

Tactile Zero-Shot Sensing of Breast Tumors: Recognition of Human Data from Phantom Data

Arpita Das¹, Dina Caroline², and Chang-Hee Won³

^{1,3}Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA-19122, USA

²Radiology Department, Temple University Hospital, Philadelphia, PA-19140, USA

¹arpita.das0001@temple.edu, ²dina.caroline@tuhs.temple.edu, ³cwon@temple.edu

Abstract— One of the biggest challenges of biomedical research is not having enough human data. This paper explored Tactile Attributes-based Zero-Shot Learning to recognize human data results from phantom data. Using this machine learning method, we trained a Multilayer Perceptron and a Support Vector Machine with the phantom data and estimated the malignancy level of the human data. The method showed that despite using phantom data for training, we successfully applied the knowledge to human data, achieving a malignancy level calculation accuracy of 74.42%, a sensitivity of 73.68%, and a specificity of 75.00%.

Keywords— *Zero-shot Learning; Machine Learning; Tumor Malignancy*

I. INTRODUCTION

Breast cancer is a leading cause of death among women worldwide [1, 2], accounting for about 30% of newly diagnosed cancers in women in 2022 [3]. Breast tumors form when breast cells multiply uncontrollably, which can lead to breast cancer, either tumorous or non-tumorous. Traditional detection methods rely on imaging modalities like ultrasound [4], MRI [5, 6], and mammograms [7-10], focusing on human patient data. However, acquiring human data is often complex, costly, and time-consuming. That is why, our research group [11, 12] has used breast phantom data to detect tumors with tactile and hyperspectral modalities. These phantoms, made from polydimethylsiloxane (PDMS) and Polyvinyl chloride (PVC), mimic human breast tissue (Fig. 1(b)). The group also employed tactile feature estimation [13] and deep neural networks [14] to detect tumors, testing the methods on both phantom and human data. These works have prompted us to evaluate the reliability of phantom results for human data because real-world conditions may differ from phantom scenarios.

To address the reliability of phantom data and assess if it can be an optimal predictor for human data, we decided to explore machine learning methods to evaluate potential correlations between phantom and human data. Previous studies [15, 16] demonstrated that semantic attributes-based Zero-Shot Learning (ZSL) can effectively classify unseen objects in the test set. This inspired us to explore semantic-based relationships between human and phantom data. Semantic attributes-based descriptions [15-19], particularly tactile attributes [16-19], have

This work was supported in part by the National Science Foundation's Grant ECCS-2114675 and the Office of Vice President for Research, Temple University. This research includes calculations carried out on HPC resources supported in part by the National Science Foundation through major research instrumentation grant number 1625061 and by the US Army Research Laboratory under contract number W911NF-16-2-0189.

been successful in object recognition. Motivated by their works, we decided to utilize Tactile Attributes-based Zero-Shot Learning for tactile sensing of breast tumors.

Zero-shot learning (ZSL) enables the identification of new objects without prior training data on those objects. It leverages data from previously encountered objects to recognize new ones by using attribute-based descriptions. Our approach, Tactile Attributes-based Zero-Shot Learning, used tactile features and descriptions of tactile attributes to classify the attributes. This method allowed the system to recognize unseen samples, such as malignant or benign tumors, based on their tactile attributes using the direct attributes prediction (DAP) method of ZSL [15, 19]. For breast tumors, these attributes can be size (e.g., small), depth (e.g., deep), and elasticity (e.g., hard). Previous works have applied ZSL to histopathology images [20], medical image artifact reduction [21], and medical image classification [22]. However, using ZSL to relate phantom data with human data remains unexplored. This paper's key contributions are: (1) mapping phantom data knowledge to human data using tactile attributes (small, deep, hard, etc.), and (2) utilizing Tactile Attributes-based Zero-Shot Learning to recognize human data (malignant or benign) based on phantom data knowledge.

II. METHODOLOGY

A. Tactile Sensing System

Using the tactile imaging sensor developed by our research group [11, 23], tactile images of phantom and human data were captured (Fig. 1). It features a soft, transparent probe measuring 20 mm × 23 mm × 14 mm with an elastic modulus of 27.16 ± 0.57 kPa. It includes four 1500 mcd white LEDs, an 18MP CMOS camera, and a force sensor (Digikey) measuring forces from 0 to 50 N. The tactile imaging sensors use total internal reflection to capture images. Without pressure on the sensing tip, the camera records no light. When pressure is applied, the incident light scatters, forming a tactile image captured by the camera (Fig. 1(a)).

