

Autonomous Selection of Energy-Based Ultrasound Speckle Tracking Parameters Using Deep Learning

Md Ashikuzzaman¹, Ahmed El-Desoky¹, MD Jahin Alam¹, and Muyinatu A. Lediju Bell^{1,2,3}

¹Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA

²Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

³Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA

Abstract—Second-order ultrasound elastography (SOUL) is a regularized speckle-tracking algorithm with demonstrated promise to produce high-quality strain images. SOUL optimizes a cost function that consists of a data fidelity term and a regularization term. However, SOUL manually selects its regularization weights to generate strain images, which limits its efficiency and usability by making it highly user-dependent. We developed a deep learning approach to autonomously select the regularization weights required to reduce user dependency. A convolutional neural network (CNN) was trained on simulated datasets and fine-tuned using phantom data to classify a strain image to be acceptable or unacceptable and predict optimal parameters based on the strain image quality. When evaluated on simulated and public phantom datasets, the CNN achieved 100.00% recall with F1 scores $\geq 95\%$. Strain images from manual and automated tuning showed similar smoothness and contrast, highlighting the potential to eliminate manual SOUL parameter selection. The CNN-based autonomous selector reduces user dependence, enabling optimal strain imaging for inexperienced users and clinicians.

Index Terms—Ultrasound strain elastography, speckle-tracking, deep learning, convolutional neural network, autonomous parameter selection

I. INTRODUCTION

Ultrasound elastography [1], [2] is a non-invasive imaging technique that assesses tissue stiffness to support the diagnosis of various conditions (e.g., tumors, cancers, benign lesions, cysts), based on the assumption that pathological tissues show altered mechanical properties from healthy tissue. To implement strain elastography with manual palpations, a handheld probe applies compression to the tissue of interest while acquiring beamformed ultrasound radiofrequency (RF) data. A speckle-tracking algorithm estimates the tissue displacement between pre- and post-compression RF frames, which is then spatially differentiated to produce a strain map. This technique has been applied in clinical applications that include breast imaging [3], [4], liver fibrosis staging [5], and cardiovascular health assessments [6].

Accurate displacement tracking between pre- and post-deformed ultrasound frames is required to generate high-quality strain images in ultrasound elastography [7]. However, this task is inherently ill-posed due to the nature of ultrasound data (e.g., numerous samples exhibit identical amplitudes, shared characteristics, and similar patterns) [4]. To overcome this challenge, window-based displacement tracking

algorithms [8], [9] divide ultrasound frames into multiple data windows and impose the constraint that all samples within a window share the same displacement. Based on this assumption, these methods estimate displacement by identifying the best-matching post-deformation window using normalized cross-correlation [10], [11] or zero-phase-crossing [12], [13]. While relatively straightforward to implement, window-based algorithms are known to be noise-sensitive [14] and involve a trade-off between accuracy and resolution depending on window size [15], [16]. Deep learning-based techniques [14], [17]–[19] were recently introduced to learn statistical patterns and correlations in pre- and post-deformed ultrasound data for displacement tracking. During testing and deployment, these methods primarily rely on inference and GPU acceleration, making them well-suited for real-time applications. However, large amounts of training data are required to develop accurate models [18], and the limited availability of clinical datasets remains a key challenge to broader adoption in ultrasound displacement tracking. Despite their promise, these models demand large annotated datasets for training, posing a barrier due to limited clinical data availability [18].

An alternative approach has been to consider energy function optimization [20]–[22], which models tissue deformation as a non-linear cost function to achieve spatially coherent displacement estimates [23]. Among energy-based methods, Second-Order Ultrasound eLastography (SOUL) [22] has shown promising results by integrating data fidelity and second-order regularization terms. However, its success heavily depends on properly selected regularization parameters. Currently, SOUL requires manual tuning of its regularization weights to produce high-quality strain images, making it highly user-dependent and limiting its efficiency and usability.

