Augmentation assisted robust fringe detection on unseen experimental signals applied to optical feedback interferometry using a deep network

Sumair Saeed Khurshid $^{1,1,1}$, Wajahat Hussain $^2$, Usman Zabit $^2$, and Olivier Bernal $^2$

$^1$Department of Electrical Engineering
$^2$Affiliation not available

October 31, 2023

Abstract

In this work, we propose using deep neural 2D networks, which have renowned generalization performance on unseen data. It specifically explains the use of Novel Augmentation technique for SMI Fringe detection and classification. Overall, it highlights the importance of data augmentation in computer vision algorithms. The results prove that computer vision can be used to cater the problems in signal processing and photoelectronics domain.
Augmentation assisted robust fringe detection on unseen experimental signals applied to optical feedback interferometry using a deep network

Sumair Saeed Khurshid¹, Wajahat Hussain¹, Usman Zabit¹ and Olivier Bernu²,³

¹Department of Electrical Engineering, National University of Sciences and Technology, Islamabad, Pakistan
²LAAS-CNRS, Toulouse, FRANCE
³University of Toulouse, Toulouse INP, Toulouse, FRANCE

corresponding author: usman.zabit@seecs.nust.edu.pk
doi:10.1109/TIM.2023.3251409

Abstract—Commercialization and industrial deployment of optical feedback interferometry or self-mixing interferometry (SMI) based displacement instruments is held back due to inaccurate fringe detection under different optical feedback or speckle induced by operating conditions. In this work, we propose using deep neural 2D networks, which have renowned generalization performance on unseen data. Nonetheless, training deep neural networks requires very large data which in our application would imply acquiring large datasets of experimental signals by operating such interferometers under as many operating conditions as possible. To circumvent this time- and resource-consuming process, we propose a novel data augmentation scheme which increases the amount of training data needed by a deep network for fringe detection/classification in SMI signals. Interestingly, this has enabled the trained deep network to acquire excellent generalization capability where it has learnt to detect SMI fringes belonging to weak-, and very strong-optical feedback regime, even when it was only trained on moderate-, and strong-feedback regime signals. Consequently, our trained model has shown robust performance for simulated weak-, moderate-, and strong-optical feedback regime SMI signals affected by additive noise. Various experimental SMI signals, acquired under different sensor- and optical-conditions, have also been successfully processed. We also implemented an established fringe detection method for comparison. Our work presents very good generalization capability as compared to this established method. Our novel augmentation scheme is generic and can be applied for other interferometric signals. We have released our dataset and implementation, with the hope that this will assist community in accelerating the commercialization of optical feedback interferometry leveraging the full potential of deep learning.

Index Terms—Fringe Detection; Fringe Classification; Data Augmentation; Self-Mixing interferometry; Optical Feedback Interferometry.

I. INTRODUCTION

SELF-MIXING interferometry (SMI) or optical feedback interferometry [1] has been demonstrated for metric applications such as displacement [2], flow [3], vibration [4], distance [5], velocity [6], range-finding [7] measurements etc. Focusing on displacement measurement with large dynamic range, detection of interferometric fringes becomes a challenge especially when the remote target surface is rough thereby causing variable optical feedback (VOF) [8] and inducing speckle [9] which result in variations in SMI signal characteristics [9]. This task is even more complicated if the SMI signal has low signal to noise ratio (SNR) [10], especially under industrial settings [2], [11]. Consequently, various methods have been proposed to ensure accurate fringe detection under these conditions, as described below.

Under moderate optical feedback (specified by $1 < C < 4.6$, where C is the optical feedback coupling factor [1]), simple derivative based thresholding of SMI signal is typically used to detect SMI fringes [12] under high SNR while double-derivative based method was proposed in [13] to counter increased noise in such SMI signals. However, these methods tend to become ineffective if C varies as this causes variations in amplitude as well as shape of SMI signals [9]. Thus, adaptive thresholding [9], Hilbert transform [14], auto-correlation [15], and custom wavelet transform [16] based processing has been proposed for VOF based SMI signals. These methods provide improved performance as compared to [12], [13] but become ineffective for $C < 0.5$ [9], [16] or might suffer from loss of direction information [14].

