

Model-based neural networks for thermographic image reconstruction

Gergő Galiger

*Department of Numerical Analysis
Eötvös Loránd University
Budapest, Hungary
galiger.gergo@inf.elte.hu*

Nazila Azadi

*Silicon Austria Labs
Linz, Austria
nazila.azadi@silicon-austria.com*

Bernhard Lehner

*Silicon Austria Labs
Linz, Austria
bernhard.lehner@silicon-austria.com*

Mario Huemer

*Institute of Signal Processing
Johannes Kepler University
Linz, Austria
mario.huemer@jku.at*

Péter Kovács

*Department of Numerical Analysis
Eötvös Loránd University
Budapest, Hungary
kovika@inf.elte.hu*

Abstract—Thermographic imaging is a specific non-destructive evaluation (NDE) approach in which the investigated object is exposed to an initial thermal excitation. The utilization of infrared cameras to detect the induced surface temperature change enables the deduction of the internal structure of the inspected material. This requires addressing a large-scale ill-posed linear inverse problem, involving regularization and forward modeling of the observed thermal diffusion process.

In this paper, we introduce a hybrid approach for thermographic imaging, which combines a deep unfolded Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) algorithm with a U-Net. Consequently, the resulting network architecture incorporates established numerical regularization heuristics, such as sparsity and smoothness. Additionally, to mitigate the data demand of DL approaches, we establish a simulated training database containing both point-like defects and line-like cracks. As a case study, we begin by considering applications in non-destructive material testing, starting with 2D problems and progressing towards thermal tomography in 3D.

Index Terms—thermal tomography, non-destructive testing, deep unfolding, FISTA-Net, U-Net

I. INTRODUCTION

Deep learning (DL) models have recently been successfully applied to a wide range of problems. Nevertheless, these data-driven methods encounter issues including heavy data dependence, inconsistent predictions on unseen examples, and the opacity of learning parameters. These limitations, which are mainly due to the model-agnostic approach used in deep learning, can be mitigated by designing neural networks inspired by mathematical models. The so-called model-based DL is an emerging field in signal processing that incorporates domain-specific knowledge into the network architecture, training process, and data [1]. Another advantage of this approach is that it reduces the demand for training data, especially in cases where real-world measurements are costly and scarce. Deep unfolding is a novel model-based DL framework that integrates domain-specific mathematical and numerical heuristics, such

as smoothness, and sparsity, into the network design. In inverse imaging problems, such as tomography, these conditions are essential to improve the reconstruction accuracy. Therefore, we utilize deep unfolding to design a DL approach for thermal tomography, which allows for more accurate defect detection in non-destructive material testing applications.

In thermal tomography, objects are inspected using flash lamps or lasers, which induce a temperature change that can be measured on the material surface. The resulting thermal pattern can be utilized to infer the internal structure of the specimen by reconstructing the heat distribution just after the excitation. Although 1D reconstruction methods have dominated this field for the past decades, they are not accurate enough when dealing with phenomena like anisotropic heat flow. To achieve a more precise reconstruction, it is necessary to consider multidimensional heat flow. Recently, Burgholzer et al. [2] proposed the so-called virtual wave concept (VWC) that accounts for lateral heat flow, leading to improved accuracy in the reconstruction process. This is a two-step reconstruction technique that first converts the surface temperature data into pressure signals and then applies ultrasonic reconstruction methods in its final step. The second stage of the VWC concept involves large-scale ill-posed inverse problems, which we addressed by deep learning in our former works [3], [4]. Specifically, we trained U-Net architectures to learn the reconstruction from pressure data provided by the numerical solvers we applied in the first step.

In this paper, we unfold the iterative reconstruction process of the first step by adopting a novel model-based DL method called FISTA-NET [5]. Additionally, we build up a new training dataset for non-destructive material testing purposes, where not just point-like defects but line-like cracks will be simulated. Then, we fuse a trained FISTA-NET with a U-Net and evaluate the reconstruction accuracy both on simulated data and real-world measurements.