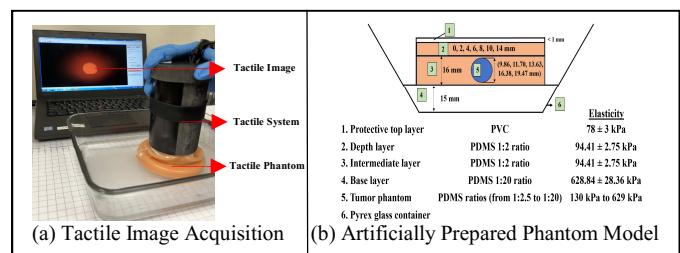


Fig. 1. Tactile Sensing System

B. Tactile Attributes-Based Zero-Shot Learning

We used tactile binary attributes to facilitate knowledge transfer and then applied DAP. We used phantom data as train sets ($D_{train} \subset X \times Y$) and human data as test sets ($D_{test} \subset X \times Z$). Size, deformation index, and risk score from [13] have been used as tactile features space (X). Additionally, depth was calculated using the same method as the size calculation described in [13]. Train data, $Y = \{y_1, \dots, y_N\}$ and test data, $Z = \{z_1, \dots, z_L\}$ are two disjoint classes' data.

Step 01 (Forming Tactile Attributes Matrix): Each sample was characterized by a set of tactile attributes $A = \{a_1, \dots, a_M\}$ (i.e., $A = \{\text{small, large, soft, hard, shallow, deep}\}$). Attributes were categorized based on their true size, depth, and stiffness (or, elasticity). Tumors with a size of 12 mm or less were considered small, and large otherwise. Tumors with a depth of 4 mm or less were deemed shallow, and deep otherwise. For stiffness, tumors with a stiffness of 330 kPa or less were identified as soft, and hard otherwise. These attributes formed the tactile attribute matrix K (Fig. 2), representing the attribute space. In Fig. 2, row i corresponds to the tactile attributes-based description $\mathbf{a}^{s_i} = [a_1^{s_i}, \dots, a_M^{s_i}]$ of a sample $s_i \in S$, where, $S = Y \cup Z$. For a specific sample, the value of a_m , $m = 1, \dots, M$, was a binary attribute such that $a_m = 1$ if the attribute a_m was present in that sample and 0 otherwise.

		A							
		a_1	\dots	\dots	a_M				
Train Samples (Phantom data) Y	y_1	0	1	\dots	1				
	y_N	\vdots	\vdots	\vdots	\vdots				
Test Samples (Unseen Human data) Z	z_1	1	0	\dots	0				
	z_L	0	1	\dots	0				
		\vdots	\vdots	\vdots	\vdots				
		1	0	\vdots	1				

Fig. 2. Tactile Attribute Matrix K

Step 02 (Computing Attribute Posteriors): Tactile attributes to be used were small, large, soft, hard, shallow, and deep. The train set and their tactile attributes-based descriptions $\{\mathbf{a}^{y_n}, y_n \in Y\}$, for $n = 1, 2, \dots, N$, were used to learn a mapping from x to A , for any $x \in X$ (Fig. 3).

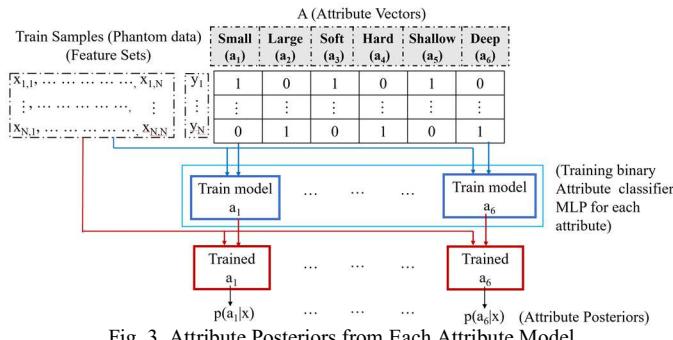


Fig. 3. Attribute Posteriors from Each Attribute Model

In Fig. 3, each attribute a_m used the tactile feature space to train an independent probabilistic binary classifier. For each classifier, the corresponding attribute column is the label. When a sample's features x were input into a_m classifier, it estimated attribute posterior $p(a_m|x)$, the probability of a_m being present in x .