Artificial neural networks (ANNs) are being increasingly used for instrumentation and measurement applications, such as localization of welding beads [17], bearing fault detection in a rocket engine [13], tyre- [19] and milling tool- [20] defect detection etc. Focusing on SMI, such ANNs have also been previously employed to process SMI signals, such as for noise removal [21], [22], fringe classification [23], modality detection [24] and displacement measurement [25]. The networks used in these approaches are either 1D (one-dimensional) networks which hinders their use for 2D spatial analysis, or shallow networks having a small number of hidden layers which limits their generalization ability to handle unseen variations in the data.

On the other hand, in this work, we propose using 2D deep convolutional neural networks (DNNs) to detect and classify SMI fringes even when a large variation in optical feedback as well as SNR occurs in SMI signals. Our aim is to utilize the spatial analysis and the generalization capability [26] of 2D

This work was supported in part by ANR-20-CE42-0010 PICSONDE
DNNs to enable robust 1D fringe detection and classification. By leveraging the outstanding generalization ability of 2D DNNs, we have been able to handle unforeseen speckle- and additive-noise corrupting real-world SMI signals.

2D DNNs are at the heart of the deep learning revolution starting from AlexNet [26]. Since then 2D DNNs have shown to match human level performance on multiple tasks or even surpass expert performance on challenging tasks, e.g., auto-learning to play video games [27], playing games against the world champion [28] and pattern detection domains requiring expert vision (medical imaging) [29] etc. Furthermore, the generalization ability of 2D DNNs is well established. Deep features, extracted from DNNs, outperformed handcrafted features on multiple visual tasks [30]. Additionally, using 2D DNNs, trained on very large but arbitrary datasets, as backbones, has achieved state of the art performance [31].

Training very deep networks suffers from overfitting issues. Data augmentation is a well known approach to increase the data in order to avoid overfitting [26]. Standard image augmentation methods include random shifting, horizontal/vertical flipping and intensity variations [26]. Recently, a simple image augmentation method managed to achieve state of the art performance on reinforcement learning tasks [32]. Deep reinforcement learning tasks generally require a test-bed or a simulator to deploy different strategies [27], [33]. The test-bed/simulator allows gathering large amounts of data. However, even when the data is abundant, augmentation scheme helps [32] to overcome the data blind-spots [34]. As mentioned earlier, deep reinforcement learning tasks have outperformed human experts [27], [28]. Even they are benefitting from simple data augmentation schemes. In this work, we propose a novel image augmentation method to train very deep networks to detect and classify laser feedback based interferometric fringes which leads to show the ability of model to generalize as depicted in Fig. 1.

Generally isotropic image scaling has been utilized for image augmentation [26]. We propose anisotropic image scaling to increase the diversity of the training data. As it will be shown later in the Fig. 8 our deep detector shows robust fringe detection on completely unseen experimental signals.

We have leveraged the well known YOLOv5s (You Only Look Once Version 5 small) object detector [35], which detects multiple classes and multiple instances of a given class, in a single forward pass. It can process 140 frames/s (fps) and above [36], hence it can be used for real-time SMI processing.

We trained our model on a dataset in which C factor was varied from 1 to 10 with step-size of 0.5. The dataset included SMI signals in the absence of noise as well as in the presence of noise (SNR was varied from 13.5 dB to 41 dB with step-size of ~1 dB). The same remote target motion (composed of a multi-tone signal with frequencies between 3 Hz and 17 Hz with maximum amplitude of ~5 µm) was used to generate the dataset, sampled with a fixed sampling frequency of 11,000 samples/s, and α. Later, we tested our novel data augmentation method on a rich unseen dataset consisting of simulated SMI signals having different C and lower SNR values. Furthermore, multiple experimental SMI signals were acquired under different sensor- or optical- (with or without speckle) conditions by using two different laser diode based sensors for the sake of experimental validation. Different experimental SMI signals were acquired with experimental target vibration varying from 5 Hz to 100 Hz and amplitude varying from ~1 µm to ~25 µm). Typical sampling frequency of experimental SMI signals was of the order of 200,000 samples/s. During processing, a signal containing 100,000 data points is in general converted into 140 images, where each image contains plot of ~714 samples. All the 140 images are processed in 1 second using the deep network thereby ensuring real-time SMI processing. Also, to avoid truncation of a fringe at the end of an image, an overlap can be introduced between the consecutive images.
Currently, to the best of our knowledge, there is no benchmark dataset for evaluating SMI algorithms. We have released our labelled dataset as well as the image augmentation implementation\(^1\) to further enhance the research in this area.

