

ACCEPTED MANUSCRIPT • OPEN ACCESS

DeepVID: deep-learning accelerated variational image decomposition model tailored to fringe pattern filtration

To cite this article before publication: Maria Cywiska *et al* 2023 *J. Opt.* in press <https://doi.org/10.1088/2040-8986/acb3df>

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2023 The Author(s). Published by IOP Publishing Ltd.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

DeepVID: deep-learning accelerated variational image decomposition model tailored to fringe pattern filtration

MARIA CYWIŃSKA*¹, KONSTANTY SZUMIGAJ¹, MICHAŁ KOŁODZIEJ¹, KRZYSZTOF PATORSKI¹, VICENTE MICO², SHIJIE FENG³, CHAO ZUO³, MACIEJ TRUSIAK¹

¹*Warsaw University of Technology, Institute of Micromechanics and Photonics, Faculty of Mechatronics, A Boboli 8, 02-525 Poland;*

²*Departamento de Óptica, Universitat de Valencia, C/Dr. Moliner, 50, 46100 Burjassot, Spain;*

³*Jiangsu Key Laboratory of Spectral Imaging & Intelligence Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

**maria.cywinska.dokt@pw.edu.pl*

Abstract: The algorithms defined as variational image decomposition (VID) constitute the state-of-the-art in solving the image prefiltration problem. However, the discussion about the advantages and disadvantages of different variational image decomposition models in the context of fringe pattern prefiltration is yet to be addressed and this work is the first one leaning into this issue. The unsupervised variational image decomposition (uVID) algorithm allows for automatic, accurate and robust preprocessing of diverse fringe patterns and introduces the parameters and stopping criterion for Chambolle's iterative projection algorithm to separate the fringes and background. However, determining the stopping criterion in each iteration is a severely time-consuming process, which is particularly important given the fact that in many cases thousands of iterations must be calculated to obtain a satisfactory fringe pattern decomposition result. Therefore, the idea of using convolutional neural network to map the relationship between the fringe pattern spatial intensity distribution and the required number of Chambolle projection iterations has emerged. That way, it is no longer required to determine the value of the stopping criterion in every iteration, but the appropriate number of iterations is known in advance via machine learning process. We showed that the calculation time is reduced on average by 3-4 times by employing the deep learning-based acceleration (DeepVID) without jeopardizing the overall accuracy of the prefiltration. This way an important progress in developing uVID algorithm features towards real-time studies of dynamic phenomena is reported in this contribution. For the sake of metrological figure of merit, we employ deep learning based solution, for the first time to the best of our knowledge, to accelerate powerful and well-established VID approach, not to bypass it completely.

Keywords: Interferometry, Fringe analysis, Fringe pattern filtering, Machine learning, Deep learning, Convolutional neural network, Variational image decomposition, Total Variation.

1. Introduction

The techniques allowing the non-invasive, label-free and precise measurement in the full-field of view are of high importance for solving the issues encountered by contemporary non-contact metrology. The state-of-the-art full-field optical measurement methods, i.e., interferometry [1-4], holographic microscopy [5-8], fringe projection [9] or the moiré technique [10], often lead to the result in a form of fringe pattern (interferogram, hologram, moiregram, shadowgram, etc.). In those techniques measurand is not given in a straightforward way but it is encoded in fringe pattern phase (or less frequently amplitude) map, which needs to be calculated

numerically from recorded intensity distribution. In the simplest way, the intensity distribution of fringe pattern can be described by [3]:

$$I(x, y) = a(x, y) + b(x, y)\text{fun}(\varphi(x, y)), \quad (1)$$

where $a(x, y)$ is background, $b(x, y)$ is fringe pattern amplitude modulation, $\varphi(x, y)$ is phase function and $\text{fun}(\varphi(x, y))$ describes the characteristic of fringe pattern, which depends on the recording method. In most cases the profile of recorded fringes is described by the cosine function. There are exceptions, however, such as quasi-trapezoid fringe patterns recorded using the moiré technique [10] or fringe patterns recorded using time-averaged interference microscopy [11], where the measured vibration amplitude is encoded in the argument of the Bessel function. It should be also noted that in the equation describing the fringe pattern intensity distribution, additional components may appear due to the noise (e.g., thermal/electronic from the camera, typically occurring additive noise, or multiplicative noise caused by the speckle phenomenon) or recording errors caused by, e.g., non-linearity of the projector/detector and resulting in the appearance of higher signal harmonics in the recorded image. After considering the above-mentioned factors, the intensity distribution of the fringe pattern can be presented in the form [3]:

$$I(x, y) = [a(x, y) + \sum_{m=1}^{\infty} b_m(x, y) \cos(m\varphi(x, y))]n_1(x, y) + n_2(x, y), \quad (2)$$

where m is the number of the harmonic periodic functions, $n_1(x, y)$ is a multiplicative noise and $n_2(x, y)$ is an additive noise. It can be clearly seen that fringe pattern analysis is a complex issue, thus algorithmic solutions used for this purpose are based on different mechanisms and there are simply a lot of them.

The most accurate solution for fringe pattern phase extraction is a multi-frame approach [1-4]. It can be clearly seen that the equation describing fringe pattern intensity distribution (Eq. (2)) has a number of components (variables), which can be considered as the unknowns. In the easiest way the phase function can be retrieved by creating the system of equations and finding the solution for phase map “unknown” $\varphi(x, y)$. This is the working principle of the algorithm called temporal phase shifting (TPS) [12-19], where the system of interferometric equations is defined by adding the phase shift (known [13-15] or unknown [16-19]) between the two coherent beams. In general, multi-frame approach has limited time resolution since multiple frames need to be collected and therefore the analysis of fast dynamic events or in an unstable environment is troublesome. This problem was solved by the hardware improvement, e.g., utilizing the polarization approaches and the multi-camera system to instantaneously record phase-shifted interferograms by different cameras [20] or use the pixelated phase-mask of a single camera [21]. Nevertheless, introduced modifications complicate and increase the cost of measurement system. For that reason, the single-frame fringe pattern analysis approach is a very attractive family of algorithmic solutions.

The most popular representative of the numerical methods allowing for phase estimation from a single frame is the Fourier transform technique [22]. This very capable and useful method has limitation by means of accepted input carrier fringes frequency. Precisely, Fourier transform method allows for the analysis of off-axis [23,24] interferograms, where cross-correlation terms are well separated from the auto-correlation term in the Fourier spectrum. The Kramers-Kronig relation [25] is a solution, which is able to deal with the slightly off-axis [26,27] interferograms, where cross-correlation terms are well separated from each other, but they are overlapping with the autocorrelation term. The localized Fourier transform representatives such as windowed Fourier transform [28], continuous wavelet transform [29] or other approaches like spatial carrier phase-shifting [30], or regularized phase tracking [31], are generally very capable but require a set of parameters to be fixed.

The algorithmic solutions dealing with the fully on-axis [32,33] interferograms are limited to multi-frame approach and there is no single-frame solution to this problem. The most

universal solution enabling the single-frame phase analysis in the widest range of fringe patterns is based on the Hilbert spiral transform (HST) [34]. It accepts quasi on-axis interferograms, where cross-correlation and auto-correlation terms are significantly overlapping but interfering beams are not perfectly aligned causing the appearance of low frequency carrier fringes in the recorded image. To perform well Hilbert spiral transform needs the input signal to be zero-mean valued, which can be achieved with the use of prefiltration algorithms.

It needs to be highlighted that the prefiltration algorithm used for preparation of the data for HST must be able to deal with the wide range of fringe patterns characteristics (shape of fringes, kind of noise and object under test etc.) and origin (different full-field optical metrology methods) so as not to nullify the advantage HST has over other known fringe pattern phase estimation solutions. Two numerical approaches are actively developed to fulfill this requirement: empirical mode decomposition (EMD) [35-37] and variational image decomposition (VID) [38-40]. The most important advantage of EMD-based solutions is the fact that they are purely data-driven and there is no need to set any parameters values to get the decomposition result in the form of a set of bidimensional intrinsic mode functions (BIMFs). Additionally, in the case of most recent solutions [41-45] the calculation time is short. On the other hand, the number of estimated BIMFs can be relatively high and the differentiation between fringe pattern background and fringes can be difficult due to the mode-mixing phenomenon, which was recently minimized in [46-48]. On the contrary, VID [49-57] provides the straightforward separation of background and fringes, but at the cost of increased numerical complexity and calculation time. Moreover, there is a number of parameters, which values need to be adjusted in order to get the accurate decomposition result.