For each attribute a_m , we trained a binary classifier Multilayer Perceptron (MLP) using the training dataset $D_{train}^m = \{(x, a_m^y) \text{ such that } (x, y) \in D_{train}\}$. We chose MLP for its ability to learn hierarchical feature representations and capture complex tactile data patterns [24]. Upon examining the tactile attribute matrix, we observed a notable imbalance in the presence versus absence counts for each attribute, leading to a skewed dataset. To address this imbalance and reduce the risk of model overfitting, we under-sampled the minority classes by randomly removing pairs $(x, a_m^{y_k})$ until the counts of class 0 and class 1 for each attribute were equalized. The final balanced dataset was used for training the attribute classifier.

Multilayer Perceptron: We trained the MLP classifier with two hidden layers, using the 'logistic' activation function and 'adam' optimizer, selected for optimal results.

Step 03 (Direct Attributes Prediction): For a test sample $x \in X$, the attribute posterior $p(a_m|x)$ for each attribute a_m was used by the DAP model along with the tactile attribute-based descriptions of test samples $\{\mathbf{a}^{z_l}, z_l \in Z\}$, for $l = 1, 2, \dots, L$, to infer the final classification z_x . x was classified as the sample having the highest posterior (Fig. 4).

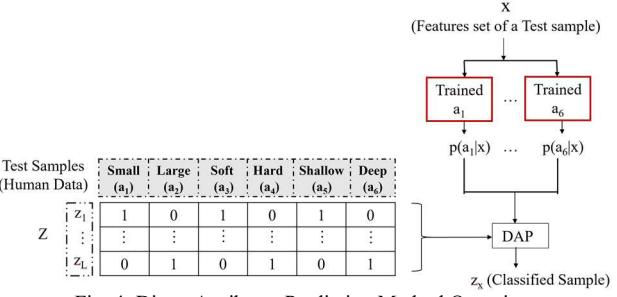


Fig. 4. Direct Attributes Prediction Method Overview

If $p(a_m = 1|x) < 0.5$, the attribute might be identified as absent from x by the model despite being truly present. To reduce misclassification effects, we replaced the attribute posterior with a uniform distribution when the attribute was identified as absent. Thus, $p(a_m = 1|x)$ had been calculated as: $p(a_m = 1|x) = \max(0.5, p(a_m = 1|x))$. DAP used Bayes' rule to compute the final sample posterior as follows [15]:

$$p(z_l|x) = \frac{p(z_l)}{p(a^{z_l})} \prod_{m=1}^M p(a_m^{z_l}|x); l = 1, 2, \dots, L \quad (1)$$

where test class priors were considered uniform, i.e., $p(z_l) = 0.5$. To facilitate zero-shot sensing, we computed attribute priors, $p(a^{z_l})$, based on its presence in the test set: $p(a_m = 1) = \frac{1}{L} \sum_{l=1}^L a_m^{z_l}$. Since ZSL assumes attribute independence [15, 19], $p(a^{z_l}|x)$ had been computed as $\prod_{m=1}^M p(a_m^{z_l}|x)$. Here, $p(a_m^{z_l}|x)$ came from each a_m trained model and $p(z_l|x)$ was the probability of x being from a particular sample z_l . Finally, the mapping function $f: X \rightarrow Z$ recognized sample z_x , as the one having the highest posterior:

$$z_x = f(x) = \operatorname{argmax}_{z_l \in Z} p(z_l|x) \quad (2)$$

Eq. (2) is the maximum a posteriori estimator (MAP) method.

Step 04 (Computing Distance Metrics and Malignancy Index): Each a_m classifier's performance was analyzed on the

test set Z by the L_1 distance between binary attribute labels and their posteriors [15]:

$$D_m = \left\| a_m - \frac{1}{N_A} \sum_{x \in D_{test}} p(a_m = 1 | x) \right\|_1 \quad (3)$$

In our case, each sample can be categorized across three binary dimensions (small/large, soft/hard, and shallow/deep), resulting in eight possible attribute combinations (Table I). Eq. (3) was computed on each such sequence of Z by averaging $p(a_m = 1 | x)$ across the number of appearances of the sequences, N_A . If x was collected from a sample, $z_l \in Z$, then $p(a_m = 1 | x)$ should be close to $a_m^{z_l}$ since $p(a_m = 1 | x) \approx 1$ if $a_m^{z_l} = 1$, and $p(a_m = 1 | x) \approx 0$ otherwise. D_{test} denotes the set of sequences.