979-8-3503-8788-9/24/\$31.00 © 2024 IEEE

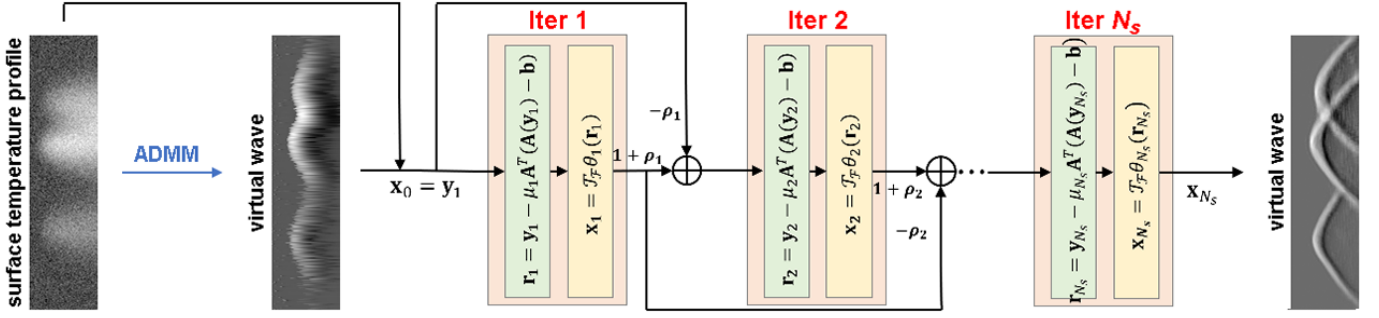


Fig. 1. Using the FISTA-Net deep unfolding network for enhancing ADMM virtual waves as an intermediary step of the two-step reconstruction process.

II. PROBLEM FORMULATION

In tomographic imaging, the object is penetrated by some form of radiation, which partially passes through the material while also being absorbed to some extent. The degree of absorption allows us to infer the internal structure of the object. Usually, the full 3D reconstruction can be reduced to a series of two-dimensional inverse problems that model the measurements taken from multiple directions and cross sections. The relationship between the measurements $\mathbf{b}$ and the local absorption coefficients $\mathbf{x}$ is often assumed to be of the form $\mathbf{Ax} = \mathbf{b}$, where the linear transformation $\mathbf{A}$ models the physical process. Since $\mathbf{A}$ is generally ill-conditioned, the solution vector $\mathbf{x}$ can only be obtained by incorporating domain knowledge into the reconstruction process. Numerical regularization offers a mathematically rigorous means to incorporate a priori information, such as sparsity or smoothness, about the solution $\mathbf{x}$:

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \{ \|\mathbf{b} - \mathbf{Ax}\|_2 + \lambda \cdot \Omega(\mathbf{x}) \}, \quad (1)$$

where infeasible solutions are penalized by the function Ω and the regularization parameter $\lambda \geq 0$.

The two stages of the thermal reconstruction problem can be modeled by two subsequent numerical regularizations described in Eq. (1). In stage one, $\mathbf{b}$ represents the surface temperature data, which is converted to so-called virtual (pressure) waves via the linear kernel matrix $\mathbf{A} = \mathbf{K}$ of the form:

$$\mathbf{K}_{i,j} = \frac{c}{\sqrt{\pi \alpha t_i}} \exp \left(-\frac{c^2 t_j^2}{4 \alpha t_i} \right) \quad (i, j = 0, \dots, N_t - 1), \quad (2)$$

where t_i denotes the measurement time instances, α stands for the thermal diffusivity, and $c > 0$ is the arbitrary chosen constant. In stage two, $\mathbf{b}$ represents the regularized solution obtained in stage one, whereas the matrix $\mathbf{A}$ implements the linear operator for wave back-projection (see e.g., [2], [4]). In the next section, we will focus on the regularization problem of stage one with $\Omega(x) = \|x\|_1$, and unfold the FISTA numerical solver into a model-based neural network.

III. METHODOLOGY

Chaining two ill-posed problems one after the other results in the magnification of the ill-posedness (i.e., obtaining a distorted result from the second step is common as its input is the

output of solving another ill-posed problem). As later shown in Section IV, this phenomenon is easily observable when working with noisy input data commonly found in real-world environments of industrial applications such as thermographic imaging. As a solution, we introduce an intermediary step in the reconstruction process, which consists in enhancing the virtual waves using the FISTA-Net model-based network.