Some of the problems were solved by the unsupervised variational image decomposition (uVID) approach [57], where all parameters were specifically tailored for the fringe pattern prefiltration and the whole calculations were automatized and significantly simplified. It was proven that uVID can be successfully used for the prefiltration of fringe patterns of biological cells [38,39,58,59], optical elements [57], arduous interferograms encoding the droplet shape [57], ESPI fringe patterns encoding cantilever deformations [40] or besselograms encoding vibrations of microsystems [60]. Nevertheless, uVID calculation time is still the issue to be addressed. The most time-consuming part of the uVID is calculation of stopping criterion called tolerance, which is also the most important parameter due to its ability to universally adjust the calculation process when the characteristic of fringe pattern changes, thus providing unique versatility and automation of uVID. All Chambolle projection iteration numbers mentioned throughout manuscript were estimated with the use of tolerance stopping criterion [57]. In short tolerance stopping criterion may be described in 4 steps (Eq. (3)). First, the difference between two consecutive steps (Δ) need to be obtained, then it should be divided by initial image (f) norm to free from the initial image pixel range influence on the final result, (Δ_{norm} , Eq. (3)):

$$\Delta = \|x_{n+1} - x_n\|_2, \quad \Delta_{norm} = \frac{\Delta}{\|f\|_2}, \quad d\Delta = \text{abs}\left(\frac{\Delta_{norm_{n+1}} + \Delta_{norm_n}}{2}\right), \quad \text{tol} = \frac{d\Delta}{\Delta_{norm_1}}. \quad (3)$$

Afterwards the derivative of Δ_{norm} is calculated ($d\Delta$, Eq. (3)). Finally, the derivative value is divided by the difference obtained for two first steps, (tol , Eq. (3)).

The question which may appear while deciding to use the VID scheme is a choice of the most suitable decomposition model. VID is a general term encompassing a family of algorithms, where image is decomposed during the minimization of some, especially tailored functionals. The definition of the minimized functional has the greatest impact on the decomposition results and each solution has its advantages and disadvantages, which need to be considered during matching the model to the specific application. Fringe pattern prefiltration is not an usual filtration problem due to the unique characteristic of the filtered images, such as its periodicity, self-similarity and clear distinction into three components: background, fringes and noise. In this contribution we perform the analysis of three the most commonly used VID

models: TV- L^2 , TV- H^{-1} and TV-G weighting their accuracy against the calculation complexity in the presence of different carrier fringes frequencies from fully off-axis to quasi on-axis setups. To the best of our knowledge such discussion appears for the first time in the context of fringe pattern background and fringes separation.

In this paper we propose the deep learning-based acceleration of uVID algorithm. Neural network-based solutions were already applied to support fringe pattern analysis on different stages, e.g., prefiltration [61-67], optimization of the window parameters for Fourier transform [68], phase estimation [69-75], phase unwrapping [76-82] and local fringe density map estimation [83]. Nevertheless, in this paper deliberately another approach was chosen than to directly determine the fringe pattern background filtration result by the neural network as an alternative to the solutions proposed in [66,67]. Developed convolutional neural network called DeepVID was taught to map the performance of the stopping criterion and to predict the number of iterations needed for uVID. We would like to highlight that the prefiltration algorithm has a crucial impact onto the final accuracy of fringe pattern analysis and therefore the outcome of the optical measurement. It should be mentioned that the use of neural network to replace the mathematically rigorous algorithmic solution such as VID may raise legitimate metrological concerns, as it was already noticed and discussed in [75, 83]. For that reason, in our work instead of developing another fully neural network-based solution the fringe pattern filtration is performed by classical, mathematically rigorous and versatile solution and only accelerated by the neural network, which in any way did not influence the final result. The working principle of proposed solution is presented in Fig. 1, which constitutes our novel approach.

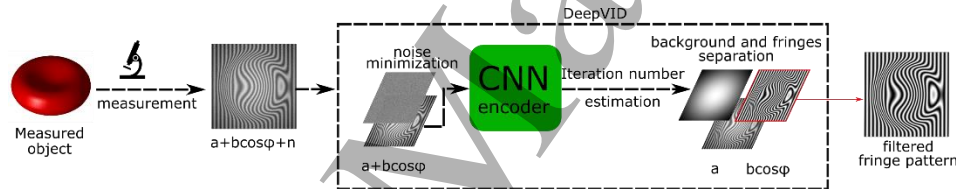


Fig. 1. Scheme of the proposed deep learning-based acceleration of the VID.

The paper is constructed as follows: section 2 provides the analysis of the VID model implemented in uVID approach comparing it with other known models and describes in detail which part of the automatized model is the most computationally expensive, section 3 describes the proposed acceleration of uVID-based on the convolutional neural network and its numerical verification, section 4 contains experimental evaluation of the proposed algorithm, whereas section 5 concludes the paper.

2. Analysis of automatized variational image decomposition model with regard to fringe pattern filtration

The estimation of the decomposition components in VID is performed by the minimization of the appropriately selected functionals. In the context of fringe pattern prefiltration in order to obtain the background-fringes separation both of the components need to be modelled in the selected well-suited function space, where minimized functional is defined as the norm. Generally, in the variational models the norm connected with the background part estimation is defined with the use of quantity called Total Variation (TV) [49]. The function space for which the TV is finite is the space of the functions with bounded variation (BV space). Because of the TV properties the BV space functions are smooth but can have discontinuities along line, which means that the image edges are not going to be smoothed.

The selection of function space for modelling the fringes part of the decomposition is a much more complicated issue. Historically first function space used for the definition of the image texture (in the context of fringe pattern prefiltration oscillatory part of the image) was L^2 space [49]. The minimization with the use of L^2 norm assumes that the texture has an overall small energy and disregards the oscillations. For that reason, the TV- L^2 decomposition model

provides the poor differentiation between the fringes (oscillations) and noise. The improvement was made using two approaches based on gradient projection minimization and frequency analysis. The first one was done with the use of the norm in the Banach space called G space [50-52]. The functions on G space can preserve the local characteristics well and have a small norm for the oscillating patterns. Because of that, fringes can be easily caught with an energy minimization process and distinguished from the noise. The second approach is performed using the Hilbert (H^{-1}) space [53], where the norm is calculated using weighted Fourier transform. This norm can favor the oscillation simply by adjusting the weights that are large for the low frequencies and small for high frequencies. Additionally, the Hilbert norm is expected to be low for locally parallel structures (fringes) and high for generally random and uncorrelated noise. In this section we will provide the comparison between three different VID models: TV- L^2 , TV-G (further applied in unsupervised VID algorithm [57]) and TV- H^{-1} expanding the understanding of their use for a specific task of fringe pattern prefiltration and the role of the number of performed iterations. It is the first such analysis reported and adds to the overall novelty of the paper.

2.1. Unsupervised variational image decomposition model performance in comparison to other known models

The state-of-the-art in the VID models, especially in the context of fringe pattern prefiltration, provides the three components decomposition, which means that the noise component is minimized prior to the background and texture components separation. For that reason, in this work, we focus solely onto the background and fringes separation and assume that the denoising was already accurately performed with the use of, e.g., block-matching 3D denoising algorithm (BM3D) [84,85]. The capability of the BM3D in the context of fringe pattern denoising was already discussed and proven in our previous works [57,86]. Here in our further numerical discussion, we simulated the noiseless fringe pattern series satisfying the relation:

$$I(x, y) = a(x, y) + \cos\left(\varphi(x, y) + \frac{2\pi}{T}x\right), \quad (4)$$

where x, y are the image coordinates simulated using Matlab function `meshgrid(1:512)`, $a(x, y)$ is a Gaussian function simulated low-frequency fringe pattern background component, T is carrier fringe period and $\varphi(x, y)$ denotes phase modulation simulated here as the shape of an unstressed red blood cell described by the equation of Evans and Skalak [87]. The dynamic range of simulated phase function was kept on the relatively small level of 1.23 rads with maximum phase difference in adjacent pixels equal to 0.02 rads. That way our numerical analysis was mainly focused onto the sensitivity of different models rather than specific local characteristics of analyzed fringe patterns since introduced phase function does not strongly disturb the shape of carrier fringes. During the simulations the fringes period is changed in the extremely wide range from 4 px (high frequency fringes close to the limits defined by Nyquist sampling theorem) to 172 px (corresponding to the case with 3 fringes in the field of view, which is a troublesome situation for background and fringes separation).

The results presented in Fig. 2 prove that the variational image decomposition keeps the prefiltration errors on the reasonable level for the whole range of simulated carrier fringes periods in the case of all analyzed decomposition models, which proves the universality of this classical fringe pattern prefiltration approach. Nevertheless among all analyzed models, the TV-G model provides the most accurate and stable results and it was proven that TV-G model outperforms other VID models in the context of fringe pattern prefiltration both considering calculation time and its accuracy. It can be seen in Fig. 2(a) that in the range of lower and medium carrier fringes periods (off-axis and slightly off-axis cases) TV-G provides the highest prefiltration accuracy and in the range of high carrier fringes periods (above 80 px) similar and the most accurate results were provided by TV-G and TV- H^{-1} . The differences in the performance of three analyzed VID models are subtle, especially in the case of low fringes periods. As it has already been mentioned the main difference between TV- L^2 , TV-G and TV-

H^{-1} is their ability to separate the fringes and noise. Changing the model parameters one can control the denoising process keeping in mind that there is a trade-off between the level of noise left in the image and preservation of low fringes periods. Once the noise has already been minimized mentioned trade-off is no longer a problem and the regularization parameters can be set arbitrarily high without the risk of preservation of the noise together with the low fringe periods. Nevertheless, the TV- H^{-1} model is the most computationally expensive, which can be clearly seen in Fig.2(b), where the calculation time of discussed models is presented. The TV- L^2 model is the simplest and the fastest one, but at the same time it provides the weakest differentiation between high fringes periods and the background. In the context of fringe pattern background and fringes separation the TV-G model was chosen for further improvement in uVID algorithm because it provides the most accurate results without adding unnecessary computational complexity.