Next, we transformed $p(a_m | x)$ into the malignancy index (MI) for each human data as follows:

$$MI = p(a_2 | x)w_1 + p(a_4 | x)w_2 + p(a_6 | x)w_3 \quad (4)$$

A tumor can be either small or large, soft or hard, and shallow or deep. In Eq. (4), we considered only the likelihood of large, hard, and deep attributes from each pair (small/large, soft/hard, and shallow/deep) of a sample because these attributes are mostly responsible for the malignancy [25].

III. RESULTS AND DISCUSSIONS

The explored method was implemented in v6.4.12 of Jupyter Notebook in a Python 3.9.13 environment, utilizing the Scikit-learn library v1.2.2 for our machine-learning solutions in a local laptop. Image processing algorithms were executed on Temple University's High-Performance Computing cluster named 'Compute', which had 20 cores and 240 GB of RAM.

Since we used the tactile attribute-based learning technique, we first assessed the performance of attribute classifiers. To transfer knowledge from phantom data (train set) to human data (novel test set), we applied Eq. (1) and Eq. (2) to determine the most likely sequence of binary attributes for each human data among the eight possible combinations in Table I, which were previously obtained from the phantom data. Our objective was to differentiate malignant from benign human samples using phantom data knowledge, focusing on these unique sequences to analyze the attribute classifiers' performance.

TABLE I. THE DISTANCE BETWEEN ATTRIBUTES BINARY LABELS AND THEIR POSTERIORS (SHOWN FROM THE MLP RESULTS)

Class	Sequences	Small	Large	Soft	Hard	Shallow	Deep	$D_m \leq 0.5$	$0.5 < D_m < 1$
Benign	101010	0.371394	0.258769	0.670329	0.397853	0.361421	0.36144		
Benign	101001	0.369663	0.435628	0.369747	0.696989	0.489205	0.568995		
Benign	011010	0.166551	0.565732	0.475415	0.478415	0.476771	0.364364		
Benign	011001	0.406991	0.430195	0.427099	0.417013	0.324746	0.424172		
Malignant	100110	0.369656	0.398011	0.397657	0.492765	0.396515	0.397421		
Malignant	100101	0.369663	0.369747	0.630279	0.407005	0.524989	0.323893		
Malignant	010110	0.316749	0.207613	0.432782	0.453771	0.300089	0.267012		
Malignant	010101	0.589123	0.397519	0.323898	0.624995	0.313679	0.443899		

Table I implemented L_1 distance using Eq. (3). It illustrates that most distances are below 0.5 (white cells), demonstrating how well the attribute classifiers can correctly classify an attribute. A smaller distance signifies better classification. Therefore, we considered $D_m \leq 0.5$ was a correct classification and a distance above 0.5 was incorrect. This approach yielded an 83.33% accuracy for the attribute classifiers.

Next, we evaluated the performance of DAP in Eq. (1) on the validation set and the 43 human data (test set). We implemented our MLP and also the SVM as described in [15] and compared their classification accuracies. Using 80% of the phantom data for training and the remaining 20% for validation, the validation set achieved an accuracy of 73.79% in the MLP setup and 68.34% for SVM. In the human dataset, the MLP again outperformed the SVM, achieving a recognition accuracy of 65.2%, compared to the SVM's accuracy of 55.8%.

In [13], the breast cancer risk score used 70% weight to the tumor stiffness because stiffness is the most significant factor in detecting malignancy and 30% to the tumor size. We followed a similar logic in Eq. (4) where the 'Hard' attribute was assigned a weight (w_2) of 70%, while the remaining two attributes each received weights (w_1 and w_3) of 15%. This method achieved an accuracy of 74.42% in distinguishing between malignant (high risk) and benign (low risk) cases in Table II, with a sensitivity of 73.68% and a specificity of 75.00%. Misclassified malignancy estimations are highlighted in Table II, considering a malignancy level (ML) of 50% or higher as a high level.

