GRAD-CAM-INSPIRED INTERPRETATION OF NEARFIELD ACOUSTIC HOLOGRAPHY USING PHYSICS-INFORMED EXPLAINABLE NEURAL NETWORK

Hagar Kafri[⋆]*, Marco Olivieri*[†]*, Fabio Antonacci*[†]*, Mordehay Moradi*[⋆]*, Augusto Sarti*[†]*, Sharon Gannot*[⋆]

[⋆]Faculty of Engineering - Bar-Ilan University, Ramat Gan, Israel
[†]Department of Electronics, Information and Bioengineering - Politecnico di Milano, Milan, Italy

ABSTRACT

The interpretation and explanation of decision-making processes of neural networks are becoming a key factor in the deep learning field. Although several approaches have been presented for classification problems, the application to regression models needs to be further investigated. In this manuscript we propose a Grad-CAM-inspired approach for the visual explanation of neural network architecture for regression problems. We apply this methodology to a recent physics-informed approach for Nearfield Acoustic Holography, called Kirchhoff-Helmholtz-based Convolutional Neural Network (KHCNN) architecture. We focus on the interpretation of KHCNN using vibrating rectangular plates with different boundary conditions and violin top plates with complex shapes. Results highlight the more informative regions of the input that the network exploits to correctly predict the desired output. The devised approach has been validated in terms of NCC and NMSE using the original input and the filtered one coming from the algorithm.

Index Terms— Grad-CAM, regression, Nearfield Acoustic Holography, Physics-Informed Neural Network

1. INTRODUCTION

In the field of vibroacoustics, Nearfield Acoustic Holography (NAH) [1] represents a powerful tool for performing modal analysis in a fully contactless way. It aims to retrieve the surface velocity field on vibrating structures starting from the radiated pressure field measured with a microphone array placed nearby the object, called holographic plane. NAH is particularly suitable in the context of fragile or lightweight objects, such as musical instruments, where the use of accelerometer sensors could damage the surface or alter the vibration due to the local change of mass.

The problem of NAH relies on the inversion of the well-known Kirchhoff-Helmholtz (KH) integral [1]. However, this inversion is a highly ill-conditioned problem, thus requiring a regularisation technique [2, 3, 4, 5].

In the last decade Deep Neural Networks (DNNs) have gained huge popularity. From computer vision [6] and sound-field analysis [7] to speech enhancement [8] and musical acoustics [9], DNNs proved their ability to learn useful representations from data.

Recently, new data-driven-based solutions have been exploited also in the context of NAH. In [10, 11, 12] DNN architectures were proposed that are able to learn compressed representation of the data inferring useful information for the estimation of the desired velocity field. Although these approaches avoid the computation of complex matrix inversions, they discard prior information related to the physical problem. For this reason, authors in [13] presented a new approach based on Physics-Informed Neural Network (PINN) for NAH.

The Kirchhoff-Helmholtz-based Convolutional Neural Network (KHCNN) proposed in [13] combines the advantages of deep learning techniques with prior knowledge coming from the physical KH model. This solution proved to provide an accurate estimate of the surface velocity field on rectangular isotropic plates with different boundary conditions and orthotropic violin top plates with complex shape, outperforming other NAH methods available in the literature [5, 11].

It is worth noting that, with the increasing utilisation of deep learning strategies, there is a strong need of interpretability [14]. While DNNs enable superior performance, their decisions become difficult to explain and cannot be directly interpreted by a human-user. Recently, several approaches [15, 16] attempt to tackle this problem. One of the most used techniques for explainable CNN is the Gradient-weighted Class Activation Mapping (Grad-CAM) method [17]. Authors applied this methodology in the context of image classification and image captioning providing useful insights into failure modes of these models. However, to date, the explainable artificial intelligence field for regression problems lacks of consolidated methodologies [18].

