Vision Transformer for Accelerated Design of Very Low-frequency Metasurface Absorber

Dev Sanghvi[1], **Prit Thakkar**[1], **Naimish Bhuva**[1], **Heet Vekariya**[1],
Bishwajit Dey[1,2], **Krupali Donda**[1*]
[1]Adani University, Ahmedabad, Gujarat, 382421, India
[2]University of Johannesburg, 2092, Johannesburg, South Africa
`krupalip18@gmail.com`

Abstract

Deep Learning (DL), a branch of machine learning, offers an effective way to comprehend and design structures by creating data-driven methodologies to supplement conventional physics and formula-based approaches. Classical approaches require deep expertise in metasurface design and highly intensive iterative computations using Finite Element Methods (FEM). To overcome this limit and accelerate non-intuitive acoustic metasurface absorber design, we implement a vision transformer (ViT) in this work. Once trained, the network outperforms traditional FEM simulation software, being five orders of magnitude faster. We envision that our implemented deep learning-based approach will be broadly used for fast acoustic designs with minimal human intervention.

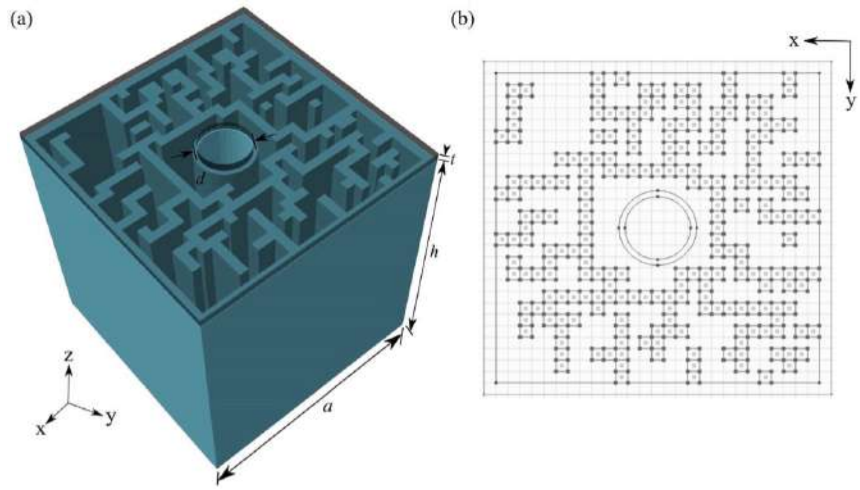
We propose a non-intuitive metasurface structure featuring a resonant cavity with a random-shaped channel and a circular aperture, as shown in Fig. 1(a). The cavity is covered by a 1 mm thick plate with a central hole, and a circular aperture attached to the hole provides powerful tunability to achieve perfect absorption at low frequencies. The structure, with a fixed size of 58 mm and a 2 mm wall thickness, is subdivided into 2 mm × 2 mm squares forming propagation channels with varying shapes and lengths. To ensure accuracy in the narrow 80–100 Hz low-frequency range, 201 frequency points are used for calculations.

Implementation begins by decomposing the input image of the metasurface into a set of 64 patches, each sized 8x8 pixels. These patches are flattened and linearly projected into 64-dimensional vectors. To retain the positional context lost during flattening, a unique positional encoding is added to each patch vector. The parallel processing allows the model to dynamically focus on different parts of the input image, assessing which features are most relevant for predicting the metasurface's spectral response. The attention outputs are then combined and normalized to stabilize the training process. Following attention, the data passes through a point-wise feed-forward network within each encoder. After processing through the prescribed number of Transformer layers, the output is fed to a global average pooling operation. The network concludes with a linear activation function that maps this vector to 20 outputs, directly linking to the metasurface's absorption properties.