A. FISTA-Net Adaptation and Optimization

To solve the ill-posed problem formulated in (1), the iterations of the FISTA numerical algorithm are unfolded in the FISTA-Net architecture into neural network layers by reformulating them as a series of gradient descent, proximal mapping, and momentum update steps. For the thermographic image reconstruction problem, the FISTA-Net architecture has been adapted by introducing the kernel matrix $\mathbf{K}$ from Eq. (2) into the gradient descent operation as:

$$\mathbf{r}^{(k)} = \mathbf{y}^{(k)} - (\mathbf{W}^{(k)})^T \left(\mathbf{K} \mathbf{y}^{(k)} - \mathbf{b} \right), \quad (3)$$

where $\mathbf{W}^{(k)}$ is the gradient operator, and $\mathbf{r}^{(k)}$, $\mathbf{y}^{(k)}$ the intermediate variables from the network mapping presented in [5]. Depending on the depth resolution (row dimension) of the $\mathbf{y}^{(k)}$ intermediary result, the matrix multiplication in (3) is limited to the first respective columns of the kernel matrix, an observation we used to reduce GPU memory allocation.

As initializing FISTA-Net using warm-starting has proven to be beneficial for ill-conditioned problems [5], we use an estimate of the target virtual waves as input of the network. This initial estimate is obtained using the ADMM numerical solver, as described in [3]. In addition, the regularization problem formulated in (1) is solved in the Abel space [6] to be able to account for the symmetry and sparsity of virtual waves in the FISTA-Net loss function. Furthermore, warm-starting enabled the simplification of the network in our experiments, resulting in smaller N_f filters in the proximal mapping layers of $N_f = 16$, and a lower iteration number of $N_s = 4$. A lazy data loading mechanism has also been implemented in the training process to enable the usage of large training datasets. The proposed pipeline is illustrated in Figure 1.

B. Training Dataset

Creating real-world data in thermography is both time-consuming and costly, as it requires the production of objects

with precisely known inner structures, and expensive equipment to perform accurate measurements. Therefore, we created a large dataset of only simulated data to train our adapted networks. For this, we used the data generator introduced in our previous work [3], which we expanded to improve the generalizability of our models. As defects in real-world objects are rarely perfectly square-shaped, we introduce a variety of defect shapes into the first binary target mask generation step, which we distinguish as point- and line-like defects.

The previously implemented square-shaped defects fall into the category of **point-like defects**, which are now generically generated as randomly parametrized polygons and ellipses. Polygons are defined by the number of edges $n \in [3, 6]$ and a scale multiplier $s \in [2, 6]$, while the parameters of the ellipses are their width and height $a, b \in [2, 6]$. Moreover, a random rotation value of $\theta \in [0, 2\pi)$ is also assigned to all polygons.

In addition to point-like defects, cracks are also commonly found in objects from real-world environments [7]. Therefore, we also include the generation of such **line-like defects** into the simulated data we train our networks on. In our implementation, cracks are modeled as a sequence of neighboring pixels characterized by their length in pixels $l \in [5, 20]$, and a direction $d \in [0, 7]$. The d direction parameter is used to limit the choice of neighboring pixels from 8 to 3 when generating a line, thus eliminating the generation of excessively curvy and looped cracks. Two representative samples from the simulated training data are shown in Figure 2.



Fig. 2. Data samples with different point-like defects and line-like cracks.

The further steps of generating the simulated training data are analogous to the methodology described in [3]. As the first step, a binary target mask is generated by randomly positioning $N \in [2, 5]$ point- and line-like defects. Then, the corresponding surface temperature data is calculated assuming adiabatic boundary conditions. Lastly, the simulated temperature measurements are used to compute the virtual waves by using ADMM and Abel transformation. The results of these steps are samples of three single-channel images of 256 by 64 pixels – target masks, temperature profiles, and virtual waves.

