Order-specific removal of nonlinearity from Optical Coherence Tomography signals

Krzysztof A. Maliszewski[a], Varvara Vetrova[a], and Sylwia M. Kolenderska[b,c]

[a]School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand
[b]Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University in Torun, ul. Grudziadzka 5, 87-100 Torun, Poland
[c]School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand

ABSTRACT

We present two neural networks: one capable of processing a raw spectrum into an A-scan with the second-order nonlinearity removed and another for processing a raw spectrum into an A-scan with the third-order nonlinearity removed. An algorithm is also proposed to enable to use these networks in a sequence for removal of both nonlinearities. The presented approaches allow for either independent switching off of each order or the simultaneous removal of all orders, offering a tool for analysing the effects of each nonlinearity order individually or simply for performing all-depth, blind OCT data linearisation.

Keywords: Optical Coherence Tomography, dispersion compensation, spectral calibration, Group Velocity Dispersion, Third Order Dispersion, nonlinearities, neural network, Machine Learning

1. INTRODUCTION

Raw spectra acquired in Optical Coherence Tomography (OCT) generally possess two types of nonlinearity, and the linearisation methods target each type separately. The first type is induced by the spectrometer: the light is dispersed in a nonlinear fashion on the detector, leading to distortions of otherwise uniform fringes. It is corrected hardware-wise with prisms[1] or in post-processing by shifting the intensity values to their correct linear positions.[2] The second type comes from the unbalanced dispersion in the interferometer – it is mainly induced by a non-symmetric amount of glass in the interferometer arms. This is mitigated by inserting the correct amount of glass in one of the arms[2] or again in postprocessing by multiplying a raw spectrum by a complex, nonlinear exponent.[3]

If a bulk object with depth-varying dispersion is imaged (such as a whole eye), the induced nonlinearity will vary with depth as well. It is challenging to remove in one go because one can only correct for just one dispersion level. This is why current methods rely on piece-wise phase corrections[4] or a priori knowledge of dispersion variability[5] – both error-prone solutions, requiring user attention.

Most of the current linearisation methods require re-application whenever experimental parameter changes (e.g. setup misalignment occurs, imaged object changes). There are also iterative optimisation algorithms for blind linearisation,[3] but they are unable to compensate for object dispersion and are limited in their real-time application.

A Machine Learning solution was proposed where OCT images with different depth linearised are stitched into a fully-linear image,[6] but it is biased towards a specific kind of object. We propose an approach similar to Ref.[6] that is unbiased and universal and does not require re-optimisation for different experimental systems or conditions. In our approach, each nonlinearity order is removed with a separate neural network, enabling switching off of a single nonlinearity order. In addition, when the networks are used in a sequence, switching off all the nonlinearity orders is possible, enabling blind, all-depth OCT data linearisation. Here, we focus on the second- and third-order nonlinearities - the orders most relevant in the context of OCT.

Further author information: (Send correspondence to K.A.M. or S.M.K)
K.A.M.: E-mail: krzysztof.maliszewski@pg.canterbury.ac.nz
S.M.K.: E-mail: skol745@aucklanduni.ac.nz

2. CONCEPT

Our networks take a two-dimensional data structure - a stack (Fig. 1) - as an input, whose rows are created in the following way.