The rest of the paper is organized as follows. After a brief overview of neural networks applied to SMI signals in section II, our proposed methodology to design and train a 2D ANN to perform fringe detection and classification is presented in section III. Then, in section IV, the obtained results on simulated- and experimental-SMI signals are presented and compared to a classical SMI fringe detection. Then, 2D detector optimization is presented in section V, followed by Discussion and Conclusion.

II. RELATED WORK

We are not the first ones to deploy ANNs for processing SMI signals. This area of research is rapidly evolving. Here we discuss the related work closest to our topic (see Table I).

\(^{[21]}\) used an ANN with a single hidden layer to remove noise from the SMI signal. More recently, \(^{[22]}\) used generative adversarial network (GAN), where encoder-decoder both comprised 22 hidden layers each, to remove different types of noise from the SMI signal. Both the aforementioned approaches used shallow networks and treated the SMI signal as 1D input. We propose to process the SMI signal as 2D input (image) by a DNN. This enables our approach to utilize the generalization ability (scale invariance) of these rich architectures and detect the fringes in a single forward pass, instead of simply focusing on noise removal.

In addition to noise removal, ANNs have also been used for laser multi-modality (uni, bi and tri) classification \(^{[24]}\).

\(^1\)https://github.com/skhurshid-data/Interferometric-Data-Augmentation

### TABLE I

<table>
<thead>
<tr>
<th>Work</th>
<th>Architecture</th>
<th>Data Augmentation</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1D 1</td>
<td>×</td>
<td>Noise Removal</td>
</tr>
<tr>
<td></td>
<td>1D 45</td>
<td>Gaussian Noise, Amplitude Noise</td>
<td>Noise Removal</td>
</tr>
<tr>
<td></td>
<td>1D 1</td>
<td>×</td>
<td>Fringe Classification</td>
</tr>
<tr>
<td></td>
<td>1D 1</td>
<td>×</td>
<td>Multi-modality Detection</td>
</tr>
<tr>
<td></td>
<td>1D 11</td>
<td>Gaussian Noise</td>
<td>Displacement Estimation</td>
</tr>
<tr>
<td></td>
<td>1D 1</td>
<td>×</td>
<td>Liquid Level Prediction</td>
</tr>
<tr>
<td></td>
<td>2D 18</td>
<td>×</td>
<td>2D Fringe Detection</td>
</tr>
<tr>
<td>Ours</td>
<td>2D 224</td>
<td>Anisotropic Scaling</td>
<td>1D Fringe Detection and Classification</td>
</tr>
</tbody>
</table>

as well as fringe (positive- and negative-fringe) classification \(^{[23]}\). These approaches also utilize shallow and 1D ANNs for classification. Furthermore, these classifiers rely on preprocessed SMI signal for classification. In opposition, our proposed approach detects multiple instances of positive- and negative-fringes efficiently in a single forward pass without requiring any preprocessing.

An interesting approach is proposed in \(^{[25]}\) to directly estimate the displacement information from a noisy SMI signal. This efficient approach manages to directly reconstruct the target displacement and more precisely its velocity. This approach relies on a 1D convolutional ANN which is trained using a set of experimental data which can be time consuming to acquire. Here, we propose instead a complete pipeline on how to apply famous 2D deep networks, with well known generalization ability, to 1D fringe detection, fringes that can later be used by algorithms to reconstruct the displacement \(^{[12]}\). In addition, the proposed 2D ANN is trained on a very large set of simulated data generated using our novel data augmentation.

Furthermore, \(^{[37]}\) have measured liquid level using optical interferometry. It leverages acoustic principles to map spectral harmonics to liquid level. Again they have used a shallow network (single hidden layer) with 1D signal as input. Generalization across several content materials, geometries and dimensions is left as future work. On the other hand, our work aims to increase generalization capabilities of a deep network by using novel data augmentation on SMI signals belonging to a very large C range.