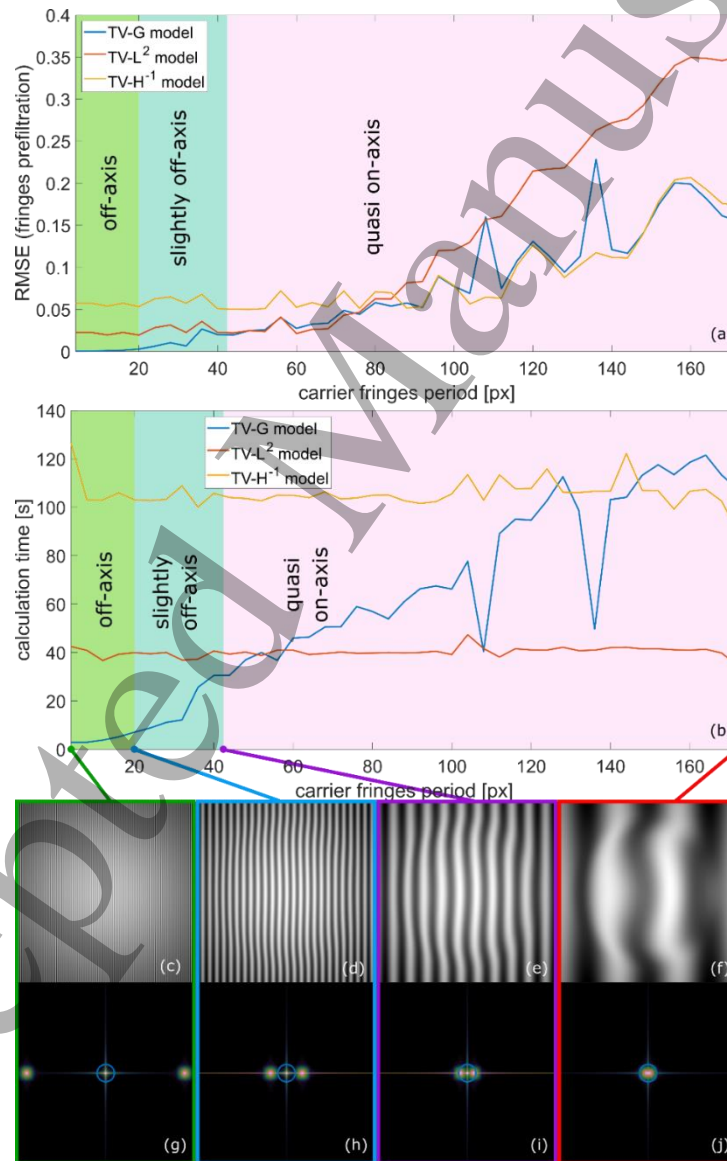


Fig. 2. The performance comparison of three different VID models: (a) fringe pattern prefiltration accuracy, (b) time of the calculations, (c), (d), (e), (f) chosen characteristic fringe patterns and (g), (h), (i), (j) their respective Fourier spectra.

2.2. Time cost of unsupervised variational image decomposition stopping criterion

The main reason why, in the past, the fringe pattern prefiltration with the use of TV-G model was an arduous process was a need of adjusting the parameters values with the change of fringe pattern characteristics. The most troublesome issue to address is the fact that with the change of carrier fringes periods the number of iterations needed for the background and fringes separation changes. The mentioned relation can be clearly seen in Fig. 3, where using the simulated data with known ground truth prefiltered images the number of iterations was defined by the minimum root-mean-squared error (RMSE). Nevertheless, the relationship between the input fringe pattern characteristics and the proper number of TV-G iterations is not straightforward since there are a lot of factors to consider, e.g., the carrier fringes period, the dynamic range and complexity of measured phase function, the contrast of fringes, the background characteristics, etc. The introduction of the stopping criterion for the TV-G iterative process [57] allowed for the automation of calculations and made its use very simple. The reliability and versatility of the stopping criterion performance was achieved at the expense of the increase of the algorithm's numerical complexity. The calculation time of a single TV-G iteration increases 9.5 times (from 0.002 seconds to 0.019 seconds) for the small (100 x 100 px) image size and 31.2 times (from 0.33 seconds to 10.3 seconds) for the image size of 2600 x 2600 px after the introduction of stopping criterion, which is caused mainly because costly matrix norms calculations. The calculation time analysis throughout this paper was performed with the use of portable personal computer with parameters: Intel Core i5-10310U(CPU), 16GB(RAM), Windows 10(OS) and Matlab computational environment. As it can be clearly seen on the chart presented in Fig. 4 the calculation time of the single TV-G iteration increases exponentially with the increase of the analyzed image size. Keeping in mind that for some fringe patterns there is a need of the calculation of few thousands of iterations to get a good quality fringes and background separation one can notice that the simplification of a stopping criterion is an issue to be addressed.

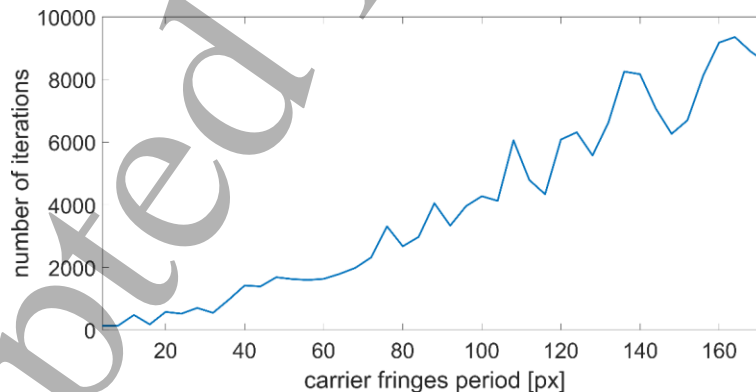


Fig. 3. Relation between the needed TV-G iteration number and the change of fringe pattern characteristic.

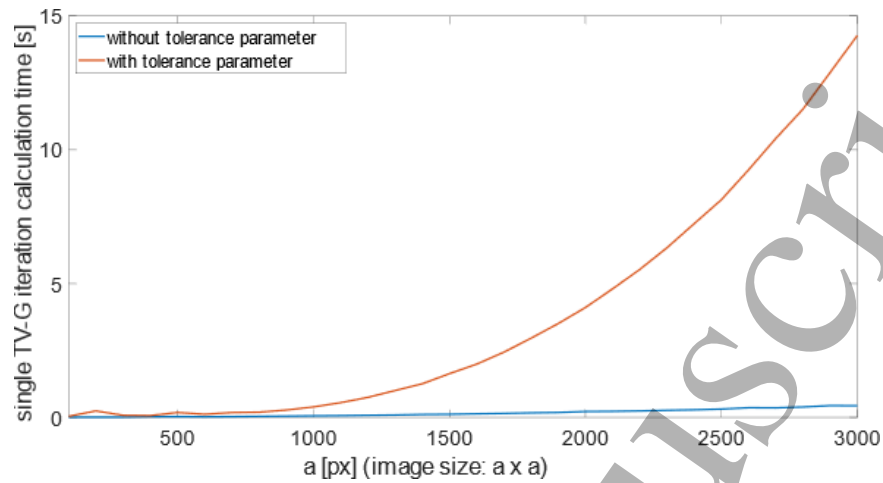


Fig. 4. Relation between the image size and the single TV-G iteration calculation time before and after introduction of tolerance parameter.

3. Fully automatic and accelerated variational image decomposition dedicated to fringe pattern filtration

As it was already mentioned the definition of the relationship between the needed number of TV-G model iterations and analyzed fringe pattern characteristics is not an easy task. For that reason instead of searching for the specific analytical and numerical solution we decided to use the advances of programming based on neural networks to define this crucial uVID parameter. In this approach the algorithm's input-output relationship is found during the neural network learning process and there is no need for its *a priori* analytical definition. In general, in the case of image analysis the most suitable and flourishing type of neural networks are convolutional neural networks (CNNs) [62-72, 74-83], where the basic mathematical operation in each neuron is defined by the convolution.

From the analytical point of view in order to find a solution to the complicated problem one should break it down to a combination of several simpler problems and find their solutions. In the case of imaging the basic numerical operation describing it is convolution. Similarly to the analytical approach the CNNs using the convolutional layers are walking towards the searched solution to the complex problem simplifying it in each layer. Keeping that in mind it is clear that the deeper the neural network is, the more complex problems it can address. The advancement of neural network-based programming over the classical one lays in the definition of the basic blocks building the searched solution, in which case they are not set arbitrarily by the algorithm's author but they are found during the network learning process (either supervised or unsupervised). In the case of the DeepVID network we decided to work with supervised learning. There are two reasons supporting the simplicity of this approach: we are dealing with noiseless images and in general the fringe patterns (interferograms, holograms, moiregrams, shadowgrams, etc.) can be described as spatially self-similar patterns, which makes it easy to receive the accurate learning outcome using purely simulated data [83].