In this work, we propose a deep learning-based framework to automate the selection of regularization parameters in SOUL for strain elastography, as illustrated in Fig. 1. We trained a convolutional neural network (CNN) on simulated and phantom datasets to learn the relationship between input RF data and optimal SOUL parameters. We then classified strain images generated with different SOUL parameter settings as either suitable or unsuitable based on learned quality characteristics. Finally, we automated the selection of the optimal strain image along with its corresponding parameter set to enhance autonomy in SOUL-based elastography.

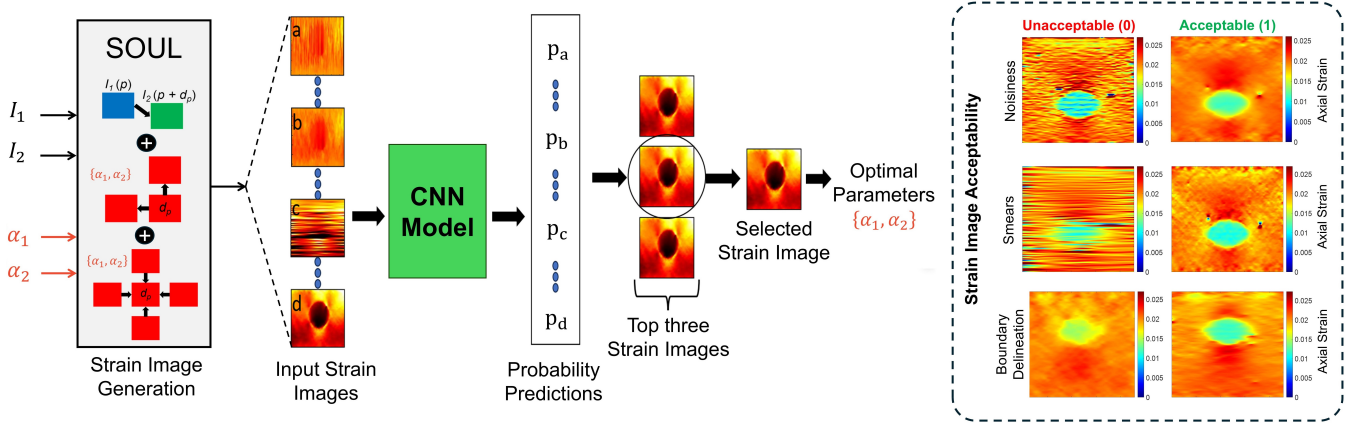


Fig. 1. Illustration of the autonomous SOUL parameter selection strategy.

II. METHODS

A. Datasets

We utilized two publicly available datasets (one simulated, one phantom, both from `data.sonography.ai` [24]) to train and evaluate our approach. The simulated dataset included 13 phantoms each modeled with background and target elastic moduli of 20 kPa and 40 kPa, respectively, and a Poisson's ratio of 0.49. A 2% uniaxial compression was applied using ABAQUS (Providence, RI), and corresponding pre- and post-compression RF data were simulated using Field II [25], [26]. The public phantom dataset consisted of RF frame pairs from an experimental phantom (CIRS Model 059, Norfolk, VA) [27]–[29] with 20 ± 5 kPa background, acquired using an Alpinion E-Cube 12R scanner with an L3-12H probe, and included a hard inclusion with a higher elastic modulus than the background. The transmit and sampling frequencies were 10 MHz and 40 MHz, respectively. Finally, four private phantoms were collected for fine-tuning the deep learning network.