To keep the ability of trained models to generalize over different levels of noise in the input images, 10 different versions of each sample were generated from -20 dB to 70 dB SNR levels in 10 dB steps. In addition, this form of data augmentation also allows a more detailed analysis of the effects of input noise on model performance. Our final simulated dataset consisted of 10,000 different samples with 10 different versions for each, resulting in a total of 100,000 images. This dataset is further divided into training (80%), validation (10%), and testing sets (10%).

C. Implementation Details

The adapted FISTA-Net networks were trained on a single NVIDIA GeForce TITAN Xp GPU, which limited the training batch size to 10 samples. The learning rate was initially set to 0.1, which has been decreased with a learning rate decay of 0.9 over the first 25 training epochs, and every 5 epochs afterward. The rest of the hyperparameters and other training details were analogous to the ones published in [5]. The source code is based on the Python codebase published in this paper with the adapted networks implemented in PyTorch.

IV. EXPERIMENTS

In this section, we examine the results of the proposed method from multiple perspectives. First, model performance is analyzed in both steps of the two-step reconstruction process on the testing subset of the newly generated simulation data. Then, physical applicability is verified by evaluating the reconstruction process on real-world measurements. Lastly, we discuss the results of introducing different defect shapes in the synthetic training dataset.

A. Synthetic dataset

The ill-conditioned problem of thermographic reconstruction was solved in [3] using various model-based regularization techniques in the problem formulation expressed in (1). The first step of virtual wave approximation was solved with the ADMM numerical solver leveraging Curvelet and Abel transformations, of which the latter showed superior results. To provide a comprehensive analysis based on the simulated data, we examine the impact of employing the proposed model-driven method in both steps of the two-step reconstruction problem. For this, we perform a comparison analysis of our method with the previously used ADMM iterative algorithm using both Abel and Curvelet transformations. In addition, output samples are also analyzed later for a more qualitative examination of results.

For the analysis of the first reconstruction step, the previously unseen surface temperature data from the testing subset is used to approximate the corresponding virtual waves using all three approaches. Then, the per-pixel mean squared error (MSE) values are averaged between the reconstructed images and the simulated expected virtual waves from the dataset, which result in per-image MSE values. Lastly, we average these image scores grouping them by the SNR values of the input images. The results of this comparison are shown in Figure 3, which clearly demonstrates the advantage of using FISTA-Net in the first reconstruction step. The model-based network is able to provide a more exact reconstruction of the expected virtual waves than the previously used numerical solvers. Moreover, the performance differences are proportional to the amount of noise in the input images, an observation that hints at the usability of this method in noisy real-world environments. Note that using ADMM with both regularization approaches provides a way to validate the results of this comparison, as they align with the observations published in [3] regarding the advantages of using Abel transformation.

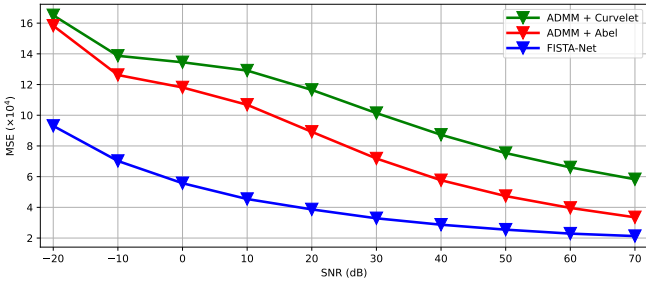


Fig. 3. Comparing model-based methods for virtual wave reconstruction in the first step.

The ability to provide a better approximation of virtual waves results in a significantly better reconstruction performance in the second reconstruction step, where these are used as input. As discussed earlier, this performance difference is mainly caused by the effects of chaining multiple ill-conditioned problems. The quantitative examination of the second step is shown in Figure 4 with the average per-image MSE differences of the reconstructed images and ground truths, grouped by SNR. The reconstruction here was performed using the plane wave ultrasound Stolt's f-k migration technique (FKMIG) [8], used as a baseline in [3]. This comparison shows a similar pattern to the first step in performance differences of the three approaches. Using FISTA-Net in the first step leads to a superior reconstruction performance, especially in the case of real-world-like noisy inputs.