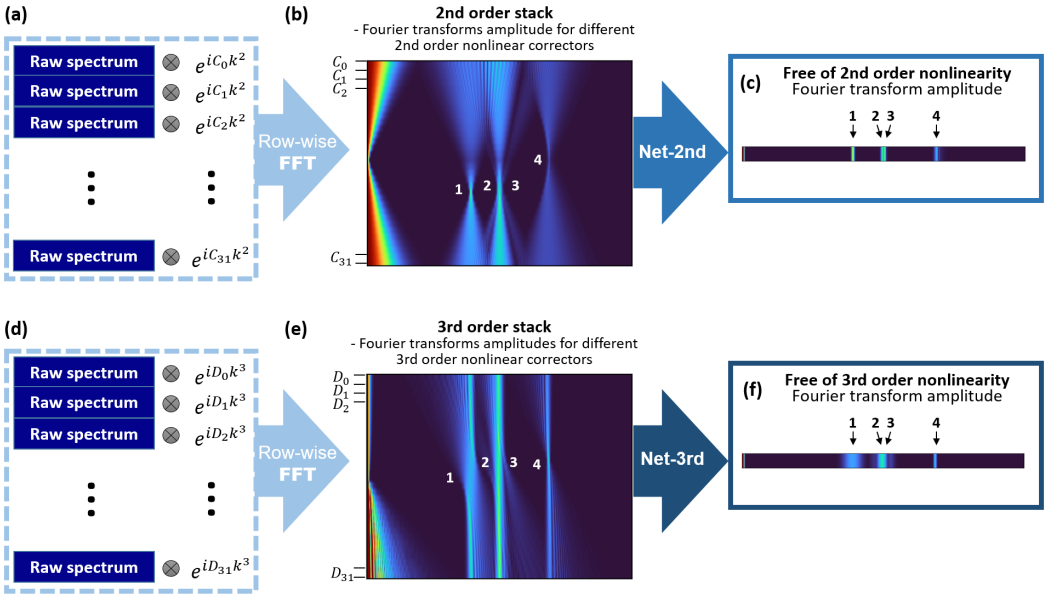

Figure 1. (a) The raw spectrum is multiplied by 32 different complex, quadratic exponents, and Fourier transformed to form (b) the 2nd order stack. This 2nd order stack is an input to Net-2nd which outputs (c) an A-scan free of the second-order nonlinearity (but containing the third-order nonlinearity). (d) The raw spectrum is multiplied by 32 different complex, cubic exponents to obtain (e) the 3rd order stack. This 3rd order stack is an input to Net-3rd which outputs (f) an A-scan free of the third-order nonlinearity (but still containing the second-order nonlinearity). 1, 2, 3, and 4 mark places in the stacks where the given nonlinearity is cancelled. The neural networks' processing can be viewed as identifying such regions and stitching them up.

First, the raw OCT spectrum is multiplied by a complex exponent. In the case of the stack used for removing the second-order nonlinearity, the 2nd order stack (Fig. 1a,b), this exponent is of the form of $e^{iC_nk^2}$, where i is the imaginary unit, C_n is a second-order coefficient with n - an integer enumerating the rows of the stack, and $k = [0, 1, 2, ..., M-1]$ with M being the length of the raw spectrum. Similarly, in the case of the stack used for removing the third-order nonlinearity, the 3rd order stack (Fig. 1d,e), the raw spectrum is multiplied by $e^{iD_nk^3}$, where D_n is the third-order coefficient.

Each row is Fourier transformed, and an absolute value is calculated to obtain a 2nd-order stack (Fig. 1b) or a 3rd-order stack (Fig. 1e). The range of the second-order and third-order coefficients is symmetrical around 0 and broad enough to ensure that each modulation's nonlinearity in the raw signal is compensated. In the 2nd-order stack, these will be the places with the smallest width (marked 1, 2, 3 and 4 in Fig. 1b) and in the 3rd-order stack, the places containing no modulation-like distortions (marked 1, 2, 3 and 4 in Fig. 1e).

2.1 Single order nonlinearity removal

The second order nonlinearity is removed by providing the neural network "Net-2nd" with the 2nd order stack. The output of this network is a Fourier transform amplitude, an A-scan, free of the second-order nonlinearity, but incorporating the third-order nonlinearity (notice the third-order-nonlinearity-related modulatory shape of peak 4 in Fig. 1c). Similarly, neural network "Net-3rd" takes the 3rd order stack as an input and outputs an A-scan free of the third-order nonlinearity, but incorporating the second-order nonlinearity (notice that peak 4 is no longer distorted by the third-order nonlinearity in Fig. 1f).

2.2 Both second- and third-order nonlinearity removal

Net-2nd and Net-3rd are used in a sequence to enable the removal of both the second- and third-order nonlinearity (Fig. 2) from an OCT spectrum. Since each network's output is a Fourier transform amplitude that is impossible to transform back to a spectrum, the input data needs to be prepared in a correct manner at the stack-generation stage to enable the proposed combined usage of the networks.

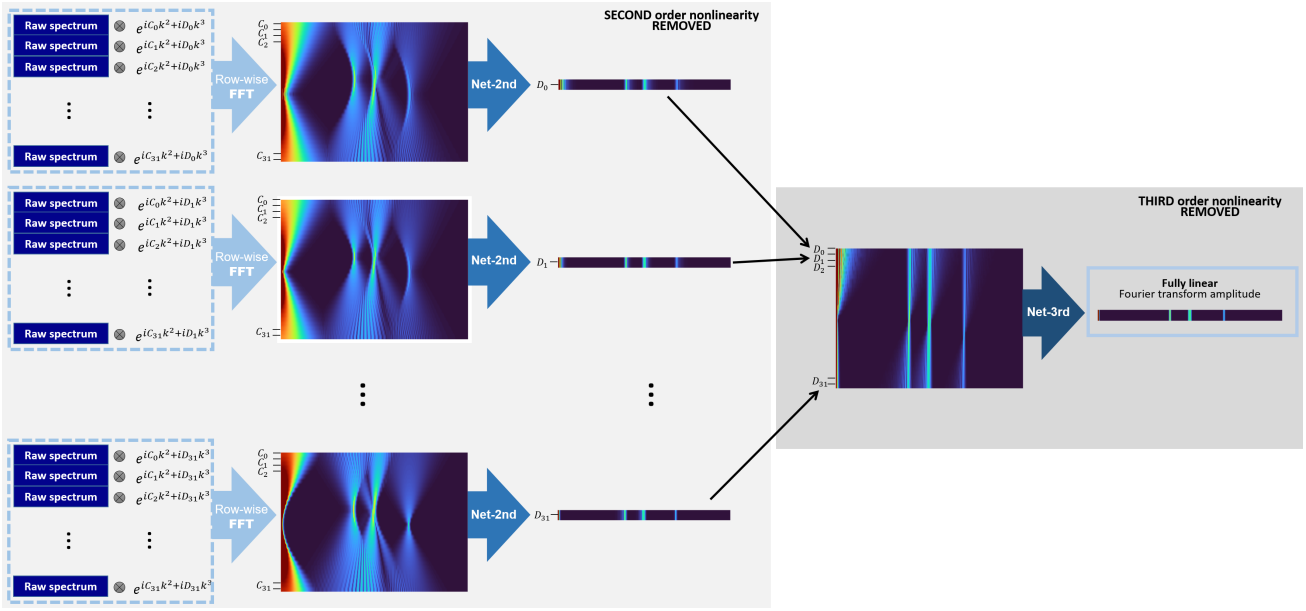

Figure 2. Multiple 2nd order stacks are generated, containing additional third-order nonlinearity introduced with a complex, cubic exponent: for each stack, a different third-order coefficient D is used. Net-2nd processes these stacks into Fourier transform amplitudes that are free of the second-order nonlinearity but still contain all the third-order nonlinearity. As a result, a 3rd order stack can be built and used with Net-3rd to provide an A-scan free of both the second- and third-order nonlinearity.