B. SOUL Parameters of Interest

The SOUL algorithm [22] uses pre- and post-compressed beamformed RF frames, $I_1(i, j)$ and $I_2(i, j)$, respectively, to estimate the incremental displacement field between tissue states. The indices i and j correspond to the axial and lateral positions, respectively, with $1 \leq i \leq m$ and $1 \leq j \leq n$. Dynamic programming [30] was employed to generate initial displacement estimates $a_{i,j}$ and $l_{i,j}$ in the axial and lateral directions. The following SOUL cost function [22] was then formulated and analytically optimized to obtain the incremental displacement fields $\Delta a_{i,j}$ and $\Delta l_{i,j}$:

$$C(\Delta a_{1,1}, \dots, \Delta a_{m,n}, \Delta l_{1,1}, \dots, \Delta l_{m,n}) = \sum_{j=1}^n \sum_{i=1}^m \{D(i, j, a_{i,j}, l_{i,j}, \Delta a_{i,j}, \Delta l_{i,j})\} + \alpha_1 \|\partial_y a\|_2^2 + \alpha_2 \|\partial_x a\|_2^2 + \beta_1 \|\partial_y l\|_2^2 + \beta_2 \|\partial_x l\|_2^2 + w\alpha_1 \|\partial_y^2 a\|_2^2 + w\alpha_2 \|\partial_x^2 a\|_2^2 + w\beta_1 \|\partial_y^2 l\|_2^2 + w\beta_2 \|\partial_x^2 l\|_2^2 \quad (1)$$

where $\partial_y(\cdot)$ and $\partial_x(\cdot)$ are first-order axial and lateral derivatives, respectively, $\partial_y^2(\cdot)$ and $\partial_x^2(\cdot)$ are second-order axial and lateral derivatives, respectively, α_1 and α_2 regularize the axial displacement field in the axial and lateral directions, respectively, β_1 and β_2 regularize the lateral displacement field in the axial and lateral directions, respectively, w is the ratio between the second-order and first-order regularization weights, and $D(\cdot)$ is a measure of data amplitude similarity.

Because axial strain estimation is less affected by lateral displacement [31], this work sets $\beta_1 = \alpha_1/2$ and $\beta_2 = \alpha_2/2$. A fixed weight of $w = 100$ is also used, as SOUL is generally less sensitive to the ratio between first- and second-order terms. These simplifications reduce the tunable parameters to α_1 and α_2 . To eliminate user-dependent (i.e., manual) tuning and improve consistency, the goal of our deep learning approach is to automate the selection of the parameter set $\{\alpha_1, \alpha_2\}$.

C. Data Label Assignment

Strain images were labeled as acceptable (1) or unacceptable (0) based on qualitative and quantitative assessments. Qualitative assessments were based on background smoothness, contrast, and lesion boundary clarity. Quantitative assessments were based on signal-to-noise ratio (SNR) and strain ratio (SR), which are two standard metrics for evaluating strain image quality, defined as:

$$\text{SNR} = \frac{\bar{s}_b}{\sigma_b} \quad (2)$$

$$\text{SR} = \frac{\bar{s}_t}{\bar{s}_b} \quad (3)$$

where $\bar{s}_b$ and $\bar{s}_t$ are the mean strain values in the background and target ROIs, respectively, and σ_b is the standard deviation in the background. Two 50×50 pixel regions of interest (ROIs), one in the target and one in the background, were used to compute SNR and SR. Acceptable SNR and SR values ranged 15–120 and 0.5–0.68, respectively, for simulated data, and 22–45 and 0.39–0.42, respectively, for public phantom data.