In this paper we present a modified version of the Grad-CAM algorithm. The goal of this work is twofold. First, we develop a gradient-weighted algorithm able to produce a visual explanation for regression model estimates. Second, we apply the devised method to KHCNN model to infer information about the NAH problem. Note that we do not aim at increasing the reliability of the model and improving its performance, but rather to exploit its interpretability.

2. SCIENTIFIC BACKGROUND

2.1. Nearfield Acoustic Holography

Let us consider a surface $\mathcal{S}$ that vibrates at a frequency f. The complex exterior radiated pressure $p(\boldsymbol{r}, \omega)$ (i.e., magnitude and phase information) measured at a point $\boldsymbol{r}$ due to the vibration of the structure at $\omega = 2\pi f$ can be formulated with the Kirchhoff-Helmholtz (KH) integral [1] as

$$\begin{aligned} p(\boldsymbol{r}, \omega) = &\int_{\mathcal{S}} p(\boldsymbol{s}, \omega) \frac{\partial}{\partial \mathbf{n}} g_\omega(\boldsymbol{r}, \boldsymbol{s}) d\mathcal{S} \\ &-j\omega\rho_0 \int_{\mathcal{S}} v_n(\boldsymbol{s}, \omega) g_\omega(\boldsymbol{r}, \boldsymbol{s}) d\mathcal{S}, \end{aligned} \tag{1}$$

where $\boldsymbol{s}$ is a point on the surface $\mathcal{S}$, j is the imaginary unit, ρ_0 is the density of the medium (for air $\rho_0 = 1.225\,\mathrm{kg}\cdot\mathrm{m}^{-3}$) and $\mathbf{n}$ is the outward vector normal to the surface at $\boldsymbol{s}$. Equation (1) models the radiated complex pressure as the superposition of the normal velocity field $v_n(\boldsymbol{s}, \omega)$ and the pressure field $p(\boldsymbol{s}, \omega)$ on the vibrating surface considering the propagation from $\boldsymbol{s}$ to $\boldsymbol{r}$ with the free-field Green's function $g_\omega(\boldsymbol{r}, \boldsymbol{s})$ [1].

Nearfield Acoustic Holography (NAH) aims at computing $v_n(\boldsymbol{s}, \omega)$ starting from $p(\boldsymbol{r}, \omega)$ acquired by a microphone array on the holographic plane $\mathcal{H}$, thus $\boldsymbol{r} \in \mathcal{H}$. Notice that to satisfy the near-field condition, $\mathcal{H}$ is required to be close to $\mathcal{S}$ [1]. Therefore, the goal of NAH relies on the inversion of (1), namely

$$\hat{v}_n(\boldsymbol{s}, \omega)\Big|_{\boldsymbol{s}\in\mathcal{S}} \approx \Gamma^{-1}\big[p(\boldsymbol{r}, \omega)\big]\Big|_{\boldsymbol{r}\in\mathcal{H}}, \tag{2}$$

where Γ is a discrete estimator that approximates the soundfield on the hologram plane. However, the inverse propagation problem (2) is highly ill-conditioned, thus often necessitates a regularisation procedure.

2.2. Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) [17] is a post-hoc explanation via visualisation of class discriminative activation for a network. Similar to gradient-based methods, Grad-CAM leverages the structure of the CNN to produce a heatmap of the pixels from the input image that contribute to the prediction of a particular class.

Grad-CAM takes advantage on a key property of deep convolutional layers. It is well-known that CNNs act as high-level feature extractors. Moreover, the feature maps of the last convolution layer reflect the structural spatial information of objects in the image [19].

Nevertheless, differently from other gradient-based methods, which propagate the gradient till the input layer, Grad-CAM propagates the value up to the last convolutional layer of the network, in order to infer the high-level feature information from the neural network point of view.

Let assume a classification network with $c \in C$ classes having K feature maps in the last convolution layer, denoted as $\mathbf{A}^k$. Grad-CAM determines the neuron importance weights α_k^c for all maps $k = 1, \ldots, K$ and class c, as the global average pooling of the gradients over the spatial dimension, namely

$$\alpha_k^c = \frac{1}{Z} \sum_i^H \sum_j^W \frac{\partial y^c}{\partial A_{ij}^k}, \tag{3}$$

where y^c represents the score for class c and $Z = H \times W$ is the spatial resolution of the feature map with height H and width W.