Finally, there has been recent surge in the deep networks, e.g., GANs \(^{[40]}\), \(^{[41]}\) etc., that augment data themselves. These data augmentation schemes have improved the performance of deep networks in various domains \(^{[42]}\), \(^{[43]}\). However, there are few challenges while leveraging these generative deep networks. Firstly, these are deep networks themselves requiring training that depends on labelled data itself. Lack of diversity in the training data, leads to homogeneous samples generated by GANs \(^{[16]}\). Note that generating training data in the industrial settings (welding beads \(^{[17]}\), bearings status...
in rocket engine [18], tyre defects [19], milling tool wear and tear [20], although limited in amount, is still an expensive and a tedious effort. In this work, our novel augmentation scheme transforms simulated data into diverse larger dataset, which shows good performance even on experimental signals. Secondly, data generated by these networks has mainly been used for classification [20], [42]. The focus of this work is detection which requires specifying the location of the fringe in the image and its type. Merely more data is not enough and manual labelling is required to detect bounding boxes [17], [19]. On the other hand, our data augmentation scheme is parametric and therefore, labels (fringe locations) can also be predicted in the augmented data.

III. OUR METHODOLOGY

A. How to train very deep 2D networks for fringe detection

In this section we detail how to train a 2D DNN to detect 1D fringes. The details of our approach are provided in Algorithm 1. We divide every 1D SMI signal of our simulated dataset, into N windows, where each window \( x \in \mathbb{R}^{\ell \times 1} \). More formally 1D SMI signal is represented as \( \mathcal{X} = \{x^1, x^2, ..., x^\ell\} \).

To diversify data, we add Gaussian noise to augment data (Line 1 in Algorithm 1). We sample a noisy vector \( \eta \) of length \( \ell \) from a Gaussian distribution \( \mathcal{N}(\mu, \sigma^2) \). Here mean (\( \mu \)) is zero, and variance (\( \sigma^2 \)) is the user provided noise level. As later shown in Fig. 8 our approach is robust to very high levels of noise in the experiments. Each window \( x \) is transformed into a noisier version \( x' \) (Line 8 to 14). We merge both versions of SMI windows \( (x, x') \) to form \( \mathcal{X}' \) (Line 13).

Next we plot each window \( x \in \mathcal{X}' \), resulting into a 2D image \( I \), (Line 2). Each image has r rows and c columns. All the images form the unlabelled image set \( \mathcal{I} = \{I^1, I^2, ..., I^M\} \) (Line 2). We manually label each \( I \in \mathcal{I} \) (Line 3). Labelling process is shown in Fig. 2. Given a sample image \( I \), a human annotator has to draw a bounding box \( b = \{x_{min}, x_{max}, y_{min}, y_{max}\} \) around a fringe and provide its label (p: positive fringe, n: negative fringe). \( (x_{min}, y_{min}) \) and \( (x_{max}, y_{max}) \) represent the 2D coordinates of the top left corner and the bottom right corner of the bounding box (see Fig. 2). We used the LabelImg tool [44] for labelling our dataset.

![Fig. 2. Manual labelling of fringes. Human annotator draws a bounding box around each fringe and provides its label.](image)

We have a set of bounding boxes \( B = \{b_1, b_2, ..., b_\alpha, b_\beta\} \) for each \( I \in \mathcal{I} \), where \( \alpha \) and \( \beta \) is the number of positive and negative fringes, respectively. Labels (bounding boxes) of the entire image set \( \mathcal{I} \) are represented as \( B = \{B^1, B^2, ..., B^M\} \).

Generating labelled data is the most laborious and resource demanding task. Very deep 2D convolution networks require a large amount of training data. To this end, we augment the image set \( \mathcal{I} \) using our novel image augmentation method (Line 4). We automatically generate the labels \( B' \) for the augmented dataset \( \mathcal{I}' \) as explained in section III-B. We divide the labelled data \( (\mathcal{I}', B') \) into training \( (\mathcal{I}_t, B_t) \) and validation set \( (\mathcal{I}_v, B_v) \) (Line 5).

Given the training data, we train the YOLOv5s object detector \( \mathcal{F}_{YOLO} \) (Line 6). This 2D DNN maps an input image \( I \) into a set of detections (bounding boxes) \( \hat{B} \) as given in (1):

\[
\hat{B} = \mathcal{F}_{YOLO}(I) \tag{1}
\]

where \( \hat{B} \) represents the detected fringes (bounding boxes) and may differ from the ground truth labels \( B \) of input image \( I \). We estimate the accuracy a of our detector \( \mathcal{F}_{YOLO} \) using the validation data \( (\mathcal{I}_v, B_v) \) (Line 7).