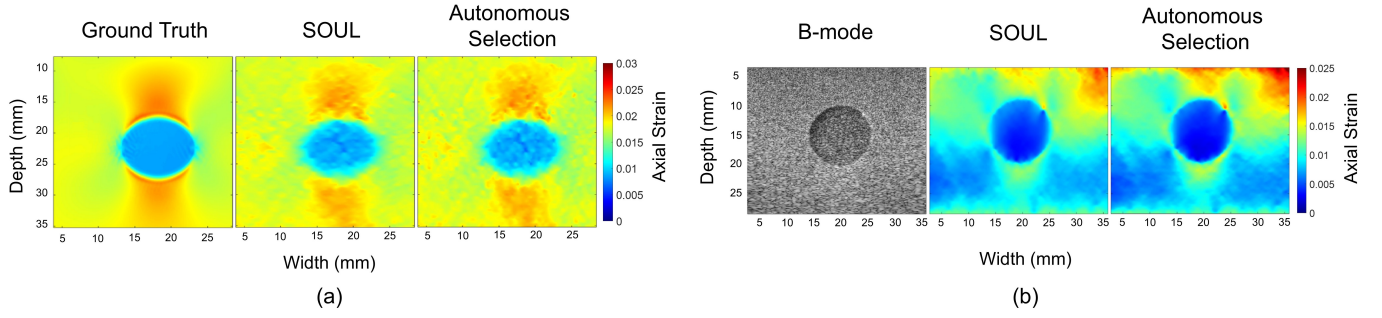


Fig. 2. Axial strain results obtained with SOUL and the proposed autonomous parameter selection strategy applied to (a) simulated and (b) phantom data.

To create labeled images for training, validation, and testing, the α_1 and α_2 parameters were varied (i.e., ranging 0.001-1000 and 0-500, respectively, for simulated data and 0.00001-1900 and 0-100, respectively, for phantom data). These variations resulted in a total of 3834 simulated and 576 phantom strain images that were acceptable (i.e., 1), and 6605 simulated and 580 phantom strain images that were unacceptable (i.e., 0). Examples of acceptable and unacceptable images are shown in Fig. 1.

D. CNN Training

We designed a binary classifier CNN to assess strain image quality. The CNN consisted of four 2D convolutional layers (3×3 kernel, stride 1, ReLU activation) with 16, 32, 64, and 128 filters, each followed by max-pooling and 30% dropout. The resulting features were flattened and passed through a fully connected layer with 256 ReLU units and 50% dropout, followed by a sigmoid-activated output unit.

The CNN was trained on simulated data (80%-20% train-validation split) using binary cross-entropy loss, Adam optimizer (learning rate 10^{-3}), and 32 batch size up to 200 epochs with early stopping. To improve generalization, the pre-trained CNN was fine-tuned on phantom data using a reduced learning rate (3×10^{-5}) and the same training configuration.

To select optimal SOUL parameters during testing, the top three strain images based on CNN confidence were inspected to safeguard our final prediction against false positives, and the corresponding parameter set $\{\alpha_1, \alpha_2\}$ producing the most acceptable strain image was chosen. To quantify performance, the following metrics were calculated:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

where TP, TN, FP, and FN refer to true positive, true negative, false positive, and false negative, respectively, based on the labels described in Section II-C.

TABLE I
PERFORMANCE METRICS OF THE CNN MODEL

Data	Accuracy	Precision	Recall	F1-score
Simulated	96.22%	90.85%	100.00%	95.21%
Experimental	99.45%	98.96%	100.00%	99.47%

III. RESULTS AND DISCUSSION

Fig. 2 shows axial strain results obtained with SOUL and the proposed method for test instances of simulated and public phantom data. With simulated data, the strain images obtained with SOUL ($\alpha_1 = 40$, $\alpha_2 = 0.05$) and with the autonomously selected method ($\alpha_1 = 38.26$, $\alpha_2 = 0.0055$) resemble the ground truth and successfully show the inclusion. With phantom data, the autonomously selected method shows sharper target boundary ($\alpha_1 = 19.35$, $\alpha_2 = 0.0125$) compared to the SOUL image obtained with manual parameters ($\alpha_1 = 40$, $\alpha_2 = 0.05$).

Table I reports the performance of the CNN when classifying strain images as acceptable or unacceptable. With the simulated and phantom data, the CNN achieved 100.00% recall, indicating that acceptable strain images were correctly identified with each network input. In addition, the F1 scores exceeded 95%, which represents an acceptable [32] performance balance between precision and recall.