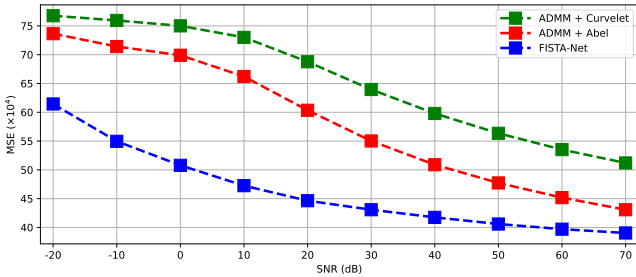


Fig. 4. Comparing model-based methods for initial temperature profile reconstruction using FKMIG in the second step.

For a qualitative demonstration of reconstruction performance differences, we compare output samples of FISTA-Net and ADMM (with Abel trf.) with ground truth images in both reconstruction steps. A representative example of this comparison is shown with two different samples in Figure 5, where both approaches are used for the reconstruction of virtual waves, and the pre-trained compact U-Net model from [3] is used in the second step.

The effects of using FISTA-Net in the virtual wave reconstruction step are illustrated on the left-side sample. Virtual waves originating from small and deep defects can not be reconstructed with ADMM at real-world-like SNR values of around 20 dB, while they are clearly separated when using FISTA-Net in the first step. At higher noise levels, even

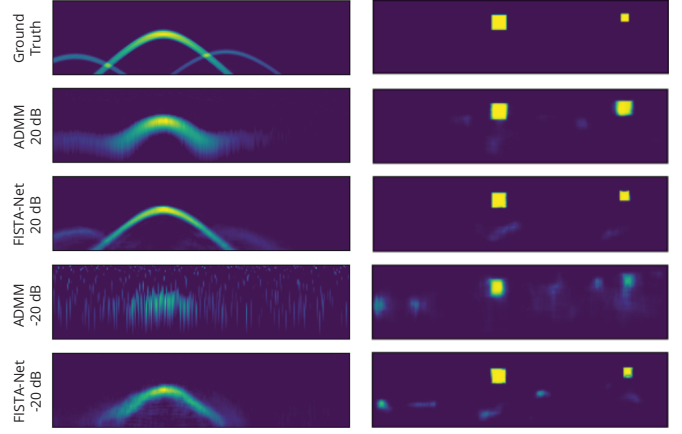


Fig. 5. Comparing different model-based approaches on two reconstruction examples of virtual waves (left) and initial temperature profiles (right) at different noise levels.

the virtual wave reconstruction of larger and close-to-surface defects, which are still visible when using FISTA-Net, is impossible with ADMM.

The effects of noise levels and defect characteristics on model usability also translate to the second reconstruction step. This is shown on the right-side sample from Figure 5, which contains two close-to-surface defects of different sizes. The ADMM reconstruction of the smaller defect is inaccurate at 20 dB SNR, with even the larger defect being hardly recognizable at higher noise. In contrast, the reconstruction precision of the FISTA-Net model is high in the case of close-to-surface defects, regardless of their shape. Furthermore, unwanted reconstruction artifacts are also less in number and smaller in size when using the model-based approach.

Although using the pre-trained U-Net model in the second reconstruction step provides a good baseline for testing the usage of FISTA-Net in the first step, further performance improvements can be obtained by retraining it on the new, more complex dataset. However, incorporating line-like defects significantly amplifies the complexity involved in the reconstruction process. This can be seen in Figure 6, where we compare reconstruction performances for point-like defects and line-like cracks separately. We observe that for line-like defects reconstructed with ADMM, the MSE increases with SNRs above 20 dB. Further analysis revealed that this trend is partly attributable to the characteristics of MSE. For instance, the sparse nature of cracks compared to larger, point-like defects diminishes their impact on the overall result. Considering MSE is utilized as the loss for the FISTA-Net reconstruction step, this can lead to the observed behavior, which may not always align with human intuition.