Algorithm 1 How to train very deep network for fringe detection?

```plaintext
INPUT: \( \mathcal{X} \) (Sample Points), \( \mu \) (mean), \( \sigma^2 \) (variance)
1: \( \mathcal{X}' \leftarrow addNoise(\mathcal{X}, \mu, \sigma^2) \)
2: \( \mathcal{I} \leftarrow getName(\mathcal{X}, r, c) \) \{r: "rows", c: "columns"\}
3: \( B \leftarrow getLabel(\mathcal{I}, p, n) \) \{"p: positive class", n: "negative class"\}
4: \( \mathcal{I}', B' \leftarrow addAugmentation(\mathcal{I}, B, s) \) \{s: scaling factor\}
5: \( \mathcal{I}_t \leftarrow (\mathcal{I}', B') \) \( \Rightarrow (B_t, B_v) \)
6: \( \mathcal{F} \leftarrow trainDetector(\mathcal{I}_t, B_t) \)
7: \( a \leftarrow evalDetector(\mathcal{F}, \mathcal{I}_v, B_v) \)

Function addNoise(\( \mathcal{X}', \mu, \sigma^2 \))
8: \( \mathcal{X}' \leftarrow \mathcal{X} \)
9: for all i do
10: \( x'_{i \times 1} \leftarrow getSample(\mathcal{X}) \)
11: \( \eta_{i \times 1} \leftarrow getNoise(\mathcal{X}', \mu, \sigma^2) \)
12: \( x'_{i \times 1} \leftarrow x'_{i \times 1} + \eta_{i \times 1} \)
13: \( \mathcal{X}' \leftarrow (\mathcal{X}', x') \)
14: end for
return \( \mathcal{X}' \)
```

B. How to augment data for 1D fringe detection?

In this section we describe our novel anisotropic image augmentation. There are a few well-known image augmentation methods (crop, horizontal/vertical flips, rotate) commonly deployed [29], [32] to increase the training data. We propose anisotropic image scaling (Fig. 3).

This anisotropic scaling is represented by 3x3 homography \( H \) matrix [45]. A pixel \( p \) on a given image, is mapped to a pixel \( p' \), using transformation \( H \) given in (2):

\[
p'_{3 \times 1} = H_{3 \times 3} \times p_{3 \times 1}, \tag{2}
\]

where \( p \) is the homogenous representation [45] of a pixel with coordinates (x, y), i.e., \( p = [x, y, 1]^T \), \( p' = [a, b, c]^T \) is the pixel where \( p \) is mapped. Its coordinates are (a/c, b/c), which are calculated using \( p' = [a, b, c]^T \).

Our anisotropic scaling consists of both horizontal and vertical scaling (Fig. 3). Vertical scaling handles changes in the SMI signal whereas horizontal scaling handles target velocity changes embedded in the SMI signal. Detailed image augmentation process is given in Algorithm 2. Already
labelled data $(I, B)$ is scaled, horizontally and vertically, by an integer scale factor $s$.

For horizontal scaling, with scale factor $s$, we sample $s$ images from the labelled data $(I, B')$. The $k$-th sampled image $I^k_{(t/s)\times c}$, is transformed into horizontally scaled image $I^k_{(t/s)\times c}$, using homography $H^k$ (Line 8-9 in Algorithm 2). Similarly the corresponding labels (bounding boxes) $B_k$ of each sampled image, are transformed to $B^k$, using the same homography $H^k$. All the transformed images are horizontally stacked to form the augmented image, i.e., $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$ ... $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, Similarly, all the transformed labels $B^k$, are merged into single set $B_A$.

For vertical scaling, each sampled image $I_{t\times c}$, is transformed into vertically scaled image $I^k_{(t/s)\times c}$, using homography $H^k$ (Line 12-13 in Algorithm 2). The transformed images $I^k_{(t/s)\times c}$ are vertically scaled to form the augmented image, i.e., $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$ ... $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, $I^k_{(t/s)\times c}$, Similarly, all the transformed labels $B^k$, are merged into single set $B_A$.