Future work will apply our approach to *in vivo* data. Additional future directions include alternative network architectures, such as transformer-based models [33], which can address long-range dependencies and complex patterns [34]. Overall, the results herein are promising to enhance the accuracy and reliability of parameter selection in SOUL strain imaging.

IV. CONCLUSION

We successfully implemented a deep learning-based framework to autonomously select regularization parameters in energy-based ultrasound strain elastography. By training a CNN on simulated and phantom datasets, the proposed approach remarkably identified optimal SOUL parameters, eliminating the need for manual tuning. The proposed approach has the potential to reduce user dependence when implementing SOUL, which is generally beneficial to simplifying associated

diagnostic, treatment monitoring, and decision-making procedures.

ACKNOWLEDGMENT

This work is supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number R01EB032960.

REFERENCES

- [1] J. Ophir *et al.*, "Elastography: A quantitative method for imaging the elasticity of biological tissues," *Ultrasonic Imaging*, vol. 13, pp. 111–134, 1991.
- [2] E. E. Konofagou *et al.*, "Myocardial elastography— A feasibility study in vivo," *Ultrasound in Medicine & Biology*, vol. 28, no. 4, pp. 475–482, 2002.
- [3] M. G. Kibria and M. K. Hasan, "A class of kernel based real-time elastography algorithms," *Ultrasonics*, vol. 61, pp. 88–102, 2015.
- [4] M. Ashikuzzaman, A. Sharma, N. Venkatayogi, E. Oluyemi, K. Myers, E. Ambinder, H. Rivaz, and M. A. L. Bell, "MixTURE: L_1 -norm-based mixed second-order continuity in strain tensor ultrasound elastography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 11, pp. 1389 – 1405, 2024.
- [5] M. Ashikuzzaman, B. Peng, J. Jiang, and H. Rivaz, "Alternating direction method of multipliers for displacement estimation in ultrasound strain elastography," *Medical Physics*, vol. 51, no. 5, pp. 3521–3540, 2024.
- [6] D. deCampo and M. Hwang, "Characterizing the neonatal brain with ultrasound elastography," *Pediatric Neurology*, vol. 86, pp. 19–26, 2018.
- [7] M. Ashikuzzaman, A. Héroux, A. Tang, G. Cloutier, and H. Rivaz, "Displacement tracking techniques in ultrasound elastography: From cross-correlation to deep learning," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 7, pp. 842–871, 2024.
- [8] M. A. Lediju, M. J. Pihl, S. J. Hsu, J. J. Dahl, C. M. Gallippi, and G. E. Trahey, "A motion-based approach to abdominal clutter reduction," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 11, pp. 2437–2449, 2009.
- [9] R. Al Mukaddim, N. H. Meshram, and T. Varghese, "Locally optimized correlation-guided bayesian adaptive regularization for ultrasound strain imaging," *Physics in Medicine & Biology*, vol. 65, no. 6, p. 065008, 2020.
- [10] R. Z. Azar, O. Goksel, and S. E. Salcudean, "Sub-sample displacement estimation from digitized ultrasound rf signals using multi-dimensional polynomial fitting of the cross-correlation function," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, no. 11, pp. 2403–2420, 2010.
- [11] J. Jiang and T. Hall, "A coupled subsample displacement estimation method for ultrasound-based strain elastography," *Physics in Medicine and Biology*, vol. 60, pp. 8347–8364, 10 2015.
- [12] P. Liu and D. Liu, "Delay estimation using instantaneous frequency and phase difference—simulation study," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 3, pp. 394–407, 2016.
- [13] L. Yuan and P. C. Pedersen, "Analytical phase-tracking-based strain estimation for ultrasound elasticity," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, no. 1, pp. 185–207, 2015.
- [14] M. Ashikuzzaman, J. Huang, S. Bonwit, A. Etemadimanesh, P. Raghavan, and M. A. L. Bell, "Deep learning-based displacement tracking for post-stroke myofascial shear strain quantification," in *IEEE International Symposium on Biomedical Imaging 2024*, 2024.