B. Real-world specimen

Synthetic data results from the two-step reconstruction process proposed in [3] translated well to a real-world problem, however, this cannot be simply assumed for the modifications presented here. Although extending the data generation process with multiple defect shapes results in a more general

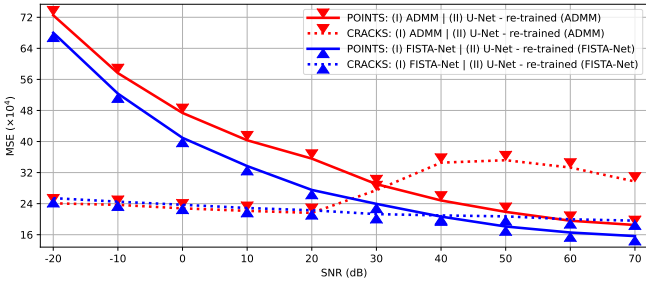


Fig. 6. Comparing model-based methods for initial temperature profile reconstruction using U-Net in the second step on point-like defects (POINTS) and line-like cracks (CRACKS).

dataset, it might also miss relevant aspects of real-world applications. In addition, the effects of the synthetically trained FISTA-Net model on data from practical applications are ambiguous as well.

To validate the practical feasibility of the proposed modifications, we evaluate the modified two-step process on a real-world specimen. For this, we use the measured data of two parallel steel rods embedded in a physical epoxy resin phantom from [3] to perform a qualitative comparison of our reconstructed output with the results published in this reference paper. The 3D reconstruction is done analogously to [3], by splitting the specimen into 2D cross sections (slices) about the vertical axis, performing the 2D reconstruction process on all cross sections, and finally merging them to obtain the 3D reconstruction. Note that the ability of this approach to accurately reconstruct 3D objects is limited, as it completely ignores heat flow between neighboring slices.

As a first experiment, we perform a comparison with the cross sections perpendicular to the orientation of the steel rods, where ignoring heat flow between slices least influences the reconstruction process. The results of this experiment are presented in Figure 7, which shows the same advantages of using the FISTA-Net model as the previous experiments on the synthetic dataset. More specifically, optimizing the virtual waves with this approach leads to a more exact reconstruction of deep-lying defects for this real-world application, as it results in uniform cross sections for both steel rods throughout the whole depth of the physical phantom. In addition, the proposed method also shows a better ability to separate these two objects inside the examined specimen, which can be observed on the deep-lying parts of the steel rod reconstructions.

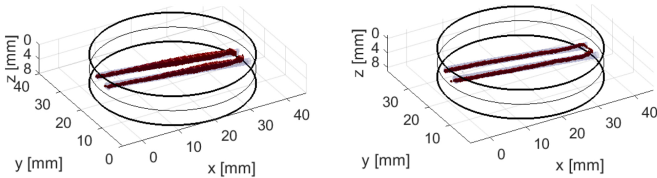


Fig. 7. Comparing the 3D reconstruction method introduced in [3] (left) with the proposed FISTA-Net virtual wave optimization approach (right) on the real-world specimen with no rotation.

In addition to the previous experiment, we also benchmark our method by perturbing the measurement data using a considerable amount of cross section rotation in Figure 8. Note that neither this process nor the original one proposed in [3] was designed for such 3D reconstruction, since the heat flow effects between the 2D cross sections significantly influence the thermographic reconstruction process in this case. Hence, the originally proposed method provided a considerably noisy reconstruction with a high number of artifacts and an inaccurate approximation of the steel rods' position. However, the proposed FISTA-Net virtual wave optimization step leads to an accurate result in the vicinity of the surface, while also eliminating unwanted reconstruction artifacts.

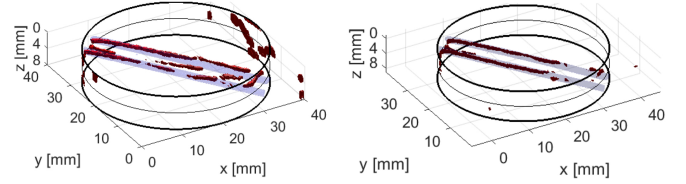


Fig. 8. Comparison from Fig. 7 with a 45° rotation of the real-world specimen.