Finally, the augmented image data $(I_A, B_A)$ is merged into manually labelled dataset $(I, B')$. The augmented image dataset is represented by $(I', B')$.

IV. RESULTS

A. Fringe detection performance in noisy conditions

We have trained our model for moderate-, and strong-optical feedback regime (by varying $C$ from 1 to 10) in the absence (SNR of $\infty$) as well as in the presence of noise (SNR was varied from 13.5 dB to 41 dB). The codes of SMI signals simulation are written and executed using Matlab while “Data Augmentation” codes are written in Python. The training is done using NVIDIA GPU (GEFORCE RTX 2080 Ti) with 16 GB RAM. The results are shown in Fig. 3 and Table II. Each row of Table II quantifies the performances for 2000 fringes taken from different simulated SMI signals having specified $C$ and SNR values.

When the noise levels match the training regime, we achieve ~99.3% on fringe detection and ~98.4% on fringe classification (first row in Fig. 3). For moderate-optical feedback regime SMI signals (C from 1 to 4), we increased the noise level (SNR from 13.5 dB to 5.6 dB) for the test data (second row in Fig. 3). Even in these nosier conditions, our approach shows a little degradation with ~98.9% in fringe detection while ~95.6% in fringe classification. We attribute this robustness to our novel augmentation scheme. For the strong-optical feedback regime SMI signals (C from 5 to 9), we also increased the noise...
Fig. 4. SMI fringe detection performance under different optical feedback and SNR conditions. Training regime (C from 1 to 10, SNR ≥ 13.5 dB). First row: test regime is the same as training regime. Second row: Test regime with increased noise (C from 1 to 4, SNR from 13.5 dB to 5.6 dB). Third row: Test regime with higher noise (C from 5 to 9, SNR from 13.5 dB to 8.2 dB). In this regime it is difficult even for a trained person to classify fringes.

Fig. 5. Cross-regime generalization across weak optical feedback regime. Training regime (1 ≤ C ≤ 10, SNR ≥ 13.5 dB). First row: test regime (0 < C ≤ 1, SNR ≥ 18.8 dB). Second row: test regime with increased noise level (SNR from 10.9 dB to 13.4 dB). Last row: test regime with higher noise (SNR from 10.9 dB to 4.82 dB). Even a human annotator may not correctly classify the detected fringes for the noise level in the last row.

### TABLE III

<table>
<thead>
<tr>
<th>C</th>
<th>SNR (dB)</th>
<th>C</th>
<th>SNR (dB)</th>
<th>Detection</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 10</td>
<td>≥ 13.5</td>
<td>0 to 1</td>
<td>≥ 18.8</td>
<td>1984/2000 (99.2%)</td>
<td>1908/2000 (95.4%)</td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>10.9 to 13.4</td>
<td>1921/2000 (96.05%)</td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>4.82 to 10.9</td>
<td>1916/2000 (95.8%)</td>
<td>1721/2000 (86.05%)</td>
<td></td>
</tr>
</tbody>
</table>

B. Cross-regime generalization of our approach

It is recalled that we have trained our model for moderate- and strong-optical feedback regime (C from 1 to 10) such that SNR ≥ 13.5 dB. We now test our model on completely different optical feedback regimes with increased noise levels.

1) Weak optical feedback regime (C from 0.1 to 1): The results are shown in Fig. 5 and Table III. Even for this completely non-overlapping regime, our approach achieves ~99.2% on fringe detection and ~95.4% on fringe classification (first row in Fig. 5). We test increasing levels of noise for this weak-optical feedback regime (Table III). Our approach shows graceful degradation in performance. Note that fringes tend to appear sinusoidal in this weak feedback regime (Fig. 5). An interesting case arises for the very high noise level (SNR from 10.9 dB to 4.82 dB). Our detector managed to detect fringes but classification decreased when it was close to 4.82 dB (~86%). On visual inspection (last row in Fig. 5) it is revealed that the cases which are close to C = 0.1 and SNR = 4.82 dB are also difficult for a trained person to classify.

2) Very strong feedback regime (C from 10 to 30): The results are shown in Fig. 6 and Table IV. In this regime, our approach surpasses the performance obtained in case of weak feedback regime. We achieve ~99.5% on fringe detection and ~99.4% on fringe classification. In this regime the fringes tend to disappear (first and second row in Fig. 6). However, our proposed approach shows high generalization capability, which we attribute to our novel image augmentation.