The activation maps $\mathbf{A}^k$ are linear weighted summed with α_k^c weights, and ReLU [20] function is then applied to consider just the positive contributions of features, thus obtaining the relevance map $\mathbf{R}^c$ as

$$\mathbf{R}^c = \mathrm{ReLU}\left(\sum_k \alpha_k^c \mathbf{A}^k\right), \tag{4}$$

where $\mathbf{R}^c \in \mathbb{R}^{H\times W}$ is a 2D map with the same spatial dimension as the feature maps of the last convolution layer. Finally, $\mathbf{R}^c$ is linear interpolated up to the input image resolution and scaled in its magnitude to the interval $[0, 1]$ to obtain the final heatmap $\mathbf{H}^c$ for the class c.

3. PROPOSED METHOD

Inspired by the Grad-CAM algorithm [17], where the use of the gradients flowing through the convolutional layers produces a coarse localisation map highlighting the important regions of the input image, here we propose a visual explanation of CNN for regression problems based on a similar gradient-weighted approach.

Although the devised algorithm can be used with different regression problems, in this manuscript we consider the NAH problem (2) as a case study.

3.1. Kirchhoff-Helmholtz-based CNN for NAH

Among several data-driven strategies for NAH [10, 11, 12], here we focus on the Kirchhoff-Helmholtz-based Convolution Neural Network (KHCNN) architecture presented in [13]. KHCNN addresses the NAH problem (2) with a physics-informed approach by combining the advantages of CNN with the KH model that governs the physical phenomenon.

The block scheme of KHCNN architecture is depicted in Fig. 1. $\mathbf{P}_\mathcal{H}(\omega), \widehat{\mathbf{P}}_\mathcal{H}(\omega) \in \mathbb{C}^{M_1\times M_2}$ are the acoustic pressure fields on $\mathcal{H}$ with dimension $M_1 \times M_2$ and $\widehat{\mathbf{V}}(\omega)$, $\widehat{\mathbf{P}}_\mathcal{S}(\omega) \in \mathbb{C}^{N_1\times N_2}$ are the normal velocity field and the pressure field on $\mathcal{S}$, respectively, with dimension $N_1 \times N_2$. The architecture is composed of two main blocks. The first consists of a CNN with one encoder $\mathcal{E}$ and two decoders $\mathcal{D}_1$ and $\mathcal{D}_2$ to

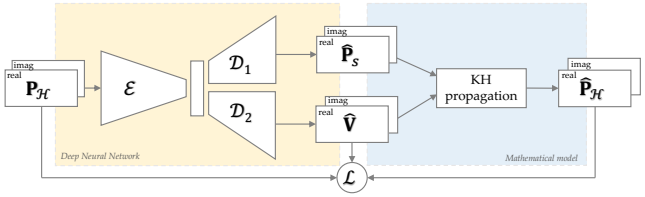

Fig. 1. Block scheme of KHCNN architecture.

estimate the latent variable $\widehat{\mathbf{P}}_{\mathcal{S}}$ and the velocity $\widehat{\mathbf{V}}$ on the surface, respectively, from the measured hologram pressure $\mathbf{P}_{\mathcal{H}}$ at the input. The second block validates the CNN outputs by applying the discretized version of the KH integral (1), thus having an estimate of $\widehat{\mathbf{P}}_{\mathcal{H}}$ from the forward propagation.

The loss function $\mathcal{L}$ considers both the desired velocity and the hologram pressure, thus imposing a physical meaning to the regularisation, as discussed in [13].