In general, to first remove the second-order nonlinearity and at the same time, enable subsequent removal of the third-order nonlinearity, Net-2nd is used multiple times (here, 32 times, as shown in Fig. 2) to output rows of a 3rd order stack needed by the Net-3rd for the third-order nonlinearity removal. More specifically, Net-2nd processes 32 2nd order stacks - each of these stacks incorporates a different third-order coefficient D introduced during the multiplication of the raw spectrum by a complex exponent. Because Net-2nd leaves the third-order nonlinearity intact, the 32 output Fourier transform amplitudes incorporate both the original, experimentally acquired third-order nonlinearity and the third-order nonlinearity introduced by the multiplication by a complex, cubic exponent. As a result, the 3rd order stack can be reconstructed where the second-order nonlinearity has already been removed. Such a 3rd order stack is used with Net-3rd to produce an A-scan free of both the second- and third-order nonlinearity.

3. TRAINING AND NETWORKS

3.1 Datasets

To facilitate the training of our neural networks, we used computer-generated datasets generated based on our OCT signal model. These datasets feature noiseless Gaussian-shaped signals that possess randomised content.

Specifically, our datasets represent a broad range of possible object configurations encountered in OCT imaging scenarios. Each simulated object has a random number of interfaces, ranging from 2 to 100, positioned at arbitrary locations. Furthermore, each object layer exhibits variability in reflectivity and has random second-order and third-order nonlinearity values. These randomised object parameters are used to synthesise modulatory signals equivalent to those acquired in actual OCT measurements.

From these in-silico signals, we construct input stacks along with corresponding A-scans along with their respective output A-scans, devoid of the corresponding nonlinearities. Each neural network is then trained using its own unique datasets. To optimise the performance, we tailored the training datasets size to contain 15,000 objects. The validation and test datasets included another unique 15,000 objects, chosen to maintain a uniform distribution across objects with varying interface counts.

3.2 Architecture

The neural networks are built upon a U-Net-inspired architecture, as depicted in Fig. 3. This architecture underwent customisation to adapt the output from a 2D structure to a 1D vector. This modification involved reducing the dimensionality of the final layers and extracting the maximum values of each column.

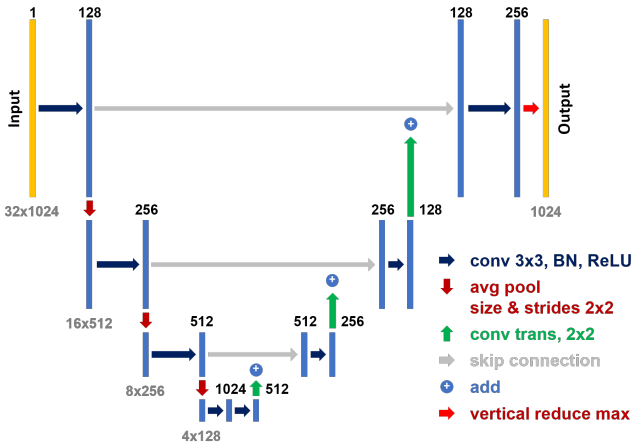

Figure 3. Net-2nd and Net-3rd neural networks utilise an optimized U-Net-inspired architecture. In contrast to the original U-Net design, these networks output a 1D vector by reducing the dimensionality of the 2D output image and extracting maximum values across columns. The architecture optimisation process was conducted using the Optuna framework.

Following the architectural modifications, the hyperparameters of the networks underwent further optimisation using the Optuna optimisation framework. We employed the Mean Absolute Error (MAE) loss function, Adam optimizer, a learning rate of 0.0002 with an exponential decay factor of 0.98 every second epoch, and a batch size of 8.

3.3 Training

During the training process, the Net-2nd and Net-3rd networks achieved convergence within just 15 epochs (Fig. 4). Subsequently, training was extended up to thirtieth epoch. By epoch 25, distinct behaviuors emerged between the two networks: while the Net-2nd network began to exhibit signs of stagnation, the Net-3rd network experienced a deceleration in loss improvement.

The observed behaviours may suggest differing sensitivities of the networks to the training data. Net-2nd's stagnation could imply that it reached its capacity to further learn from the small dataset. Conversely, the deceleration in loss improvement for Net-3rd might suggest that it is approaching a performance plateau.