The data simulated for the DeepVID supervised learning process consist of 10000 images from which 9400 images are building the training dataset and 600 images are used for the validation during training, but are of the same origin as training dataset (simulated with the phase function simulated by tailored polynomials). Additionally, the test dataset of 100 images was simulated, which were not seen by the network during training and are of different origin than training dataset (the phase function defined as unstressed red blood cell shape [87], which is not a polynomial). Test dataset will be used throughout the paper for DeepVID performance analysis. The task given to the DeepVID neural network is to map the performance of tolerance stopping criterion [57] but without the troublesome and time-consuming calculation of criterion

value in every Chambolle projection iteration [88]. The accuracy and robustness of proposed stopping criterion was discussed and proven in [57] and in the scope of this paper we are not aiming at improving the accuracy of uVID solution but to accelerate it while maintaining its high universality. Based on that the DeepVID training data was labelled calculating the tolerance stopping criterion and finishing the iterative process after appropriate (defined in [57]) level of criterion value is achieved. Both training and validation datasets consisted of different carrier periods so 3 to dozens of fringes were visible in the image. Dataset phase functions were simulated with the use of two-dimensional polynomial functions of orders from 2 to 5 sampled across X and Y axes. In order to simplify the dataset in the simulated interferograms the background function was set as constant, as the uVID iteration number is mostly influenced by the characteristic of the fringe pattern rather than the background modulation. This statement was confirmed by the analysis presented in Fig. 5, where all four presented curves calculated for different types of background modulation are highly similar. The maximum difference between the estimates of the number of iterations in the presence of different background types does not exceed 3% of the nominal value.

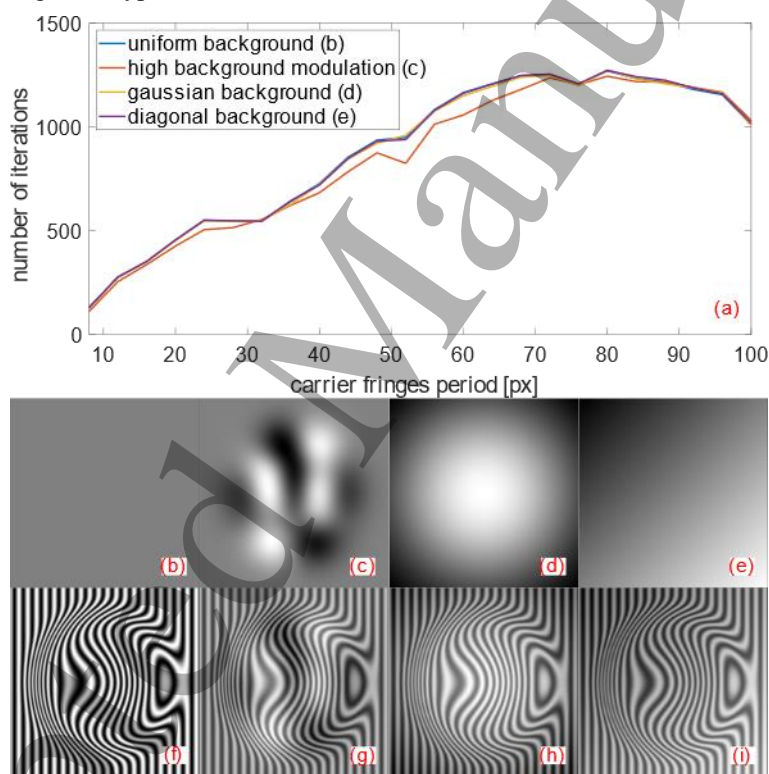


Fig. 5. Influence of the background modulation onto estimated Chambolle projection iteration number: (a) relation between iteration number and carrier fringes period with the presence of different types of background, (b) uniform background, (c) high background modulation, (d) gaussian background, (e) diagonal background and examples of the analyzed fringe patterns with (f) uniform background, (g) high background modulation, (h) gaussian background, (i) diagonal background.

3.1. Proposed DeepVID network architecture

As it was already highlighted the purpose of the proposed neural network is to predict the number of TV-G model iterations needed for successful fringe term decoupling for a given fringe pattern. The approach that has been taken is different than in majority of neural network related papers, because neural network is not going to replace the mathematically reliable

classical fringe pattern prefiltration approach. Another application of neural networks has been presented here, similar to the one presented in [68], where neural network was used for optimization of window shape in windowed Fourier transform instead of fully replacing the Fourier transform performance. The input of DeepVID neural network is assumed to be 512x512 px grayscale image while the returned output is a single floating number. Even though the single-valued output could indicate that we are facing a classification problem, the problem to be solved by the neural network is regression in nature. The neural network input-output relationship we are looking for during the learning process has a continuous distribution depending on fringe pattern characteristic rather than some discrete, categorizing values.

The proposed network architecture is a simple CNN encoder, shown in Fig. 6, consisting of 8 convolutional layers with ReLU as an activation unit and batch normalization. The input of the network is a 512x512 px grayscale image, while with the last convolutional layer the network proceeds to the output of the network as 1x1. In order to prevent overfitting, as it was suggested in [89], the 50% dropout regulation was applied on the first convolutional layer. Preventing overfitting is extremely important in our approach, because the training process was performed using purely simulated data with the phase function defined by tailored polynomials, while the target is to use DeepVID on experimentally recorded data, where phase function does not necessarily have to be defined using polynomials.

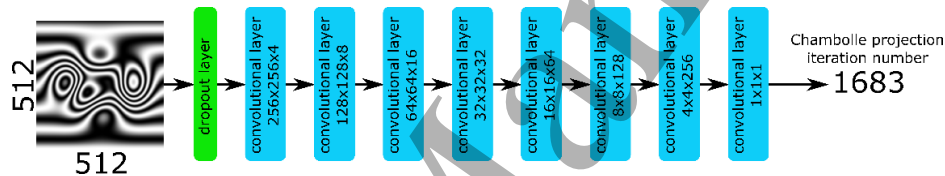


Fig. 6. Scheme of the proposed DeepVID architecture.

The network architecture examined in this paper was trained during the process of the supervised learning with the use of a computer-generated fringe pattern data set. Platform used for deep network training: Intel Core i7-11800H(CPU), NVIDIA GeForce 3050Ti(GPU), 64GB(RAM), Windows 10(OS), Matlab Deep Learning Toolbox™(Software). The learning parameter values, like learning rate (0.001), type of optimizer (ADAM), number of training (9400) and validation (600) data samples have been selected after many trial iterations. During the learning process proposed network learn to match the input image (fringe pattern) to a target label (number of Chambolle projection iterations [88]). The loss function used during the training is a half-mean-squared-error and is defined as follows:

$$loss = \frac{1}{2N} \sum_{i=1}^N (y(i) - t(i))^2,$$

where N is mini-batch size, $y(i)$ is the neural network estimate and $t(i)$ is a known ground truth for i -th data instance. Network trained that way can subsequently be used to perform a prediction on a new input image without knowing the ground-truth iteration number and therefore significantly reduces the calculation time. During the analysis of different sets of learning parameters values we found the interesting relationship between the neural network accuracy and the mini-batch size, which is presented in Fig. 7 using the test dataset. The bigger the mini-batch size, the more accurate learning outcome of DeepVID network. This relationship between neural network accuracy and mini-batch size may seem unusual at a first sight, since we would expect that mini-batch size has negligible effect on learning accuracy, but it can be easily explained. The very small mini-batch size connected with the dropout regularization results in insufficient generalization of neural network learned model due to the reduced cross-correlation between individual data instances from the training dataset. In other words, we are preventing the overfitting of the neural network so effective that we introduced underfitting.

As it can be seen in Figs. 7(a)-7(c) small mini-batch size results in fixation of the network outcome around a certain constant value since during the learning process the network does not comprehend the general features of the whole training dataset. Nevertheless, underfitting is not an issue that cannot be easily handled by adjusting the training parameters. With the increase of the mini-batch size the neural network learning outcome starts to fit to the expected line defined by the ground truth. The best alignment was achieved for the mini-batch size equal to 350, see Fig. 7(f), which at the same time was a learning platform computational capacity limit and that parameter defines the final DeepVID neural network learning process. We chose this way to deal with the underfitting instead of simply removing the dropout layer, because without the dropout layer regardless the mini-batch size the neural network overfits to the simulated training dataset and cannot generalize correctly for the experimental reality.

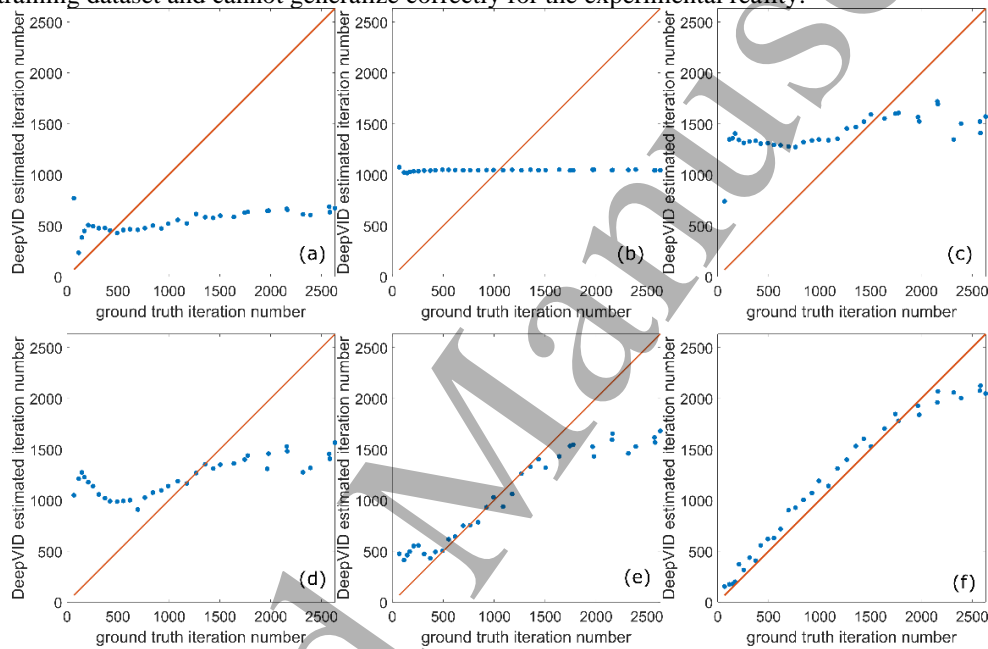


Fig. 7. Scatter plot (blue dots) of the predicted value of iterations vs their true values (red line) for mini-batch size equal to (a) 1, (b) 2, (c) 10, (d) 100, (e) 200 and (f) 350.