3.2. Grad-CAM-inspired for KHCNN

The proposed algorithm aims to visualise the heatmap $\mathbf{H}(\omega)$ to highlight the important regions of the input $\mathbf{P}_{\mathcal{H}}(\omega)$ needed to produce the KHCNN estimate $\widehat{\mathbf{V}}(\omega)$. Notice that, differently from Grad-CAM [17], the resulting heatmap $\mathbf{H}$ does not depend on class c, but on the same vibrating frequency ω of the input $\mathbf{P}_{\mathcal{H}}$.

The devised algorithm can operate with different resolution of the KHCNN. Hence, with input in $M = M_1 \times M_2$ points and output in $N = N_1 \times N_2$ points. Moreover, it produces a heatmap $\mathbf{H}$ for each of the two decoders, thus disclosing the regions of the input that drives the outputs, $\widehat{\mathbf{P}}_{\mathcal{S}}$ and $\widehat{\mathbf{V}}$ respectively.

We use the loss function defined in [13] as $\mathcal{L} = 0.5 \cdot \mathcal{L}_{\mathrm{Re}} + 0.5 \cdot \mathcal{L}_{\mathrm{Im}}$, with

$$\begin{aligned}\mathcal{L}_{\mathrm{Re}} &= \left\|\mathrm{Re}\left(\mathbf{V}\right) - \mathrm{Re}\left(\widehat{\mathbf{V}}\right)\right\|_2^2 + \left\|\mathrm{Re}\left(\mathbf{P}_{\mathcal{H}}\right) - \mathrm{Re}\left(\widehat{\mathbf{P}}_{\mathcal{H}}\right)\right\|_2^2, \\ \mathcal{L}_{\mathrm{Im}} &= \left\|\mathrm{Im}\left(\mathbf{V}\right) - \mathrm{Im}\left(\widehat{\mathbf{V}}\right)\right\|_2^2 + \left\|\mathrm{Im}\left(\mathbf{P}_{\mathcal{H}}\right) - \mathrm{Im}\left(\widehat{\mathbf{P}}_{\mathcal{H}}\right)\right\|_2^2,\end{aligned} \tag{5}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ are operators that take the real and imaginary part of the complex field, respectively. The pressure and velocity fields in (5) are represented without the dependence of ω for the sake of simplicity. Moreover, the true values $\mathbf{V}$ come from the synthesised datasets [13] computed with Finite Element simulations.

In order to tackle the regression problem, we modify Equation (3) by deriving the loss function given the activation map, namely

$$\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial \mathcal{L}}{\partial A_{ij}^k}, \tag{6}$$

where $\mathbf{A}^k$ is the activation map of the last layer of KHCNN decoder. In Equation (6) we replace the score with the loss function based on the assumption that the loss function holds the information of the meaningful parts of the input with respect to the output.

Finally, the resulting heatmap $\mathbf{H}$ is linear interpolated to the input resolution $M_1 \times M_2$ with values in $[0, 1]$.

4. RESULTS

4.1. Setup

We analyse the explainability of KHCNN by applying the proposed method to decoder $\mathcal{D}_2$, thus obtaining the heatmap $\mathbf{H}$ related to the estimate $\widehat{\mathbf{V}}$.

Using the pre-trained KHCNN architecture of [13], we focus on two datasets: 672 aluminium rectangular plates and 1568 violin top plates made of Sitka spruce. The datasets was generated using *COMSOL Multiphysics*® software simulating the radiated pressure in M points and the normal velocity field in N points of different plates excited at different ω.

The aluminium plate dataset available in [21] comprises 15 570 samples of rectangular vibrating plates with different dimensions and boundary conditions (BCs): simply supported, clamped and free BCs, that characterised the vibration based on the conditions imposed to the edges of the structure [22]. The input resolution of $\mathbf{P}_{\mathcal{H}}$ is $M = M_1 \times M_2 = 16 \times 64$ points. On the other hand, the violin dataset comprises 7256 samples of plates with free BCs and input resolution of $M = M_1 \times M_2 = 8 \times 8$. In both cases KHCNN estimates the desired velocity on $N = N_1 \times N_2 = 16 \times 64$ points.