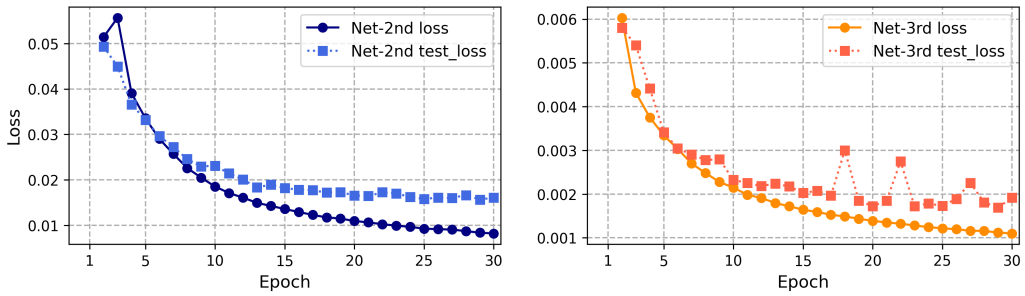

Figure 4. Performance evaluation of the neural networks using the Mean Absolute Error loss function across 30 training epochs. (a) illustrates the Net-2nd performance, and (b) showcases the Net-3rd performance. The solid lines represent the training loss, and the dotted lines depict the test loss. Results of epoch 1 are not depicted on the plots due to outlier values in the loss data.

4. RESULTS

We evaluated the performance of our networks on two experimental datasets. The first one represents a stack of glasses - a 50-μm-thick quartz on top of a 350-μm-thick sapphire - and contains 199 spectra. The other dataset represents an onion and contains 419 spectra. The datasets were obtained using a laboratory OCT setup employing a pulsed light source (Menlo Systems, central wavelength 1550 nm, total bandwidth 160 nm) and a fibre-based, spectral detection[7] (5km-long fibre spool and a single-pixel detector). The axial resolution of the OCT setup is 17 μm in air and the imaging range is 0.9 mm in air.

To evaluate the outputs of the proposed networks, the reference images were calculated using standard OCT image calculation approaches.[1] First, the raw spectra undergo traditional linearisation (spectral calibration and then dispersion compensation), and finally Fourier transformation is performed and the amplitude of the resulting complex signal is calculated to provide the image.

We perform the following 3 tests using the experimental datasets:

1. the datasets are processed with Net-2nd into images free of the second-order nonlinearity (all other orders, including the third-order nonlinearity, remain intact);

2. the datasets are processed with Net-3rd into images free of the third-order nonlinearity (all other orders, including the second-order nonlinearity, remain intact);

3. the datasets are processed with both Net-2nd and Net-3rd:

a) they are first processed with Net-2nd, and then with Net-3rd, using the approach presented in Fig. 2;

b) they are first processed with Net-3rd, and then with Net-2nd, using the modified approach presented in Fig. 2 where a series of 3rd order stacks are calculated and used in Net-3rd to provide a 2nd order stack suitable for application in Net-2nd.

While the first two tests enable evaluation of the networks in terms of their specific goal, Net-2nd for its capability to remove the second-order nonlinearity and Net-3rd for its capability to remove the third-order nonlinearity, the third test provides insight into the networks' combined performance.

4.1 Set of glasses

(Net-2nd testing) Fig. 5a shows an image obtained by directly Fourier transforming the spectra from the first dataset. Without prior spectral calibration and dispersion compensation, the uncompensated nonlinearities cause severe resolution degradation, making the image indecipherable.

If spectral calibration and dispersion compensation are applied to the spectra, Fourier transformation reveals an object (Fig. 5b). Here, the tilted lines correspond to the actual layer boundaries, whereas the perpendicular lines are "false" boundaries, called autocorrelation artefacts, and occur as a result of the interference of light reflected from the object interfaces.