3.2. Deep learning-based acceleration of VID

The goal of the work presented in this paper was the acceleration of the algorithmic solution called uVID without compromising the quality of estimated filtration results. In order to fully validate the efficiency of proposed DeepVID solution we presented in Fig. 8 three charts estimated with the use of test dataset: first verifying the accuracy of our neural network in mapping the performance of tolerance stopping criterion (Fig. 8(a)), second presenting the gain in the calculation time (Fig. 8(b)) and third proving that the VID-based fringe pattern filtration accuracy was not decreased in comparison with uVID (Fig. 8(c)). Analyzing the results presented in Fig. 8(a) and 8(c) it can be clearly seen that DeepVID maps the performance of the tolerance parameter without compromising the filtration accuracy. Additionally, for all images included in the test dataset the calculations were significantly accelerated. The calculation time needed for the fringe pattern prefiltration was shortened around 2-4 times (3 times on average). In general, the reduction of the calculation time remains at the level of 60-75% for the whole range of analyzed fringe patterns.

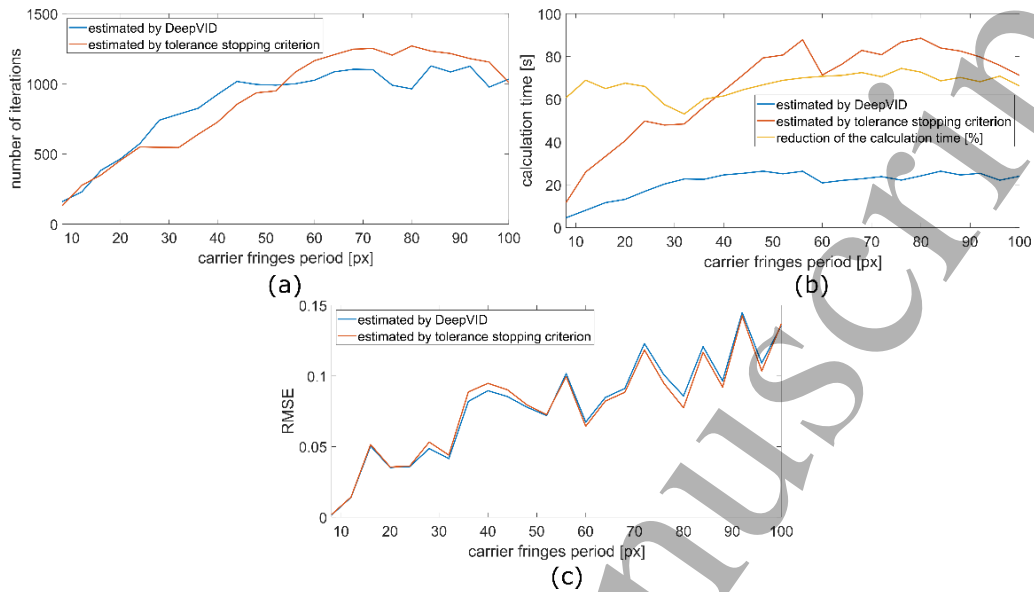


Fig. 8. Analysis of DeepVID performance on simulated test dataset. Relationship between carrier fringes period and (a) estimated number of iterations, (b) calculation time and (c) RMSE of filtration results.

4. Experimental verification

Even though our neural network-based solution was taught on the simulated fringe patterns (with fixed phase function and varying carrier fringes period) it should ultimately work on the experimentally recorded data. As it was already mentioned the biggest advantage of using the tolerance stopping criterion was its universality and independence from the specific fringe pattern characteristic and origin. For that reason, we decided to verify the performance of DeepVID with the use of four contrasting cases: the TPS series of interferograms of single polystyrene microsphere of approx. 90 μm in diameter, single interferogram of healthy prostate cells (RWPE) [90], single interferogram encoding the droplet shape [91-93] and besselogram obtained as the amplitude from the TPS series of temporal averaging interferometry based interferograms [11, 60]. It should be highlighted that DeepVID solution does not perform purely neural network-based fringe pattern prefiltration. It is not dedicated to the specific type of data as it is in the case of the filtration performed fully by the neural network for, e.g., fringe projection [66]. DeepVID is thus as versatile as the capable and well-established uVID itself [57]. As it was explained, in this paper we are focused onto the fringes-background separation with previously minimized noise component and for that reason all experimentally recorded data was prefiltered with the use of BM3D denoising method.

In Table 1 the results of the analysis of 5 phase-shifted interferograms are presented. The complexity of the analyzed data should be highlighted since we are dealing here with closed fringes, which are very troublesome for single-frame fringe pattern analysis techniques, especially around central fringe. Additionally, some disruptions and diffraction effects are clearly visible at the boundary between the microsphere surface and the medium. Despite this, the DeepVID solution provided sufficiently accurate results since the maximum difference between number of iterations defined by tolerance stopping criterion (considered as a ground truth) and neural network is 66 iterations. Furthermore, the prediction inaccuracy did not influence the filtration results since RMSE stays on the same level for both approaches. Similarly to the case of simulated test dataset the calculations were shortened by 3-4 times.

The interferogram of healthy prostate cells presented in Table 2 is a representative of different type of fringe patterns, where carrier fringes period is sufficiently low not to cause the

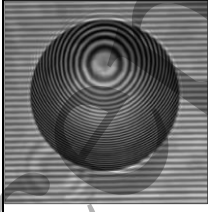
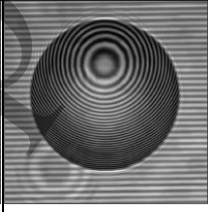
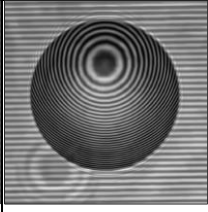
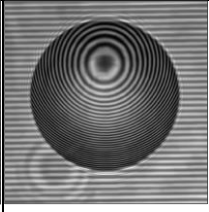
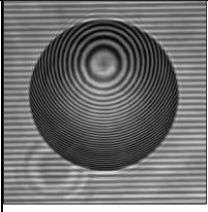
appearance of closed fringes. In that case the Chambolle projection iteration [88] number predicted by DeepVID is highly accurate since the difference between neural network-based estimate and number defined by tolerance stopping criterion is only 6 (~2% of nominal value). The reduction of calculation time is significant and consistent with our previous analyses since it was shortened by 2.4 times.

Highly cumbersome case of fringe pattern is presented in Table 3, where the interferogram encoding the droplet shape is presented. It contains of regions with very dense fringes and at the same time of regions of extremely low density. Nevertheless, even in that case the accuracy of the uVID iterations number estimated by DeepVID was proven. The uVID-based number of iterations is equal to 519, while estimate retrieved by DeepVID is equal to 547. The capability to deal with that type of data in the case of uVID was already discussed [57], while here it is demonstrated that DeepVID does not limit in any way the analysis options available with the use of standard uVID approach. The gain in the meaning of calculation time reduction is also clearly noticeable since the calculations were shortened by 4.3 times.

Finally, in Table 4 completely different fringe pattern is presented where the profile of fringes is not described by cosine function but by the Bessel function. Even though the DeepVID neural network was trained with the use of typical, cosine fringe patterns the prediction of uVID iterations number is highly accurate. The uVID number of iterations estimated with the use of already tested and validated for besselograms filtration [60] tolerance parameter is equal to 1272, while DeepVID predicted iteration number is equal to 1203. The reduction of calculation time is the most noticeable and because of the high number of needed iterations calculations were shortened by 4.7 times.

The examples of fringe pattern background and fringes separation with the use both DeepVID and uVID approach of the analyzed images are presented in Fig. 9, where it was proven that there is no noticeable difference between uVID and its neural network accelerated version. To prove that, we used Quality Index (QI) [94]. QI takes values from -1 to 1, where maximum similarity between analyzed and reference images is received for QI=1. For fringes presented in Fig. 9(b) and Fig. 9(d) QI=0.9991, for fringes presented in Fig. 9(f) and Fig. 9(h) QI=0.9998, for fringes presented in Fig. 9(j) and Fig. 9(l) QI=0.9935, for fringes presented in Fig. 9(n) and Fig. 9(p) QI=0.9923. All important uVID attributes, such as details preservation and ability to work with wide range of fringe pattern characteristics, are well preserved. Computation time is greatly reduced, thus the goal of this study is successfully achieved. It should also be highlighted that with the DeepVID approach the neural network estimate does not need to be highly precise unlike in the case where the neural network provides fully background and fringes estimation. In the case of complex fringe patterns and high iteration number the RMSE curve of uVID filtration results flattens and the estimates does not change significantly even for several dozens of iterations.