- [15] F. Viola and W. F. Walker, "A comparison of the performance of time-delay estimators in medical ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 50, no. 4, pp. 392–401, 2003.
- [16] W. F. Walker and G. E. Trahey, "A fundamental limit on delay estimation using partially correlated speckle signals," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 42, no. 2, pp. 301–308, 1995.
- [17] A. K. Z. Tehrani, M. Ashikuzzaman, and H. Rivaz, "Lateral strain imaging using self-supervised and physically inspired constraints in unsupervised regularized elastography," *IEEE Transactions on Medical Imaging*, vol. 42, no. 5, pp. 1462–1471, 2023.
- [18] M. G. Kibria and H. Rivaz, "Gluenet: Ultrasound elastography using convolutional neural network," in *MICCAI POCUS*, 2018.
- [19] R. Delaunay, Y. Hu, and T. Vercauteren, "An unsupervised learning approach to ultrasound strain elastography with spatio-temporal consistency," *Physics in Medicine & Biology*, vol. 66, no. 17, p. 175031, 2021.
- [20] M. Ashikuzzaman and H. Rivaz, "Incorporating multiple observations in global ultrasound elastography," in *IEEE EMBC*, 2020.
- [21] M. Honarvar, R. Sahebjavaheer, S. Salcudean, and R. Rohling, "Sparsity regularization in dynamic elastography," *Physics in Medicine & Biology*, vol. 57, no. 19, p. 5909, 2012.
- [22] M. Ashikuzzaman *et al.*, "Combining first- and second-order continuity constraints in ultrasound elastography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 68, no. 7, pp. 2407–2418, 2021.
- [23] M. Ashikuzzaman and H. Rivaz, "Second-order ultrasound elastography with L_1 -norm spatial regularization," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 3, pp. 1008–1019, 2022.
- [24] A. K. Z. Tehrani and H. Rivaz, "Displacement estimation in ultrasound elastography using pyramidal convolutional neural network," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, pp. 1–1, 2020.
- [25] J. Jensen, "Field: A program for simulating ultrasound systems," *Medical and Biological Engineering and Computing*, vol. 34, pp. 351–352, 1996.
- [26] J. A. Jensen, "Simulation of advanced ultrasound systems using Field II," in *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro*. IEEE, 2004, pp. 636–639.
- [27] A. K. Tehrani *et al.*, "Bi-directional semi-supervised training of convolutional neural networks for ultrasound elastography displacement estimation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2022.
- [28] M. Mirzaei, A. Asif, and H. Rivaz, "Combining total variation regularization with window-based time delay estimation in ultrasound elastography," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2744–2754, 2019.
- [29] A. K. Z. Tehrani, M. Mirzaei, and H. Rivaz, "Semi-supervised training of optical flow convolutional neural networks in ultrasound elastography," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Cham: Springer International Publishing, 2020, pp. 504–513.
- [30] H. Rivaz *et al.*, "Ultrasound elastography: A dynamic programming approach," *IEEE Transactions on Medical Imaging*, vol. 27, no. 10, pp. 1373–1377, 2008.
- [31] M. Ashikuzzaman, A. K. Tehrani, and H. Rivaz, "Exploiting mechanics-based priors for lateral displacement estimation in ultrasound elastography," *IEEE Transactions on Medical Imaging*, vol. 42, no. 11, pp. 3307–3322, 2023.
- [32] Z. C. Lipton, C. Elkan, and B. Naryanaswamy, "Optimal thresholding of classifiers to maximize f1 measure," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2014, pp. 225–239.
- [33] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, October 2021, pp. 10 012–10 022.
- [34] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang, A. Xiao, C. Xu, Y. Xu, Z. Yang, Y. Zhang, and D. Tao, "A survey on vision transformer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 87–110, 2023.