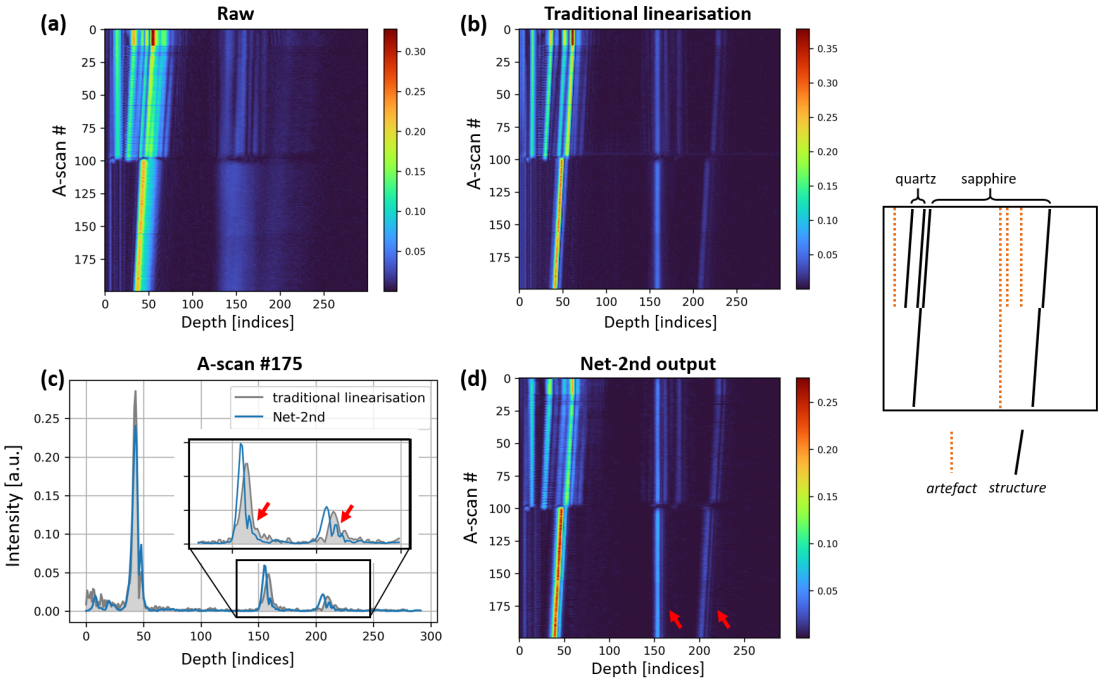

Figure 5. **(a)** Raw image, obtained by directly Fourier transforming the spectra, is of low quality due to uncompensated second- and third-order nonlinearities. **(b)** After linearisation of the spectra, Fourier transformation reveals the image elements: the imaged object structure and artefacts. **(d)** Processing of the raw, unlinearised dataset with Net-2nd removes the second-order nonlinearity, leaving the third-order nonlinearity intact. **(c)** Comparison of the 175th A-scan from the tradionally linearised image (gray line) with the same 175th A-scan from the Net-2nd output (blue line) shows that indeed the third-order nonlinearity is present in the Net-2nd output. The deeper peaks, marked with the red arrows, possess the characteristic one-sided modulatory shape. The **inset** shows schematically which image elements are artefacts (dashed orange lines) and which are the structure (solid black lines).

The raw spectra from the first dataset are processed using Net-2nd, providing an image in Fig. 5d. The second-order nonlinearity - which accounts for most of the image resolution degradation - is removed, but, as expected, one can see the third-order nonlinearity. It is observed as a modulation-like distortion of the deepest peaks (red arrows in Fig. 5d and in the 175th A-scan presented in Fig. 5c, where the gray line is the A-scan obtained as a result of traditional processing and the blue line is the Net-2nd output).

(Net-3rd testing) The same dataset was processed with Net-3rd. The output (shown in Fig. 6b) shows an image with the third-order nonlinearity removed. This statement is supported by two observations.

One, the part of the image corresponding to the thin layer of quartz is indiscernible in the raw image (marked with the yellow arrow in Fig. 6a), but it is discernible in the Net-3rd output (Fig. 6b, also yellow arrow). The third-order nonlinearity present in the raw image causes the distortion of the peaks (i.e. introduces a modulatory shape), leading to the scrambling of the quartz image. With the third-order nonlinearity gone in the Net-3rd output, the distortion is gone as well, enabling to distinguish both quartz interfaces. It should be noted that these two peaks are not resolved as well as they could be due to the remaining second-order nonlinearity.

Two, the elongation of a deeper element (marked with the green arrow in Fig. 6a) is gone and replaced with a symmetric Gaussian-like peak in the Net-3rd output (Fig. 6b, also a green arrow). Since the second-order nonlinearity leads to peak broadening and the third-order nonlinearity leads to the appearance of a modulatory "tail" at one side (as observed in Fig. 5c and d), the superposition of the two orders of nonlinearity creates a broad peak that is elongated on one side. The application of Net-3rd removes the third-order nonlinearity contribution, "restoring" the symmetric broadening caused by the second-order nonlinearity.

The closer inspection of these effects can be done by comparing A-scans from the raw image and from the