Table 1. Analysis of the experimentally recorded TPS series of interferograms with phase shift equal to $\pi/2$.

				
number of Chambolle projection iterations estimated with the use of tolerance stopping criterion (ground truth)				
583	591	572	572	581
number of Chambolle projection iterations estimated with the use of DeepVID neural network				
638	655	647	622	635
time needed for the calculations with the use of tolerance stopping criterion [s]				

66	42	40	40	41
time needed for the calculations with the use of DeepVID prediction [s]				
16	14	15	15	15
reduction of the calculation time				
>4 times	3 times	~3 times	~3 times	~3 times
RMSE of the filtration results estimated after the iteration number defined by tolerance stopping criterion				
0.1044	0.102	0.098	0.0907	0.1454
RMSE of the filtration results estimated after the iteration number defined by DeepVID				
0.1065	0.1053	0.1027	0.0935	0.1468

Table 2. Analysis of the experimentally recorded interferogram of healthy prostate cells (RWPE).

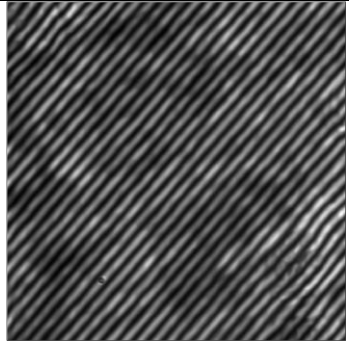
	number of Chambolle projection iterations estimated with the use of	
	tolerance stopping criterion	DeepVID prediction
	312	318
	time needed for the calculations with the use of	
	tolerance stopping criterion	DeepVID prediction
	18.2 s	7.6 s
	reduction of the calculation time: 2.4 times	

Table 3. Analysis of the interferogram encoding the droplet shape.

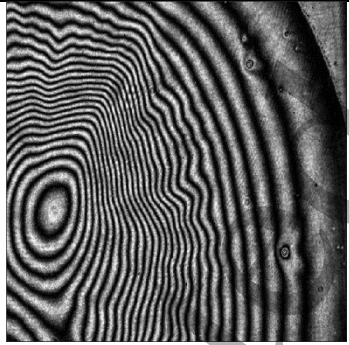

	number of Chambolle projection iterations estimated with the use of	
	tolerance stopping criterion	DeepVID prediction
	519	547
	time needed for the calculations with the use of	
	tolerance stopping criterion	DeepVID prediction
	43 s	10 s
	reduction of the calculation time: 4.3 times	

Table 4. Analysis of the besselogram encoding the amplitude of the higher vibration mode of square micromembrane.

	number of Chambolle projection iterations estimated with the use of	
	tolerance stopping criterion	DeepVID prediction
	1272	1203
	time needed for the calculations with the use of	
	tolerance stopping criterion	DeepVID prediction
	168 s	36 s

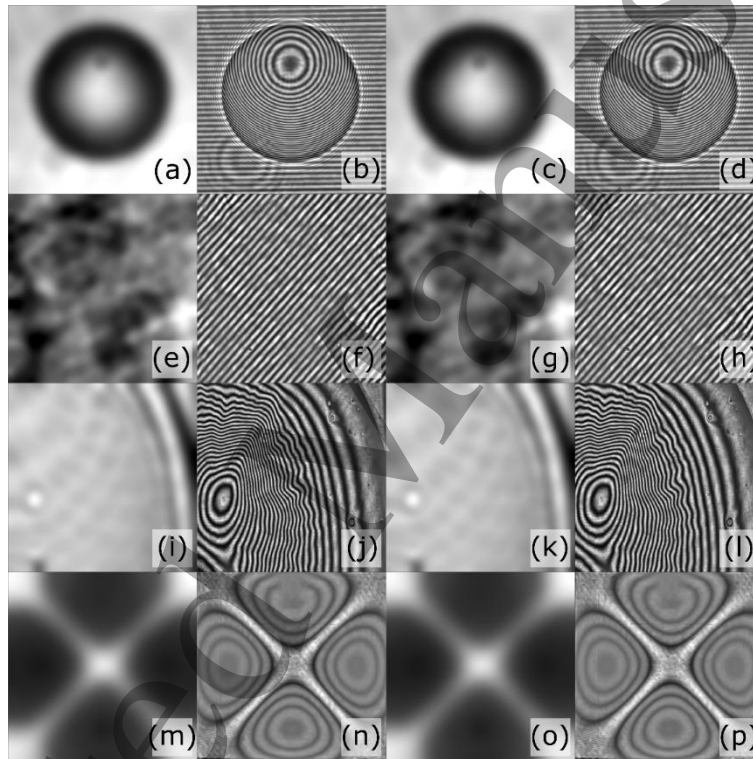
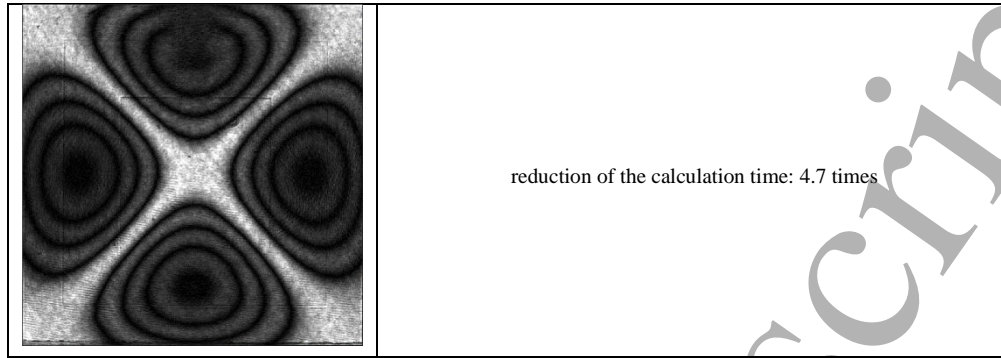


Fig. 9. The prefiltration results of fringe patterns presented in Tables 1-4. Microsphere interferogram background and fringes separation using (a), (b) DeepVID approach and (c), (d) uVID approach. RWPE interferogram background and fringes separation using (e), (f) DeepVID approach and (g), (h) uVID approach. Droplet interferogram background and fringes separation using (i), (j) DeepVID approach and (k), (l) uVID approach. Besselogram background and fringes separation using (m), (n) DeepVID approach and (o), (p) uVID approach.

5. Conclusion

The main novelty of this paper is connected with the meaningful acceleration of unsupervised VID algorithm [57] with the use of convolutional neural network. The biggest advantage of uVID over other known VID-based fringe pattern prefiltration solutions is its automation, versatility and robustness. It was achieved through the introduction of the tolerance stopping criterion allowing the automatic definition of proper number of algorithm's iterations. On the other hand, introduced stopping criterion significantly increased the computation time of a single iteration and by doing so, it extended the prefiltration time. Proposed DeepVID neural network has been taught to map the performance of the previously introduced [57] tolerance

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

stopping criterion. It is important to emphasize that the purpose of using neural network was not to increase the accuracy of already highly precise uVID fringe pattern prefiltration algorithm, but to accelerate the uVID calculations without jeopardizing its efficiency. DeepVID allows to instantly determine the proper Chambolle projection iteration [88] number omitting the time-consuming process of stopping criterion calculation. The performance of DeepVID solution was successfully tested using the experimentally recorded data. It was proven that the time of the VID prefiltration was shortened up to 4-fold.

Another important novelty of this paper is the analysis and comparison of three VID models: TV-L², TV-H¹ and TV-G in the context of a specific task of fringe pattern prefiltration. It is the first time when the issue is discussed considering the calculation complexity of different models and their performance for different types of fringe patterns (with high and low frequency components). Conclusions drawn from mentioned analysis allow to determine TV-G model as the most suitable for prefiltering fringe patterns, which confirmed the efficiency of uVID approach.

Acknowledgement

The authors would like to thank Dr Sam Dehaeck for the access to experimental data presented in Table 3 and Fig. 9.

Funding

This work has been partially funded by the National Science Center Poland (OPUS 2020/37/B/ST7/03629 and PRELUDIUM 2021/41/N/ST7/04057), Grant PID2020-120056GB-C21 funded by MCIN/AEI/10.13039/501100011033. Studies were funded by FOTECH-1 project granted by Warsaw University of Technology under the program Excellence Initiative: Research University (ID-UB). MC work was supported by the Foundation for Polish Science (FNP) and by the Polish National Agency for Academic Exchange under the Iwanowska programme.

Disclosures

The author declares no conflicts of interest.

Data Availability.

Data may be obtained from the authors upon reasonable request.