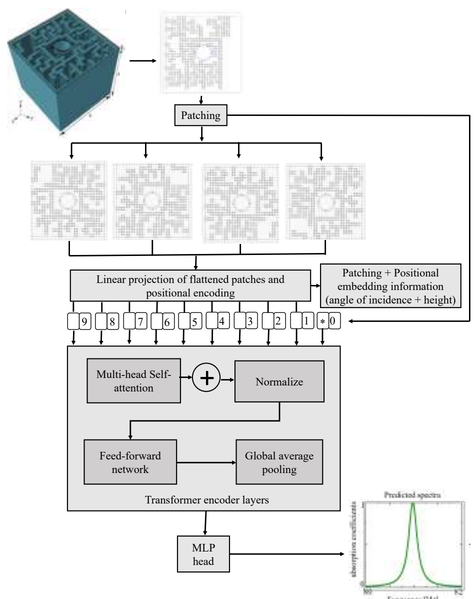
Training incorporates techniques such as model check pointing to save the best-performing models and learning rate reduction on plateaus to adjust the learning rate dynamically based on validation loss. Through these sophisticated mechanisms, the ViT effectively captures and processes complex spatial relationships in metasurface images. After the training process was completed, the model achieved an average MSE of 0.01200 for the predicted absorption coefficients on the validation set. To improve the stability of the learning process and increase training efficiency, layer normalization is employed. Each encoder layer, which consists of multi-head self-attention and feed-forward networks,

[*]Corresponding author: `krupalip18@gmail.com`, ORCID: 0000-0001-6802-0145

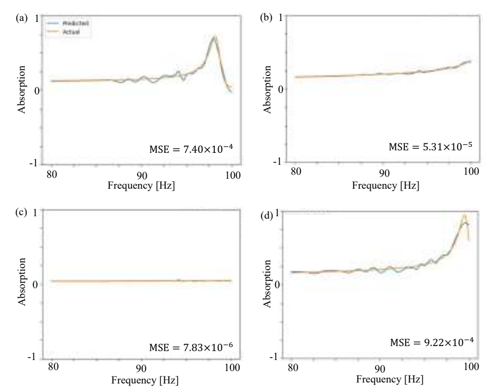
1st Conference on Applied AI and Scientific Machine Learning (CASML 2024).



(a) 3D view of proposed metasurface absorber for the oblique angle incidence.



(b) Deep learning network architecture for the metasurface absorber design. The implemented network consists of transformer encoder layers with a multi-head self-attention mechanism.



(c) Comparison of predicted (blue curve) and actual values (orange curve) of absorption spectra for four different cases.

is augmented with layer normalization to mitigate internal covariate shifts. Figure 1c shows the comparison of predicted (represented with blue curves) and actual values (represented with orange curve) of absorption spectra for four different cases. For each case, the MSE is listed in the inset of the graph. The presented MSE values indicate the predicted absorption spectra by the ViT (orange curves) closely match the FEM simulations (blue curves), for a variety of different propagation channels and incidence angles.

In this article, we develop and implement a forward simulator using a based on Vision Transformer, to design the behaviour of complex and non-intuitive acoustic metasurface absorbers. The proposed approach addresses the limitations of traditional design methods by handling a large set of input parameters. Once adequately trained, this deep learning simulator achieves performance that is over four orders of magnitude faster than traditional Finite Element Method-based solvers.

References

Badreddine Assouar, Bin Liang, Ying Wu, Yong Li, Jian-Chun Cheng, and Yun Jing. Acoustic metasurfaces. *Nature Reviews Materials*, 3(12):460–472, 2018.

Krupali Donda, Ying Zhu, Shi Wei Fan, Lei Cao, Yong Li, and Badreddine Assouar. Extreme low-frequency ultrathin acoustic absorbing metasurface. *Applied Physics Letters*, 115(17), 2019.

Yong Li and Badreddine M. Assouar. Acoustic metasurface-based perfect absorber with deep subwavelength thickness. *Applied Physics Letters*, 108(6), 2016.

Victor Romero-García, Georgios Theocharis, Olivier Richoux, and Vincent Pagneux. Use of complex frequency plane to design broadband and sub-wavelength absorbers. *The Journal of the Acoustical Society of America*, 139(6):3395–3403, 2016.