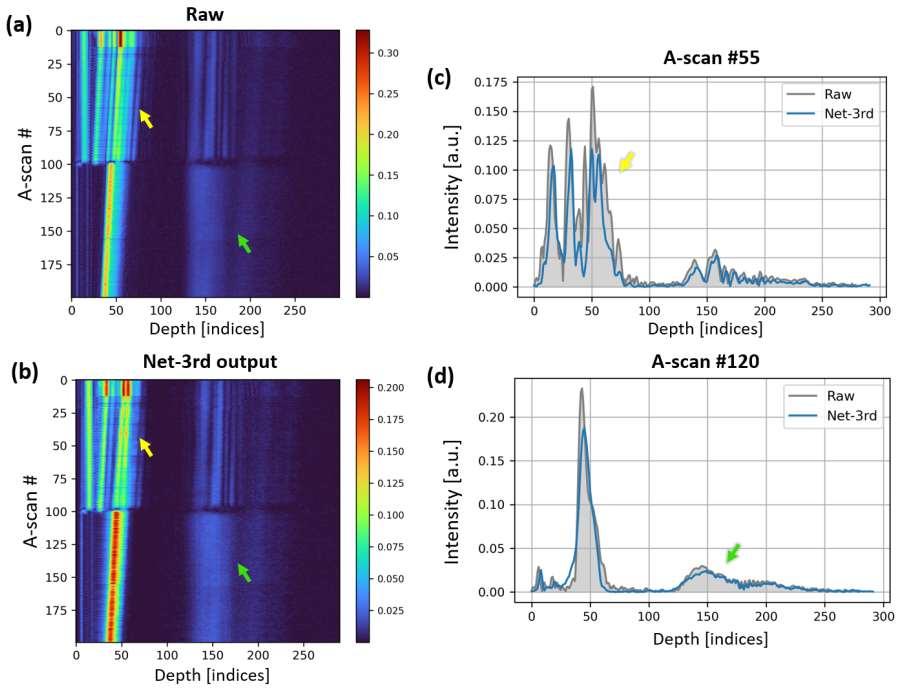

Figure 6. **(a)** The raw image is blurred and distorted by the second- and third-order nonlinearities. **(b)** Processing of the raw spectra with Net-3rd results in the removal of the third-order nonlinearity, leading to the retrieval of the structural details of quartz (yellow array) and restoring a symmetrical Gaussian-like shape of deeper peaks (green arrow). This can be better observed on plotted **(c)** 55th A-scan and **(d)** 120th A-scan from the raw image (gray line) and Net-3rd output (blue line).

Net-3rd output, presented in Fig. 6c and d (gray and blue lines, respectively).

(Net-2nd / Net-3rd combination tests) Finally, Net-2nd and Net-3rd were used in a sequence to enable the removal of both the second- and third-order nonlinearity. At first glance, the result of applying Net-2nd and then Net-3rd (Fig. 7b) is very similar to the result of first applying Net-3rd and then Net-2nd (Fig. 7c) and both approaches look to be in a good agreement with the image obtained using traditional linearisation (Fig. 7a).

A closer inspection of the area corresponding to quartz shows that the Net-3rd —> Net-2nd solution processed the dataset more favourably than the Net-2nd —> Net-3rd solution, and that both solutions underperformed as compared to the result of the traditional linearisation. This is also confirmed when A-scans from the image obtained using traditional linearisation and A-cans from the networks' output are compared (see A-scan #55 in Fig. 7d and e and A-scan #120 in Fig. 7f and g). Such comparison of A-scans also shows that the compound networks processing leads to intensity equalisation. More specifically, the height of the peaks placed at the front of the A-scans is much smaller.

4.2 Onion

The spectra in the second dataset were Fourier transformed to produce the raw image in Fig. 8a. These raw spectra were also linearised using the traditional approaches and Fourier transformed, providing a tradionally linearised image in Fig. 8b. Note that due to the low axial resolution of the OCT setup used for imaging, the cellular structure of the onion is not resolved as well as when high-resolution OCT setups are used. Also, the visual quality is decreased due to the presence of the autocorrelation artefacts cluttering the front of the image and possibly overlapping the onion structure.

The application of Net-3rd improved a bit the resolution of the image by removing the third-order nonlinearity (the subsurface layer starts to be discernible in Fig. 8c), but as expected, it was the application of Net-2nd - so the removal of the second-order nonlinearity - that provided a better quality image (Fig. 8d). Both combined

