retin Eye Res 57:46–75. https://doi.org/10.1016/j.preteyeres.2016.12.002
10. Mitchell P, Liew G, Gopinath B, Wong TY (2018) Age-related macular degeneration. The Lancet 392(10153):1147–1159. https://doi.org/10.1016/s0140-6736(18)31550-2
11. Vidal PL, de Moura J, Novo J, Penedo MG, Ortega M (2018) Intraretinal fluid identification via enhanced maps using optical coherence tomography images. Biomed Opt Express 9(10):4730. https://doi.org/10.1364/boe.9.004730
12. Borrelli E, Sarraf D, Freund KB, Sadda SR (2018) OCT angiography and evaluation of the choroid and choroidal vascular disorders. Prog Retin Eye Res 67:30–55. https://doi.org/10.1016/j.preteyeres.2018.07.002
13. Spaide RF, Fujimoto JG, Waheed NK, Sadda SR, Staurenghi G (2018) Optical coherence tomography angiography. Prog Retin Eye Res 64:1–55. https://doi.org/10.1016/j.preteyeres.2017.11.003
14. de Moura J, Novo J, Rouco J, Penedo MG, Ortega M (2017) Automatic detection of blood vessels in retinal OCT images. In: International work-conference on the interplay between natural and artificial computation. Springer, pp 3–10
15. Kashani AH, Chen CL, Gahm JK, Zheng F, Richter GM, Rosenfeld PJ, Shi Y, Wang RK (2017) Optical coherence tomography angiography: a comprehensive review of current methods and clinical applications. Prog Retin Eye Res 60:66–100. https://doi.org/10.1016/j.preteyeres.2017.07.002
16. de Moura J, Novo J, Charlón P, Barreira N, Ortega M (2017) Enhanced visualization of the retinal vasculature using depth information in OCT. Med Biol Eng Comput 55(12):2209–2225. https://doi.org/10.1007/s11517-017-1660-8
17. Schmitt JM, Xiang SH, Yung KM (1999) Speckle in optical coherence tomography. J Biomed Opt 4(1):95. https://doi.org/10.1117/1.429925
18. Chan HP, Hadjiiski LM, Samala RK (2020) Computer-aided diagnosis in the era of deep learning. Med Phys 47(5):e218–e227. https://doi.org/10.1002/mp.13764
19. Singh LK, Pooja, Garg M, Khanna M, Bhadoria RS (2021) An enhanced deep image model for glaucoma diagnosis using feature-based detection in retinal fundus. Med Biol Eng Comput 59(2):333–353. https://doi.org/10.1007/s11517-020-02307-5
20. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JA, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88. https://doi.org/10.1016/j.media.2017.07.005
21. Lee JH, Kim YT, Lee JB, Jeong SN (2020) A performance comparison between automated deep learning and dental professionals in classification of dental implant systems from dental imaging: a multi-center study. Diagnostics 10(11):910. https://doi.org/10.3390/diagnostics10110910
22. Ting DSW, Cheung CYL, Lim G, Tan GSW, Quang ND, Gan A, Hamzah H, Garcia-Franco R, Yeo IYS, Lee SY, Wong EYM, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Sivaprasad S, Varma R, Jonas JB, He MG, Cheng CY, Cheung GCM, Aung T, Hsu W, Lee ML, Wong TY (2017) Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA 318(22):2211. https://doi.org/10.1001/jama.2017.18152

23. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316(22):2402. https://doi.org/10.1001/jama.2016.17216

24. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1). https://doi.org/10.1186/s40537-019-0197-0

25. Adler DC, Ko TH, Fujimoto JG (2004) Speckle reduction in optical coherence tomography images by use of a spatially adaptive wavelet filter. Opt Lett 29(24):2878. https://doi.org/10.1364/ol.29.002878

26. Wong A, Mishra A, Bizheva K, Clausi DA (2010) General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery. Opt. Express 18(8):8338–8352. https://doi.org/10.1364/OE.18.008338, http://www.opticsexpress.org/abstract.cfm?URI=oe-18-8-8338

27. Cameron A, Lui D, Boroomand A, Glaister J, Wong A, Bizheva K (2013) Stochastic speckle noise compensation in optical coherence tomography using non-stationary spline-based speckle noise modelling. Biomed Opt Express 4(9):1769–1785. https://doi.org/10.1364/BOE.4.001769, http://www.osapublishing.org/boe/abstract.cfm?URI=boe-4-9-1769

28. Aum J, hyun Kim J, Jeong J (2015) Effective speckle noise suppression in optical coherence tomography images using nonlocal means denoising filter with double gaussian anisotropic kernels. Appl Opt 54(13):D43. https://doi.org/10.1364/ao.54.000d43

29. Li M, Idoughi R, Choudhury B, Heidrich W (2017) Statistical model for OCT image denoising. Biomed Opt Express 8(9):3903–3917. https://doi.org/10.1364/BOE.8.003903, http://www.osapublishing.org/boe/abstract.cfm?URI=boe-8-9-3903