Here, our test data contains noise such that SNR ≥ 22 dB. It becomes difficult to spot fringes in the SMI signal if the noise level is further increased (see last row in Fig. 6). This higher SNR requirement under these high C value based SMI
signals is understandable as SMI fringe amplitude increasingly reduces as C is increased. This contrasts with weak regime SMI signals in which fringe amplitude does not decrease as C is reduced from 1 towards 0.

C. Experimental Setup

The experimental setup is depicted in Fig. 7. Two different SMI sensors were used to acquire SMI signals. One is based on the Hitachi HL7851G laser diode (LD) emitting laser wavelength of 785 nm. This laser is driven by a constant LD drive current. Other SMI sensor uses a L1550P5DFB Telcordia laser diode, with wavelength of 1550 nm, is shown. A variable optical attenuator is used to change C. Use of a rough remote target surface caused VOF and speckle. Consequently, its performance degrades in the absence of speckle. Consequently, its performance degrades for such experimental SMI signals.

We also compared the latency (last two columns in Table V) of our deep approach with that of the classic fringe detection method [12]. We converted 100000 data points into 140 images (frames) each consisting of ~714 data points. It took 0.007 s to process one image on our GPU, so the total processing time is 0.007 x 140 = 0.98 s. This means our approach can be considered real-time when using a GPU, given our sampling frequency is 100 kHz. On the other hand, approach of [12] when executed on a CPU (Intel Core i7, 8th Gen, 8GB RAM) is 3.26 times faster, however, its accuracy is ~25% less than that of our approach. It is also noted that YOLOv5s has constant processing time for each image, irrespective of the number of fringes or the fringe amplitude (Fig. 9).

Finally Table VI shows that our approach has high precision (precision is the fraction of correct detections among the total detections), recall (recall is the fraction of correct detections among total instances) and F1 score (weighted average of precision and recall) indicating limited number of false positives and false negatives.

V. OPTIMIZING 2D DETECTOR FOR BINARY IMAGES

YOLOv5s detector was designed considering natural images consisting of scenes and objects. In this work, we have shown that well known detector works even if we convert 1D signal into a plot or an image, consisting of white background and a curve of single color. Is it possible to further optimize
Fig. 8. Generalization across variety of experimental SMI signals. Our proposed approach shows robust performance on fringes with different optical feedback strength, remote target frequency and amplitude, additive-noise, and optical speckle conditions causing varying signal shape and SNR.

In this work, we have quantized weights of YOLOv5s network (Table VII). Memory footprint considerably reduces from 57 MB for float32, to 7 MB for int8, with only a slight reduction in performance. Furthermore, this quantized YOLOv5s version can benefit from embedded platforms (e.g., Nvidia Jetson Nano platform). In our opinion, this detector can be further optimized using binary networks [48].

TABLE VI

<table>
<thead>
<tr>
<th>Occupied Bits</th>
<th>Data Type</th>
<th>Model Footprint (MBs)</th>
<th>F1 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>float32</td>
<td>57</td>
<td>0.96</td>
</tr>
<tr>
<td>16</td>
<td>float16</td>
<td>14.3</td>
<td>0.96</td>
</tr>
<tr>
<td>8</td>
<td>int8</td>
<td>7.4</td>
<td>0.93</td>
</tr>
</tbody>
</table>

VI. DISCUSSION

Use of data augmentation may become redundant if automated labeling of simulated data is possible. However, in the present case, it may not be straightforward to automatically label the simulated SMI fringes especially in the presence of noise. Note that the labeling is to be done within the images which are obtained after converting the 1D SMI signals provided by the SMI mathematical model to images. So, automatically placing the appropriate bounding boxes within the image of an SMI signal whose fringes may have different duration and amplitude is non-trivial.
Finally, in our opinion, well known and mature pipelines (like YOLO) come with lot of community support and proven track record, and therefore, should be further explored and optimized. In order to assist the community, we share our data and implementation to further explore this direction.

VIII. ACKNOWLEDGMENT

The authors would like to thank Agence Nationale de la Recherche (ANR) : ANR-20-CE42-0010 PICSONDE in cooperation with ACOEM, Thierry MAZOYER.

REFERENCES