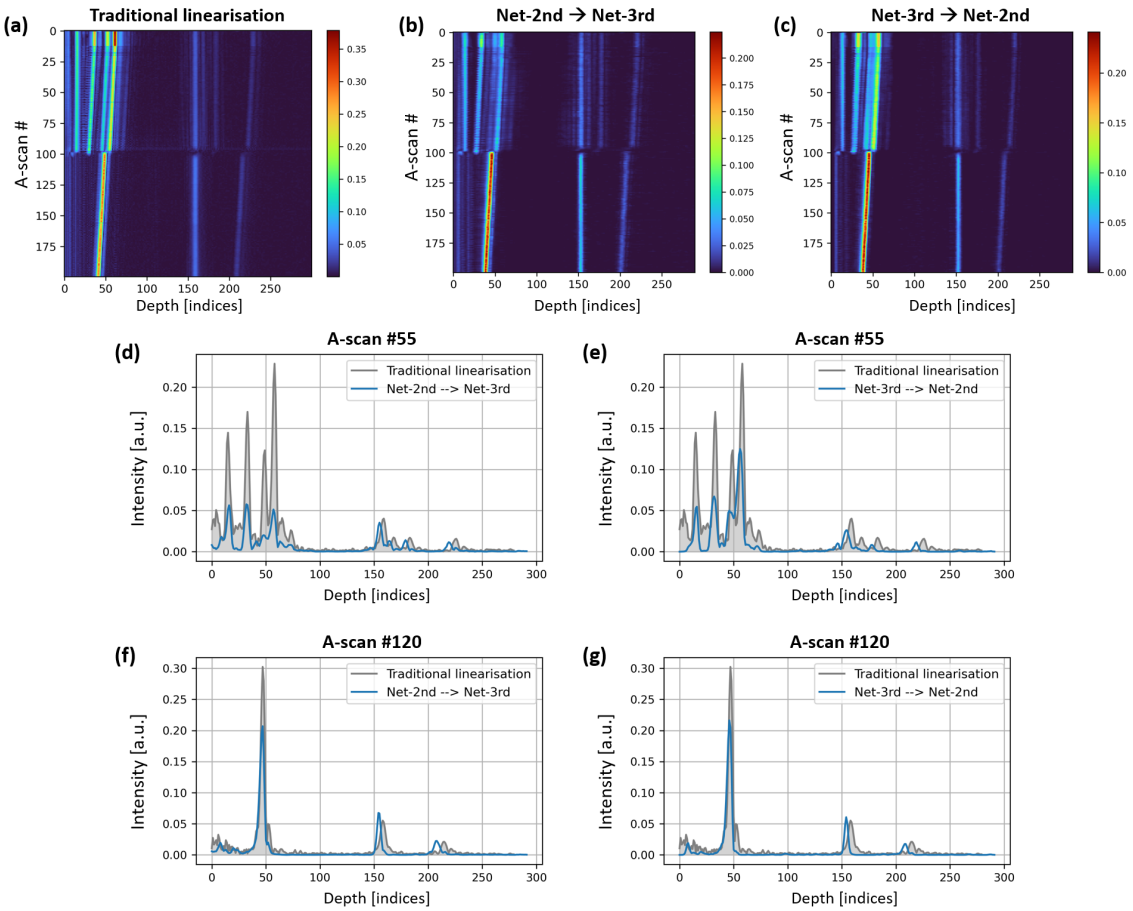

Figure 7. **(a)** Traditional linearisation removes both the second- and third-order nonlinearity, enabling to retrieve the structure of the imaged object. **(b)** Processing with Net-2nd and then with Net-3rd according to the algorithm presented in Fig. 2 **(c)** and processing with Net-3rd first and then with Net-2nd reveal the structure of the object free of the second- and third-order nonlinearity, but with a slightly decreased quality. **(d)** Comparison of the 55th A-scans from (a) and (b), **(e)** and from (a) and (c), as well as **(f)** the comparison of the 120th A-scan from (a) and (b), **(g)** and from (a) and (c) shows superiority of the Net-3rd —> Net-2nd solution over the Net-2nd —> Net-3rd solution.

networks solutions (Fig. 8a and f) were able to provide good-quality predictions of the fully linear image. The area marked with the red rectangle was zoomed in for the Net-2nd output and showed in Fig. 8d1. Again, as expected, one can see peak distortions related to the third-order nonlinearity. These distortions are gone in the Net-2nd —> Net-3rd solution (Fig. 8e1) as well as in the Net-3rd —> Net-2nd solution (Fig. 8f1). Comparison with the zoomed-in area from the traditionally linearised image (Fig. 8b1) suggests that, again, the quality of predictions is the Net-3rd —> Net-2nd solution is better than the quality of the other solution.

5. SUMMARY & DISCUSSION

We presented two neural networks for raw data linearisation in Optical Coherence Tomography (OCT). The first network, Net-2nd, removes the second-order nonlinearity, the biggest contributor to the resolution degradation in OCT images. The other network, Net-3rd, removes the third-order nonlinearity which leads to the peak distortions and is observed in high-resolution OCT or OCT systems incorporating a significant amount of unmatched dispersion. Using the datasets representing a stack of glasses and an onion, we confirmed that each network successfully removes the given nonlinearity order from the raw spectra (see Fig. 5 and Fig. 6).

A combined solution was also proposed (Fig. 2), aimed at the removal of both orders. The tests on both the