V. CONCLUSIONS

Thermal NDE methods are essential in numerous industrial applications. However, the mathematical modeling of such approaches requires solving highly ill-posed linear inverse problems. As a solution, we present a regularized two-step hybrid approach combining the FISTA-Net and U-Net architectures, which we train on an extensive simulated dataset. To test the practical usability of the proposed method, we evaluate it on both synthetic and real-world data. Our approach improves the reconstruction accuracy of previous methods on the simulated data for a wide range of input noise levels. Specifically, the proposed hybrid process provides an enhanced reconstruction of small and deep defects, while also reducing unwanted artifacts. Moreover, our experiments demonstrate the same performance improvements in real-world applications.

For future work, there is a clear opportunity to enhance the evaluation procedure, as already pointed out during the discussion of Figure 6 in Section IV. Specifically, uncertainties remain regarding which metrics accurately reflect performance in alignment with human expert assessments. To address these concerns, we propose a thorough analysis of potential metrics, similar to what was done for segmentation in [9]. However, as thermographic imaging entails specific pitfalls and challenges distinct from segmentation, such as the rise in entropy with depth, we anticipate the need to introduce novel metrics, loss functions, as well as methodological enhancements.

ACKNOWLEDGMENT

G. Galiger and P. Kovács were supported by the ÚNKP-23-3 and the ÚNKP-22-5 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund, respectively. The authors also thank the Márton Áron College of Eötvös Loránd University, and the Collegium Talentum Programme of Hungary for supporting this

work. This project is co-financed by research subsidies granted by the government of Upper Austria under grant Wi-2022-597365/18-Au, and supported by Silicon Austria Labs (SAL), owned by the Republic of Austria, the Styrian Business Promotion Agency (SFG), the federal state of Carinthia, the Upper Austrian Research (UAR), and the Austrian Association for the Electric and Electronics Industry (FEEI).

REFERENCES

- [1] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proceedings of the IEEE*, pp. 1–35, 2023.
- [2] P. Burgholzer, M. Thor, J. Gruber, and G. Mayr, "Three-dimensional thermographic imaging using a virtual wave concept," *Journal of Applied Physics*, vol. 121, no. 10, p. 105102, 2017.
- [3] P. Kovács, B. Lehner, G. Thummerer, G. Mayr, P. Burgholzer, and M. Huemer, "Deep learning approaches for thermographic imaging," *Journal of Applied Physics*, vol. 128, no. 15, pp. 155 103–1–16, 2020.
- [4] —, "Surfing virtual waves to thermal tomography: From model- to deep learning-based reconstructions," *IEEE Signal Processing Magazine*, vol. 39, no. 1, pp. 55–67, 2022.
- [5] J. Xiang, Y. Dong, and Y. Yang, "FISTA-net: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging," *IEEE Transactions on Medical Imaging*, vol. 40, no. 5, pp. 1329–1339, may 2021. [Online]. Available: <https://doi.org/10.1109/tmi.2021.3054167>
- [6] G. Thummerer, G. Mayr, M. Haltmeier, and P. Burgholzer, "Photoacoustic reconstruction from photothermal measurements including prior information," *Photoacoustics*, vol. 19, p. 100175, 2020.
- [7] Y. Kong, C. Bennett, and C. Hyde, "A review of non-destructive testing techniques for the in-situ investigation of fretting fatigue cracks," *Materials & Design*, vol. 196, p. 109093, 2020.
- [8] D. Garcia, L. L. Tarnec, S. Muth, E. Montagnon, J. Porée, and G. Cloutier, "Stolt's f-k migration for plane wave ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 60, no. 9, pp. 1853–1867, 2013.
- [9] A. Reinke, M. D. Tizabi, M. Baumgartner, M. Eisenmann, D. Heckmann-Nötzel, A. E. Kavur, T. Rädtsch, C. H. Sudre, L. Acion, M. Antonelli *et al.*, "Understanding metric-related pitfalls in image analysis validation," *Nature Methods*, vol. 21, no. 2, p. 182–194, feb 2024. [Online]. Available: <http://dx.doi.org/10.1038/s41592-023-02150-0>