30. Chong B, Zhu YK (2013) Speckle reduction in optical coherence tomography images of human finger skin by wavelet modified BM3D filter. Opt Commun 291:461–469. https://doi.org/10.1016/j.optcom.2012.10.053, https://www.sciencedirect.com/science/article/pii/S0030401812012199

31. Kafieh R, Rabbani H, Selesnick I (2015) Three dimensional data-driven multi scale atomic representation of optical coherence tomography. IEEE Trans Med Imaging 34(5):1042–1062. https://doi.org/10.1109/TMI.2014.2374354

32. Apostolopoulos S, Salas J, Ordóñez JLP, Tan SS, Ciller C, Ebneter A, Zinkernagel M, Sznitman R, Wolf S, Zanet SD, Munk MR (2020) Automatically enhanced OCT scans of the retina: a proof of concept study. Sci Rep 10(1). https://doi.org/10.1038/s41598-020-64724-8

33. Xu M, Tang C, Hao F, Chen M, Lei Z (2020) Texture preservation and speckle reduction in poor optical coherence tomography using the convolutional neural network. Med Image Anal 64:101,727. https://doi.org/10.1016/j.media.2020.101727, https://www.sciencedirect.com/science/article/pii/S1361841520300918

34. Seebök P, Romo-Bucheli D, Waldstein S, Bogunovic H, Orlando JI, Gerendas BS, Langs G, Schmidt-Erfurth U (2019) Using Cyclegans for effectively reducing image variability across OCT devices and improving retinal fluid segmentation. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), pp 605–609, https://doi.org/10.1109/ISBI.2019.8759158

35. Huang Y, Lu Z, Shao Z, Ran M, Zhou J, Fang L, Zhang Y (2019) Simultaneous denoising and super-resolution of optical coherence tomography images based on generative adversarial network. Opt Express 27(9):12,289–12,307. https://doi.org/10.1364/OE.27.012289, http://www.opticsexpress.org/abstract.cfm?URI=oe-27-9-12289

36. Gende M, de Moura J, Novo J, Ortega M (2022) High/low quality style transfer for mutual conversion of OCT images using contrastive unpaired translation generative. In: Image analysis and processing – ICIAP 2022, Lecture notes in computer science. Springer International Publishing, Cham, pp 210–220

37. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S (2019) PyTorch: an imperative style, high-performance deep learning library. In: Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox E, Garnett R (eds) Advances in neural information processing systems, vol 32. Curran Associates, Inc, pp 8026–8037. https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf

38. Bradski G (2000) The opencv library. Dr Dobb's. J Softw Tools 25:120–125. https://ci.nii.ac.jp/naid/10028167478/en/

39. Park T, Efros AA, Zhang R, Zhu JY (2020) Contrastive learning for unpaired image-to-image translation. In: Computer Vision – ECCV 2020. Springer International Publishing, Cham, pp 319–345, https://doi.org/10.1007/978-3-030-58545-7_19

40. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. Commun ACM 63(11):139–144. https://doi.org/10.1145/3422622

41. Chen T, Kornblith S, Norouzi M, Hinton G (2020) A simple framework for contrastive learning of visual representations. In: III HD, Singh A (eds) Proceedings of the 37th international conference on machine learning, Proceedings of Machine Learning Research, vol 119. PMLR, pp 1597–1607. https://proceedings.mlr.press/v119/chen20j.html

42. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778. https://doi.org/10.1109/CVPR.2016.90

43. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: Bengio Y, LeCun Y (eds) 3rd international conference on learning representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015 Conference Track Proceedings. arXiv:1412.6980

44. Mittal A, Moorthy AK, Bovik AC (2012) No-reference image quality assessment in the spatial domain. IEEE Trans Image Process 21(12):4695–4708. https://doi.org/10.1109/tip.2012.2214050

45. Horé A, Ziou D (2010) Image quality metrics: PSNR vs. SSIM. In: 2010 20th international conference on pattern recognition, pp 2366–2369. https://doi.org/10.1109/ICPR.2010.579

46. Wang Z, Bovik A, Sheikh H, Simoncelli E (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612. https://doi.org/10.1109/TIP.2003.819861

47. Zhang Z, Dai G, Liang X, Yu S, Li L, Xie Y (2018) Can signal-to-noise ratio perform as a baseline indicator for medical image quality assessment. IEEE Access 6:11,534–11,543. https://doi.org/10.1109/access.2018.2796632

48. Yu S, Dai G, Wang Z, Li L, Wei X, Xie Y (2018) A consistency evaluation of signal-to-noise ratio in the quality assessment of human brain magnetic resonance images. BMC Med Imaging 18(1). https://doi.org/10.1186/s12880-018-0256-6

49. Chaabouni A, Gaudeau Y, Lambert J, Moureaux JM, Gallet P (2014) Subjective and objective quality assessment for H264 compressed medical video sequences. In: 2014 4th international conference on image processing theory, tools and applications (IPTA), pp 1–5. https://doi.org/10.1109/IPTA.2014.7001922

50. Moorthy AK, Bovik AC (2010) A two-step framework for constructing blind image quality indices. IEEE Sig Process Lett 17(5):513–516