4.2. Evaluation

Fig. 2 shows three examples of rectangular plates having different BCs along with the input hologram pressure in $M = 1024$ points, the velocity estimate and its ground truth in $N = 1024$ points. From the output of decoder $\mathcal{D}_2$ we apply the proposed algorithm to obtain the heatmap $\mathbf{H} \in \mathbb{R}^{16 \times 64}$ as depicted in the last row of the figure.

In general, inspecting the resulting heatmaps $\mathbf{H}$, we noticed that the network focuses on the regions where vibrational lobes of the velocity fields are present. In particular, KHCNN exploits the symmetry of vibrations and the edges for different BCs inferring the estimates as interpolation of the vibrational patterns.

For the simply supported and clamped cases in Fig. 2(a) and Fig. 2(b), respectively, the regions with maximum amplitude correspond to the areas of the largest magnitude velocity $|\widehat{\mathbf{V}}|$ discarding the edges, where the velocity is zero. Conversely, for the free BCs, where the edges are free to move, the resulting heatmap highlights that the network based the decision by taking into account also the external regions of the input, as shown in Fig. 2(c)

For the violin top plate dataset we analyse the KHCNN architecture using $M = 64$ points for the input pressure. In Fig. 3 we show three examples of violin top plates vibrating

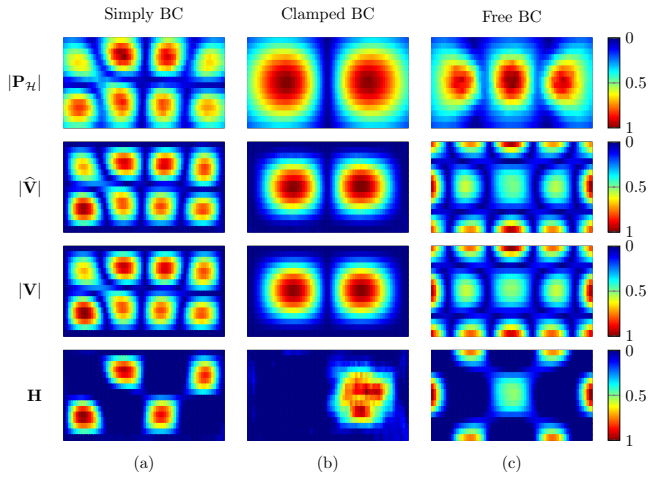

Fig. 2. Reconstruction examples on rectangular plates for simply (a), clamped (b), and free (c) BCs. First row is the magnitude hologram pressure at the input with $M_1 = 16$ and $M_2 = 64$. Second and third rows are the magnitude of the velocity estimate and ground truth, respectively. The last row depicts the resulting heatmap of the algorithm coming from decoder $\mathcal{D}_2$.

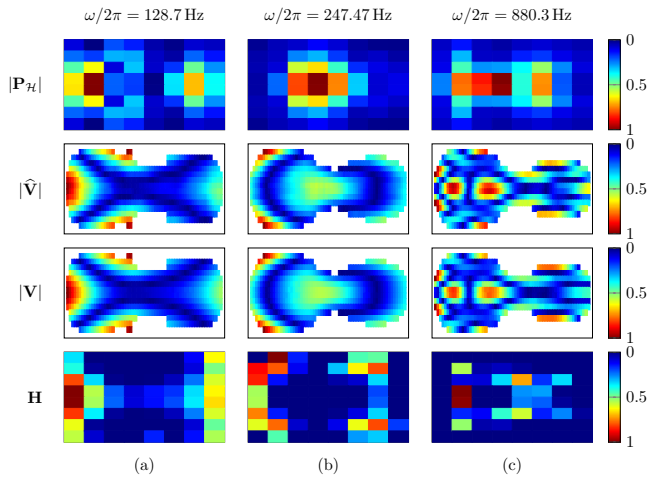

Fig. 3. Reconstruction examples on violin top plates with free BCs with different shapes and vibrating frequencies. The magnitude of the input hologram pressure with $M_1 = 8$ and $M_2 = 8$ is depicted in the first row. Second and third rows are the magnitude of the velocity estimate and ground truth, respectively. The resulting heatmap of the algorithm is shown in the last row.

at different frequencies with the realtive heatmaps $\mathbf{H} \in \mathbb{R}^{8\times 8}$ computed by the proposed algorithm from decoder $\mathcal{D}_2$.