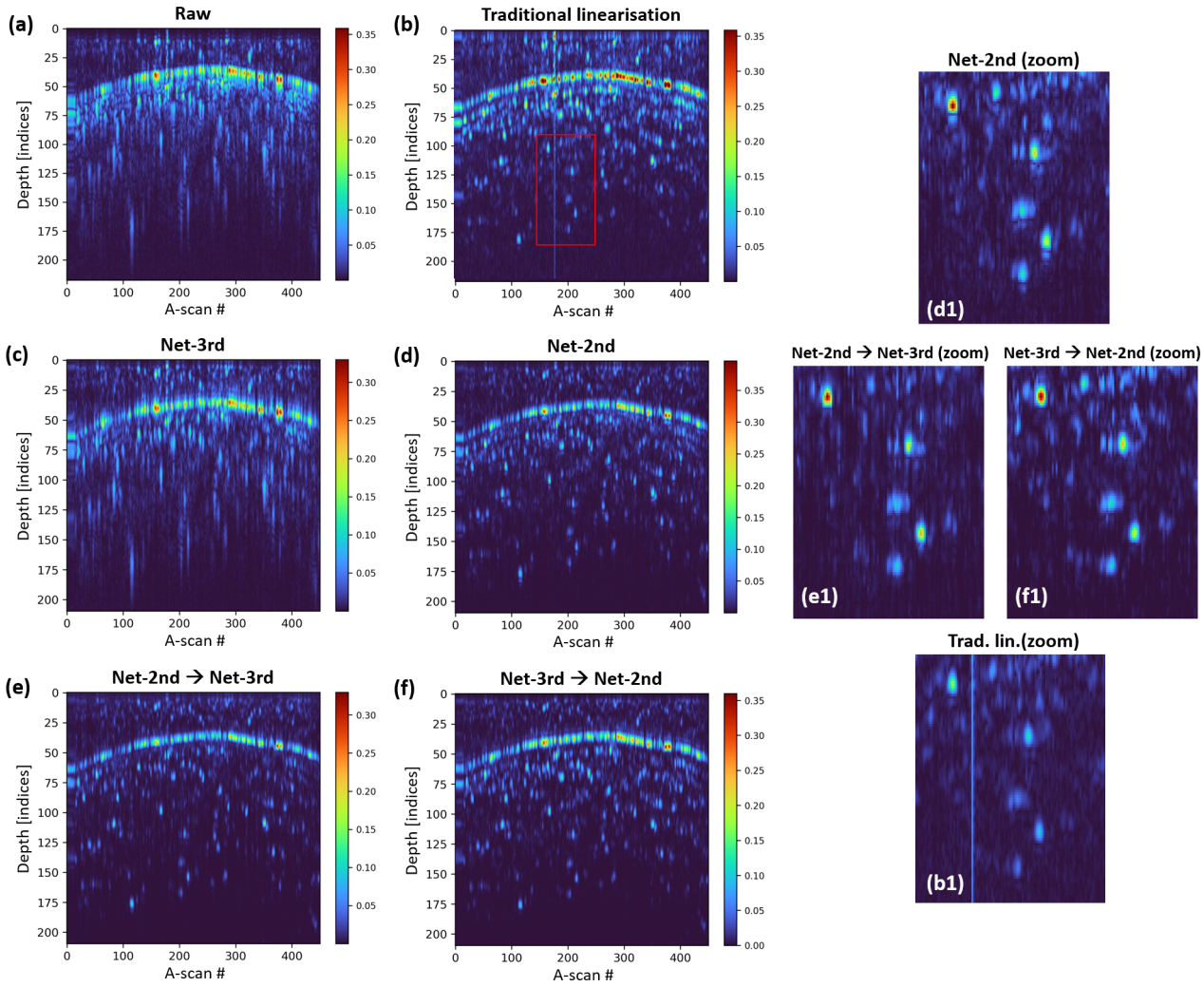

Figure 8. **(a)** Raw image of the onion. **(b)** Traditionally linearised image. **(c)** The output of Net-3rd where the third-order nonlinearity is removed. **(d)** The output of Net-2nd where the second-order nonlinearity is removed. **(e)** Compound solution where Net-2nd is first applied then Net-3rd, **(f)** compund solution where Net-3rd is applied first and then Net-2nd. An area marked with a red rectangle is zoomed for **(d1)** Net-2nd, revealing uncompensated third-order nonlinearity distortions, **(e1)** for the Net-2nd —> Net-3rd solution and **(f1)** for the Net-3rd —> Net-2nd solution, both showing no distortions and the quality similar to that of **(b1)** the traditionally linearised image.

stack of glasses and the onion showed that the application of Net-3rd before Net-2nd provides better predictions than the application of Net-2nd and then Net-3rd (Fig. 7b versus Fig. 7c and Fig. 8e1 versus Fig. 8f1). This assymetricity in the combined performance of the proposed networks might be related to the networks' individual performance (MAE smaller than 0.02 for Net-2nd versus MAE smaller than 0.002 for Net-3rd, Fig. 4). Net-3rd outputs high-fidelity signals which, when combined, produce a stack (here, a 2nd order stack) whose appearance is within the expectations of Net-2nd. More specifically, the reconstructed 2nd order stack does not deviate from the type of stacks Net-2nd was trained on (for example in terms of the peak shape) and consequently, Net-2nd can infer a better quality signal.

Further work will consist in optimizing the performance of the networks. The tests will also be done on datasets which are out of reach of traditional linearisation methods. More specifically, we plan to use the datasets which represent dispersive objects for which all-depth linearisation is impossible using traditional approaches.

ACKNOWLEDGMENTS

Sylwia M. Kolenderska would like to acknowledge the support of the New Zealand Ministry of Business, Innovation and Employment (MBIE), Smart Ideas grant "Extending the Boundaries of Digital Signal Processing: AI-powered Fourier Transformation Alternative (E7943)". All authors thank Oliver Batchelor, François Bissey, and Piotr Kolenderski for sharing their computational resources.

REFERENCES

[1] Wang, K. and Ding, Z., "Spectral calibration in spectral domain optical coherence tomography," *Chinese Optics Letters* **6**(12), 902–904 (2008).

[2] Lan, G. and Li, G., "Design of ak-space spectrometer for ultra-broad waveband spectral domain optical coherence tomography," *Scientific reports* **7**(1), 42353 (2017).

[3] Wojtkowski, M., Srinivasan, V., Ko, T., Fujimoto, J., Kowalczyk, A., and Duker, J., "Ultrahigh-resolution, high-speed, fourier domain oct and methods for dispersion compensation," *Optics express* **12**(11), 2404 (2004).

[4] Marks, D. L., Oldenburg, A. L., Reynolds, J. J., and Boppart, S. A., "Digital algorithm for dispersion correction in optical coherence tomography for homogeneous and stratified media," *Applied optics* **42**(2), 204–217 (2003).

[5] Pan, L., Wang, X., Li, Z., Zhang, X., Bu, Y., Nan, N., Chen, Y., Wang, X., and Dai, F., "Depth-dependent dispersion compensation for full-depth oct image," *Optics express* **25**(9), 10345–10354 (2017).

[6] Ahmed, S., Le, D., Son, T., Adejumo, T., Ma, G., and Yao, X., "Adc-net: an open-source deep learning network for automated dispersion compensation in oct," *Frontiers in Medicine* **9**, 864879 (2022).

[7] Kolenderska, S. M., Vanholsbeeck, F., and Kolenderski, P., "Quantum-inspired detection for spectral domain optical coherence tomography," *Optics Letters* **45**(13), 3443–3446 (2020).