References

1. J. Schwider, "Advanced evaluation techniques in interferometry," in *Progress in Optics*, E. Wolf, ed., (North Holland, 1990).
2. D. Malacara, *Optical Shop Testing* (John Wiley, 2007).
3. D. W. Robinson and G. Reid, *Interferogram Analysis: Digital Fringe Pattern Measurement* (Institute of Physics Publishing, 1993).
4. D. Malacara, M. Servin, and Z. Malacara, *Interferogram Analysis for Optical Testing* (Marcel Dekker, 1998).
5. M. K. Kim, *Digital Holographic Microscopy: Principles, Techniques, and Applications* (Springer-Verlag, New York, 2011).
6. B. Kemper and G. von Bally, "Digital holographic microscopy for live cell applications and technical inspection," *Appl. Opt.* 47(4), A52-A61 (2008).
7. M. K. Kim, "Principles and techniques of digital holographic microscopy," *SPIE Rev.* 1(1), 018005 (2010).
8. M. Mikuła, T. Kozacki, M. Józwick and J. Kostencka, „Accurate shape measurement of focusing microstructures in Fourier digital holographic microscopy,” *Appl. Opt.* 57(1), A197-A204 (2018).
9. S. S. Gorthi and P. Rastogi, "Fringe projection techniques: Whither we are?" *Opt. Lasers Eng.* 48(2), 133–140 (2010).
10. K. Patorski and M. Kujawińska, *Handbook of the Moiré Fringe Technique* (Elsevier, 1993).
11. L. Sałbut, K. Patorski, M. Józwick, J. Kacperski, C. Gorecki, A. Jacobelli, and T. Dean, „Active micro-elements testing by interferometry using time-average and quasi-stroboscopic techniques,” *Proc. SPIE* 5145, 23-32 (2003).
12. R. Juarez-Salazar, C. Mendoza-Rodriguez, J. E. Hernandez-Beltran, and C. Robledo-Sanchez, "How do phase shifting algorithms work?" *Eur. J. Phys.* 39(6), 065302 (2018).

13. J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, "Digital wavefront measuring interferometer for testing optical surfaces and lenses," *Appl. Opt.* 13(11), 2693–2703 (1974).
14. P. Hariharan, B. F. Oreb and T. Eiju, "Digital phase-shifting interferometry: a simple error compensating phase calculation algorithm," *Appl. Opt.* 26(13), 2504–2505 (1987).
15. Y. Surrel, "Phase stepping: a new self-calibrating algorithm," *Appl. Opt.* 32(19), 3598–3599 (1993).
16. J. E. Greivenkamp, "Generalized data reduction for heterodyne interferometry," *Opt. Eng.* 23(4), 234350 (1984).
17. Z. Wang and B. Han, "Advanced iterative algorithm for phase extraction of randomly phase-shifted interferograms," *Opt. Lett.* 29(14), 1671–1673 (2004).
18. P. Carré, "Installation et utilisation du comparateur photoélectrique et interférentiel du Bureau International des Poids et Mesures," *Metrologia* 2(1), 13–23 (1966).
19. J. Vargas, J. A. Quiroga, and T. Belenguer, "Phase-shifting interferometry based on principal component analysis," *Opt. Lett.* 36(8), 1326–1328 (2011).
20. R. Smythe and R. Moore, "Instantaneous Phase Measuring Interferometry," *Opt. Eng.* 23(4), 361–364 (1984).
21. J. E. Millerd, N. J. Brock, J. B. Hayes, M. B. North-Morris, M. Novak and J. C. Wyant, "Pixelated phase-mask dynamic interferometer," *Proc. SPIE 5531, Interferometry XII: Techniques and Analysis* (2004).
22. M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," *J. Opt. Soc. Am.* 72(1), 156–160 (1982).
23. E. N. Leith, and J. Upatnieks, "Reconstructed Wavefronts and Communication Theory," *J. Opt. Soc. Am.* 52(10), 1123–1130 (1962).
24. E. Cucho, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," *Appl. Opt.* 39(23), 4070–4075 (2000).
25. Y. Baek, K. Lee, S. Shin, and Y. Park, "Kramers–Kronig holographic imaging for high-space-bandwidth product," *Optica* 6(1), 45–51 (2019).
26. N. T. Shaked, Y. Zhu, M. T. Rinehart, and A. Wax, "Two-step-only phase-shifting interferometry with optimized detector bandwidth for microscopy of live cells," *Opt. Express* 17(18), 15585–15591 (2009).
27. L. Xue, J. Lai, S. Wang, and Z. Li, "Single-shot slightly-off-axis interferometry based Hilbert phase microscopy of red blood cells," *Biomed. Opt. Express* 2(4), 987–995 (2011).
28. Q. Kemao, "Windowed Fourier transform for fringe pattern analysis," *Appl. Opt.* 43(13), 2695–2702 (2004).
29. K. Pokorski, and K. Patorski, "Processing and phase analysis of fringe patterns with contrast reversals," *Opt. Express* 21(19), 22596–609 (2013).
30. M. Pirga, and M. Kujawińska, "Two directional spatial carrier phase-shifting method for analysis of crossed and closed fringe patterns" *Opt. Eng.* 34(8), 2459–66 (1995).
31. L. Kai, and Q. Kemao, "Improved generalized regularized phase tracker for demodulation of a single fringe pattern," *Opt. Express* 21(20), 24385–97 (2013).
32. D. Gabor, and W. P. Goss, "Interference Microscope with Total Wavefront Reconstruction," *J. Opt. Soc. Am.* 56(7), 849–858 (1966).
33. I. Yamaguchi, and T. Zhang, "Phase-shifting digital holography," *Opt. Lett.* 22(16), 1268–1270 (1997).
34. K. G. Larkin, D. J. Bone, and M. A. Oldfield, "Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform," *J. Opt. Soc. Am. A* 18(8), 1862–1870 (2001).
35. M. Wielgus and K. Patorski, "Evaluation of amplitude encoded fringe patterns using the bidimensional empirical mode decomposition and the 2D Hilbert transform generalizations," *Appl. Opt.* 50(28), 5513–5523 (2011).
36. M. Trusiak, Ł. Służewski, and K. Patorski, "Single shot fringe pattern phase demodulation using Hilbert-Huang transform aided by the principal component analysis," *Opt. Express* 24(4), 4221–4238 (2016).
37. M. Trusiak, K. Patorski, and K. Pokorski, "Hilbert-Huang processing for single-exposure two-dimensional grating interferometry," *Opt. Express* 21(23), 28359–28379 (2013).
38. M. Cywińska, M. Trusiak, V. Mico and K. Patorski, "Single-frame fringe pattern analysis using modified variational image decomposition aided by the Hilbert transform for fast full-field quantitative phase imaging," *Proc. SPIE 10677, Unconventional Optical Imaging, 106772B* (2018).
39. M. Trusiak, M. Cywińska, V. Mico, J.-A. Picazo-Bueno, C. Zuo, P. Zdankowski and K. Patorski, "Variational Hilbert Quantitative Phase Imaging," *Sci Rep* 10, 13955 (2020)
40. M. Cywińska, M. Trusiak, K. Patorski, "Analysis of fringe patterns with variable density using modified variational image decomposition aided by the Hilbert Transform," *Proc. SPIE 10976, 21st Czech-Polish-Slovak Optical Conference on Wave and Quantum Aspects of Contemporary Optics, 1097609* (2018).
41. S. M. A. Bhuiyan, R. R. Adhami, and J. F. Khan, "Fast and adaptive bidimensional empirical mode decomposition using order-statistics filter based envelope estimation," *EURASIP J. Adv. Signal Process.* 2008(164), 725356 (2008).
42. M. Trusiak, M. Wielgus, and K. Patorski, "Advanced processing of optical fringe patterns by automated selective reconstruction and enhanced fast empirical mode decomposition," *Opt. Lasers Eng.* 52(1), 230–240 (2014).
43. M. Trusiak, K. Patorski, and M. Wielgus, "Adaptive enhancement of optical fringe patterns by selective reconstruction using FABEMD algorithm and Hilbert spiral transform," *Opt. Express* 20(21), 23463–23479 (2012).