We can notice that the network is more affected by the edges of the input pressure mainly at low frequencies, as shown in Fig. 3(a) and Fig. 3(b). On the other hand, when the frequency increases the network focuses on the internal regions of the pressure, as in Fig. 3(c). However, also in this case the main regions of interest correspond to the maximum lobes of the velocity field.

4.3. Validation

In order to assess the performance of the proposed explainable algorithm we computed the Normalised Cross Correlation (NCC) and Normalised Mean Square Error (NMSE) between the KHCNN estimates coming from the acquired input pressure and the filtered one by the resulting heatmap $\mathbf{H}$, namely

$$\mathrm{NCC} = \frac{|\hat{\mathbf{v}}_f^H \cdot \hat{\mathbf{v}}|}{\|\hat{\mathbf{v}}_f\|_2 \cdot \|\hat{\mathbf{v}}\|_2}, \quad \mathrm{NMSE} = 10\log_{10}\left(\frac{\mathbf{e}^H \cdot \mathbf{e}}{\hat{\mathbf{v}}^H \cdot \hat{\mathbf{v}}}\right), \tag{7}$$

where ^{H} is the Hermitian transpose operator, $\mathbf{e} = \hat{\mathbf{v}}_f - \hat{\mathbf{v}}$, and $\hat{\mathbf{v}}$, $\hat{\mathbf{v}}_f$ are the vectorized form of $\widehat{\mathbf{V}} = \mathrm{KHCNN}(\mathbf{P}_{\mathcal{H}})$ and $\widehat{\mathbf{V}}_f = \mathrm{KHCNN}(\mathbf{P}_{\mathcal{H}} \odot \mathbf{H})$, respectively, where $\odot$ denotes the Hadamard product. Notice that, $\hat{\mathbf{v}}_f$ and $\hat{\mathbf{v}}$ should ideally match if the computed $\mathbf{H}$ highlights correct regions of the input, which drives the output decision. Therefore, $\mathrm{NCC} \in [0, 1]$ is optimum when it is close to 1 and NMSE, in decibel, should be as lower as possible. We computed Equation (7) for the 7256 samples of the violin dataset reaching an average NCC value of 0.98 and $\mathrm{NMSE} = -13.93\,\mathrm{dB}$ on average.

5. CONCLUSIONS

In this manuscript we proposed a modified version of the Grad-CAM algorithm for regression problems. We applied the devised method to a recent physics-informed approach for NAH, called KHCNN. We focused on the explainable KHCNN when operating with vibrating rectangular plates having different boundary conditions and with violin top plates with different shapes. The proposed algorithm is able to compute visual heatmaps that highlight the important regions of the input responsible of the KHCNN velocity estimates. Moreover, the methodology is flexible enough to be able to work with different resolutions of the input hologram pressure field. We validated the performance in terms of NCC and NMSE by comparing the KHCNN estimates stemming from the acquired hologram pressure and the filtered one with the computed heatmap, thus proving the reliability of the method. Results help in the interpretation of the KHCNN decision process inspecting how the network infers the symmetry of vibrational patterns and the behaviour at the edges. Future works can exploit this knowledge for the implementation of new architectures. In particular, we foresee new compact and lightweight architectures defined with well-designed constraints to archive good results for more complex objects, such as industrial machineries.

6. REFERENCES

[1] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic press, 1999.

[2] E. G. Williams, "Regularization methods for near-field acoustical holography," *The Journal of the Acoustical Society of America (JASA)*, vol. 110, no. 4, pp. 1976–1988, 2001.