51. Mittal A, Soundararajan R, Bovik AC (2013) Making a "completely blind" image quality analyzer. IEEE Sig Process Lett 20(3):209–212

52. Huang G, Liu Z, Maaten LVD, Weinberger KQ (2017) Densely connected convolutional networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 2261–2269. https://doi.org/10.1109/cvpr.2017.243

53. Nugroho KA (2018) A comparison of handcrafted and deep neural network feature extraction for classifying optical coherence tomography (OCT) images. In: 2018 2nd international conference on informatics and computational sciences (ICICoS), pp 1–6. https://doi.org/10.1109/ICICOS.2018.8621687

54. Al-Bander B, Williams BM, Al-Nuaimy W, Al-Taee MA, Pratt H, Zheng Y (2018) Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis. Symmetry 10(4). https://doi.org/10.3390/sym10040087, https://www.mdpi.com/2073-8994/10/4/87

55. Wang S, Tang C, Sun J, Zhang Y (2019) Cerebral micro-bleeding detection based on densely connected neural network. Front Neurosci 13. https://doi.org/10.3389/fnins.2019.00422

56. Yildirim O, Talo M, Ay B, Baloglu UB, Aydin G, Acharya UR (2019) Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals. Comput Biol Med 113:103,387. https://doi.org/10.1016/j.compbiomed.2019.103387, http://www.sciencedirect.com/science/article/pii/S0010482519302641

Mateo Gende received his degree in Computer Engineering in 2018 from the University of A Coruña (Spain), and his M.Sc. degree in Bioinformatics for Health Sciences in 2021. In 2020, he started pursuing his Ph.D. degree in Computer Science and Artificial Intelligence. He has also worked as a visiting researcher in the analysis of ophthalmic images at University College London. His main research interests lie in computer vision, medical image analysis, pattern recognition and machine learning.

Joaquim de Moura received his degree in Computer Engineering in 2014 from the University of A Coruña (Spain). In 2016, he received his M.Sc. degree in Computer Engineering from the same university and in 2019 his Ph.D. degree (cum Laude) in Computer Science and Artificial Intelligence. He has also worked as a visiting researcher at the INESC-TEC (Portugal), in the development of different methodologies for DME diagnosis with OCT images. He is currently a postdoctoral researcher in the Centre for Research in Information and Communication Technologies (CITIC) at the University of A Coruña. His main research interests include computer vision, biomedical image processing and video processing.


Jorge Novo received his M.Sc. and Ph.D. degrees (cum Laude) in Computer Science from the University of A Coruña in 2007 and 2012, respectively. He has also worked as a visiting researcher with CMR images in the detection of landmark points at Imperial College London, and as a postdoctoral research fellow at the INEB and INESC-TEC research institutes in the development of CAD systems for lung cancer diagnosis with chest CT images. He is currently an Associate Professor in the Department of Computer Science and Information Technologies at the University of A Coruña, and his main research interests lie in the fields of computer vision, pattern recognition and biomedical image processing.

Manuel G. Penedo is currently the Director of Research Centre for Information Technology and Communications (CITIC), University of A Coruña, Centre recognised by the Autonomous Government. He also is responsible for the area of medical image processing at the Institute of Biomedical Research of A Coruña (INIBIC) and he heads the VARPA research group (Computer Vision and Pattern Recognition) in the University of A Coruña. Belongs to various international scientific comitees (journals and conferences). He is a Professor of the Department of Computer Science in the University of A Coruña, Spain. He is author of over 200 papers in international books, journals and conference proceedings.
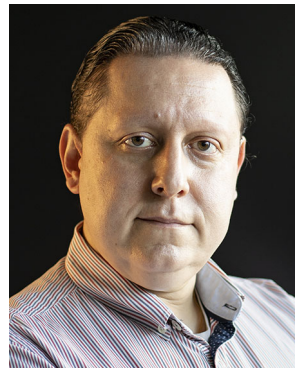
Marcos Ortega received his degree in Computer Science in 2004 and a doctor's degree in Computer Science in 2009. He currently serves as Associate Professor at the University of A Coruña, teaching mainly in the Faculty of Computer Science, also serving as his Secretary. He is a researcher in Research Centre in Information and Communication Technologies (CITIC), member of its scientific committee and its representative in the ECHAlliance. He is also a member of the Institute of Biomedical Research of A Coruña (INIBIC) with the rank of Principal Investigator. His research focuses on computer vision, medical image processing and medical informatics.

Affiliations

Mateo Gende[1,2] · **Joaquim de Moura**[1,2] · **Jorge Novo**[1,2] · **Manuel G. Penedo**[1,2] · **Marcos Ortega**[1,2]

Mateo Gende
m.gende@udc.es

Jorge Novo
jnovo@udc.es

Manuel G. Penedo
mgpenedo@udc.es

Marcos Ortega
mortega@udc.es

[1] Grupo, VARPA, Instituto de Investigación Biomédica de A Coruña (INIBIC), Universidade da Coruña, Xubias de Arriba, 84, A Coruña, 15006, A Coruña, Spain

[2] Centro de investigación, CITIC, Universidade da Coruña, Campus de Elviña s/n, A Coruña, 15071, A Coruña, Spain