44. X. Zhou, A. G. Podoleanu, Z. Yang, T. Yang, and H. Zhao, "Morphological operation-based bi-dimensional empirical mode decomposition for automatic background removal of fringe patterns," *Opt. Express* 20(22), 24247–24262 (2012).
45. C. Wang, Q. Kema, and F. Da, "Automatic fringe enhancement with novel bidimensional sinusoids-assisted empirical mode decomposition," *Opt. Express* 25(20), 24299–24311 (2017).
46. Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.* 1(1), 1–41 (2009).
47. P. Gocłowski, M. Trusiak, A. Ahmad, A. Styk, V. Mico, B. S. Ahluwalia and Krzysztof Patroski, "Automatic fringe pattern enhancement using truly adaptive period-guided bidimensional empirical mode decomposition," *Opt. Express* 28(5), 6277–6293 (2020).
48. P. Gocłowski, M. Cywińska, A. Ahmad, B. Ahluwalia and Maciej Trusiak, "Single-shot fringe pattern phase retrieval using improved period-guided bidimensional empirical mode decomposition and Hilbert transform," *Opt. Express* 29(20), 31632–31649 (2021).
49. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* 60(1–4), 259–268 (1992).
50. L. Vese and S. Osher, "Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing," *J. Sci. Comput.* 19(1–3), 553–572 (2003).
51. Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations* (American Mathematical Society, 2001).
52. A. Aujol and A. Chambolle, "Dual norms and image decomposition models," *Int. J. Comput. Vis.* 63(1), 85–104 (2005).
53. S. Osher, A. Sole and L. Vese, "Image decomposition and restoration using total variation minimization and the H^{-1} norm," *Multiscale Model. Simul.*, 1(3), 349–370 (2003).
54. X. Zhu, Z. Chen, and C. Tang, "Variational image decomposition for automatic background and noise removal of fringe patterns," *Opt. Lett.* 38(3), 275–277 (2013).
55. X. Zhu, C. Tang, B. Li, C. Sun, and L. Wang, "Phase retrieval from single frame projection fringe pattern with variational image decomposition," *Opt. Lasers Eng.* 59(8), 25–33 (2014).
56. B. Li, C. Tang, G. Gao, M. Chen, S. Tang, and Z. Lei, "General filtering method for electronic speckle pattern interferometry fringe images with various densities based on variational image decomposition," *Appl. Opt.* 56(16), 4843–4853 (2017).
57. M. Cywińska, M. Trusiak, K. Patroski, "Automatized fringe pattern preprocessing using unsupervised variational image decomposition," *Opt. Express* 27(16), 22542–22562 (2019).
58. M. Cywińska, M. Trusiak, K. Patroski, "Biological phase sample study using variational Hilbert imaging technique," *Proc. SPIE 10887, Quantitative Phase Imaging V*, 108872F (2019).
59. M. Cywińska, M. Trusiak, K. Patroski, J. A. Picazo-Bueno, V. Mico, "Modified variational image decomposition algorithm aided by the Hilbert transform as an alternative to 2D Hilbert-Huang transform for fringe pattern phase retrieval," *Proc. SPIE 10834, Speckle 2018: VII International Conference on Speckle Metrology*, 1083422 (2018).
60. M. Cywińska, M. Trusiak, A. Styk, K. Patroski, "Full-field vibration profilometry using time-averaged interference microscopy aided by variational analysis," *Optics Express* 28(1), 435–450 (2020).
61. S. Cho, "A Neural Network for Denoising Fringe Patterns with Nonuniformly Illuminating Background Noise," *J. Korean Phys. Soc.* 75(6), 454–459 (2019).
62. B. Lin, S. Fu, C. Zhang, F. Wang and Y. Li, "Optical fringe patterns filtering based on multi-stage convolution neural network," *Opt. Lasers Eng.* 126, 105853 (2020).
63. K. Yan, J. Shi, T. Sun, J. Li and Y. Yu, "Fringe pattern filtering using convolutional neural network," *Proc. SPIE 11205, Seventh International Conference on Optical and Photonic Engineering (icOPEN 2019)*, 112050O (2019).
64. K. Yan, Y. Yu, C. Huang, L. Sui, K. Qian and A. Asundi, "Fringe pattern denoising based on deep learning," *Opt. Commun.* 437, 148–152 (2019).
65. Z. Zhang, Y. Zheng, T. Xu, A. Upadhy, Y. J. Lim, A. Mathews, L. Xie, and W. M. Lee, "Holo-UNet: hologram-to-hologram neural network restoration for high fidelity low light quantitative phase imaging of live cells," *Biomed. Opt. Express* 11(10), 5478–5487 (2020).
66. J. Shi, X. Zhu, H. Wang, L. Song, L. and Q. Guo, "Label enhanced and patch based deep learning for phase retrieval from single frame fringe pattern in fringe projection 3D measurement," *Opt. Express* 27(20), 28929–28943 (2019).
67. A. Reyes-Figueroa, V. H. Flores, and M. Rivera, "Deep neural network for fringe pattern filtering and normalization," *Appl. Opt.* 60(7), 2022–2036 (2021).
68. W. Xiao, Q. Wang, F. Pan, R. Cao, X. Wu, and L. Sun, "Adaptive frequency filtering based on convolutional neural networks in off-axis digital holographic microscopy," *Biomed. Opt. Express* 10(4), 1613–1626 (2019).
69. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin and C. Zuo, "Fringe pattern analysis using deep learning," *Adv. Photon.* 1(2), 025001 (2019).
70. S. Van der Jeught and J. J. J. Dirckx, "Deep neural networks for single shot structured light profilometry," *Opt. Express* 27(12), 17091–17101 (2019).
71. J. Qian, S. Feng, Y. Li, T. Tao, J. Han, Q. Chen, and C. Zuo, "Single-shot absolute 3D shape measurement with deep-learning-based color fringe projection profilometry," *Opt. Lett.* 45(7), 1842–1845 (2020).

72. H. Yu, X. Chen, Z. Zhang, C. Zuo, Y. Zhang, D. Zheng, and J. Han, "Dynamic 3-D measurement based on fringe-to-fringe transformation using deep learning," *Opt. Express* 28(7), 9405-9418 (2020).
73. H. Nguyen, N. Dunne, H. Li, Y. Wang, and Z. Wang, "Real-time 3D shape measurement using 3LCD projection and deep machine learning," *Appl. Opt.* 58(26), 7100-7109 (2019).
74. S. Feng, C. Zuo, L. Zhang, W. Yin, and Q. Chen, "Generalized framework for non-sinusoidal fringe analysis using deep learning," *Photonics Res.* 9(6), 1084-1098 (2021).
75. S. Feng, C. Zuo, Y. Hu, Y. Li, and Q. Chen, "Deep-learning-based fringe-pattern analysis with uncertainty estimation," *Optica* 8(12), 1507-1510 (2021).
76. K. Wang, Y. Li, Q. Kemao, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," *Opt. Express* 27(10), 15100-15115 (2019).
77. G. E. Spoorthi, S. Gorthi, R. K. Sai and S. Gorthi, "PhaseNet: A Deep Convolutional Neural Network for Two-Dimensional Phase Unwrapping," *IEEE Signal Processing Letters* 26(1), 54-58 (2019).
78. T. Zhang, S. Jiang, Z. Zhao, K. Dixit, X. Zhou, J. Hou, Y. Zhang, and C. Yan, "Rapid and robust two-dimensional phase unwrapping via deep learning," *Opt. Express* 27(16), 23173-23185 (2019).
79. J. Zhang, X. Tian, J. Shao, H. Luo, and R. Liang, "Phase unwrapping in optical metrology via denoised and convolutional segmentation networks," *Opt. Express* 27(10), 14903-14912 (2019).
80. G. Dardikman-Yoffe, D. Roitshtain, S. K. Mirsky, N. A. Turko, M. Habaza, and N. T. Shaked, "PhUn-Net: ready-to-use neural network for unwrapping quantitative phase images of biological cells," *Biomed. Opt. Express* 11(2), 1107-1121 (2020).
81. C. Wu, Z. Qiao, N. Zhang, X. Li, J. Fan, H. Song, D. Ai, J. Yang, and Y. Huang, "Phase unwrapping based on a residual en-decoder network for phase images in Fourier domain Doppler optical coherence tomography," *Biomed. Opt. Express* 11(4), 1760-1771 (2020).
82. Y. Qin, S. Wan, Y. Wan, J. Weng, W. Liu, and Q. Gong, "Direct and accurate phase unwrapping with deep neural network," *Appl. Opt.* 59(24), 7258-7267 (2020).
83. M. Cywińska, F. Brzeski, W. Krajnik, K. Patorski, C. Zuo, and M. Trusiak, "DeepDensity: Convolutional neural network based estimation of local fringe pattern density," *Opt. Lasers Eng.* 145, 106675 (2021).
84. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarjan, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* 16(8), 2080-2095 (2007).
85. M. Lebrun, "An Analysis and Implementation of the BM3D Image Denoising Method," *Image Processing Online* 2, 175-213 (2012).
86. M. Cywińska, M. Trusiak, V. Mico and K. Patorski, "Single-frame fringe pattern analysis using modified variational image decomposition aided by the Hilbert transform for fast full-field quantitative phase imaging," *Proc. SPIE* 10677, Unconventional Optical Imaging, 106772B (2018).
87. E. A. Evans and R. Skalak, *Mechanics and Thermodynamics of Biomembranes*, (CRC Press, 1980).
88. A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," *J. Math. Imaging Vis.* 20(1), 89-97 (2004).
89. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research* 15(56), 1929-1958 (2014).
90. V. Mico, C. Ferreira, Z. Zalevsky and J. Garcia, "Spatially-multiplexed interferometric microscopy (SMIM): converting a standard microscope into a holographic one," *Opt Express* 22(12), 14929-43 (2014).
91. Y. Tsoumpas, S. Dehaeck, A. Rednikov, and P. Colinet, "Effect of Marangoni Flows on the Shape of Thin Sessile Droplets Evaporating into Air," *Langmuir* 31(49), 13334-13340 (2015).
92. S. Dehaeck, Y. Tsoumpas, and P. Colinet, "Analyzing closed-fringe images using two-dimensional Fan wavelets," *Appl. Opt.* 54(10), 2939-2952 (2015).
93. S. Dehaeck and P. Colinet, "Improving speed and precision of local frequency analysis using Gaussian ridge interpolation for wavelet and windowed Fourier ridge algorithms," *Opt. Lasers Eng.* 77, 54-63 (2016).
94. Z. Wang and A. Bovik, "A Universal Image Quality Index," *IEEE Signal Process. Lett.* 9(3), 81-84 (2002).