[3] A. Schuhmacher, J. Hald, K. Rasmussen, and P. Hansen, "Sound source reconstruction using inverse boundary element calculations," *The Journal of the Acoustical Society of America (JASA)*, vol. 113, pp. 114–27, 02 2003.

[4] G. Chardon, L. Daudet, A. Peillot, F. Ollivier, N. Bertin, and R. Gribonval, "Near-field acoustic holography using sparse regularization and compressive sampling principles," *The Journal of the Acoustical Society of America (JASA)*, vol. 132, no. 3, pp. 1521–1534, 2012.

[5] E. Fernandez-Grande, A. Xenaki, and P. Gerstoft, "A sparse equivalent source method for near-field acoustic holography," *The Journal of the Acoustical Society of America (JASA)*, vol. 141, no. 1, pp. 532–542, 2017.

[6] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational Intelligence and Neuroscience*, vol. 2018, 2018.

[7] M. Pezzoli, D. Perini, A. Bernardini, F. Borra, F. Antonacci, and A. Sarti, "Deep prior approach for room impulse response reconstruction," *Sensors*, vol. 22, no. 7, pp. 2710, 2022.

[8] A. Pandey and D. Wang, "A new framework for cnn-based speech enhancement in the time domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 7, pp. 1179–1188, 2019.

[9] M. Olivieri, R. Malvermi, M. Pezzoli, M. Zanoni, S. Gonzalez, F. Antonacci, and A. Sarti, "Audio information retrieval and musical acoustics," *IEEE Instrumentation & Measurement Magazine*, vol. 24, no. 7, pp. 10–20, 2021.

[10] M. Olivieri, M. Pezzoli, R. Malvermi, F. Antonacci, and A. Sarti, "Near-field acoustic holography analysis with convolutional neural networks," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Seoul, Korea, August 2020, Institute of Noise Control Engineering, vol. 261, pp. 5607–5618.

[11] M. Olivieri, M. Pezzoli, F. Antonacci, and A. Sarti, "Near field acoustic holography on arbitrary shapes using convolutional neural network," in *European Signal Processing Conference (EUSIPCO)*, Dublin, Ireland, August 2021.

[12] J. Wang, Z. Zhang, Z. Li, and Q. Huang, "Research on joint training strategy for 3d convolutional neural network based near-field acoustical holography with optimized hyperparameters," *Measurement*, p. 111790, 2022.

[13] M. Olivieri, M. Pezzoli, F. Antonacci, and A. Sarti, "A physics-informed neural network approach for nearfield acoustic holography," *Sensors*, vol. 21, no. 23, pp. 7834, 2021.

[14] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining explanations: An overview of interpretability of machine learning," in *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*. IEEE, 2018, pp. 80–89.

[15] G. Montavon, W. Samek, and K. Müller, "Methods for interpreting and understanding deep neural networks," *Digital signal processing*, vol. 73, pp. 1–15, 2018.

[16] Y. Zhang, P. Tiňo, A. Leonardis, and K. Tang, "A survey on neural network interpretability," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2021.

[17] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

[18] S. Letzgus, P. Wagner, J. Lederer, W. Samek, K. Müller, and G. Montavon, "Toward explainable artificial intelligence for regression models: A methodological perspective," *IEEE Signal Processing Magazine*, vol. 39, no. 4, pp. 40–58, 2022.

[19] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[20] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *International Conference on Machine Learning (ICML)*, Haifa, Israel, June 2010, pp. 807–814.

[21] M. Olivieri, M. Pezzoli, F. Antonacci, and A. Sarti, "Nah rectangular plate dataset (nearfield acoustic holography)," `https://doi.org/10.5281/zenodo.5702615`, (accessed on 20 October 2022).

[22] A. W. Leissa, *Vibration of plates*, vol. 160, Scientific and Technical Information Division, National Aeronautics and Space Administration, 1969.