TABLE II. THE MALIGNANCY LEVEL (ML) OF THE HUMAN SAMPLES (MI ≥ 0.5 WAS CONSIDERED HIGH) (SHOWN FROM THE MLP RESULTS) [M = MALIGNANT AND B = BENIGN]

Patient	MI	ML	Biopsy	Patient	MI	ML	Biopsy
1.	0.73	High	M	23.	0.40	Low	B
2.	0.16	Low	B	24.	0.16	Low	M
3.	0.40	Low	B	25.	0.16	Low	B
4.	0.26	Low	B	26.	0.52	High	M
5.	0.10	Low	B	27.	0.71	High	M
6.	0.37	Low	B	28.	0.91	High	M
7.	0.27	Low	M	29.	0.56	High	B
8.	1.00	High	M	30.	0.29	Low	B
9.	0.60	Low	M	31.	0.16	Low	B
10.	0.26	Low	M	32.	0.59	High	M
11.	0.18	Low	B	33.	0.46	Low	B
12.	0.16	Low	B	34.	0.47	Low	B
13.	0.20	High	B	35.	0.44	Low	M
14.	0.59	High	M	36.	0.43	Low	B
15.	0.99	High	M	37.	0.40	Low	B
16.	0.46	Low	B	38.	0.39	Low	B
17.	0.80	High	B	39.	0.52	High	B
18.	0.90	High	M	40.	0.61	High	M
19.	1.00	High	M	41.	0.51	High	B
20.	0.90	High	B	42.	0.71	High	M
21.	0.80	High	M	43.	0.61	High	M
22.	0.40	Low	B				

IV. CONCLUSIONS

We utilized Tactile Attributes-based Zero-Shot Learning to transfer knowledge from phantom data to unseen human samples. Two binary classifiers, MLP and SVM, were trained for each attribute and compared their recognition accuracies. We calculated the L_1 distance for each unique attribute sequence using the attribute posteriors to assess the classifiers' performance and achieved 83.33% accuracy. We conclude that even though phantom data was used to train the data, we could relate the phantom knowledge with human data with an accuracy of 74.42%, a sensitivity of 73.68%, and a specificity of 75.00% malignancy level calculation.

ACKNOWLEDGMENT

The authors thank Dr. Vira Oleksyuk for helping in the phantom and human data acquisition.

V. REFERENCES

- [1] J. Ferlay, H. R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008," *Int. J. Cancer*, vol. 127, no. 12, pp. 2893–2917, 2010, doi: 10.1002/ijc.25516.
- [2] World Cancer Research Fund, "Cancer Facts and Figures 2021," *World Cancer Research Fund International*, pp. 1–4, 2021. [Online]. Available: <http://www.wcrf.org/int/cancer-facts-figures/worldwide-data>.
- [3] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA. Cancer J. Clin.*, vol. 72, no. 1, pp. 7–33, 2022, doi: 10.3322/caac.21708.
- [4] Y. Luo, Q. Huang, and X. Li, "Segmentation information with attention integration for classification of breast tumor in ultrasound image," *Pattern Recognit.*, vol. 124, Apr. 2022, doi: 10.1016/j.patcog.2021.108427.
- [5] H. J. Kim *et al.*, "High-resolution diffusion-weighted MRI plus mammography for detecting clinically occult breast cancers in women with dense breasts," *Eur. J. Radiol.*, vol. 175, Jun. 2024, doi: 10.1016/j.ejrad.2024.111440.
- [6] F. E. Zerrad *et al.*, "Microwave Imaging Approach for Breast Cancer Detection Using a Tapered Slot Antenna Loaded with Parasitic Components," *Materials (Basel.)*, vol. 16, no. 4, 2023, doi: 10.3390/ma16041496.
- [7] L. Singh and A. Alam, "An efficient hybrid methodology for an early detection of breast cancer in digital mammograms," *J. Ambient Intell. Humaniz. Comput.*, vol. 15, no. 1, pp. 337–360, Jan. 2024, doi: 10.1007/s12652-022-03895-w.
- [8] S. S. Boudouh and M. Bouakkaz, "New enhanced breast tumor detection approach in mammogram scans based on pre-processing and deep transfer learning techniques," *Multimed. Tools Appl.*, vol. 83, no. 9, pp. 27357–27378, Mar. 2024, doi: 10.1007/s11042-023-16545-w.
- [9] F. Yan, H. Huang, W. Pedrycz, and K. Hirota, "Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms," *Expert Syst. Appl.*, vol. 227, no. April, p. 120282, 2023, doi: 10.1016/j.eswa.2023.120282.
- [10] J. G. Melehoodappattu, A. S. Dhas, B. K. Kandathil, and K. S. Adarsh, "Breast cancer detection in mammogram: combining modified CNN and texture feature based approach," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 9, pp. 11397–11406, 2023, doi: 10.1007/s12652-022-03713-3.
- [11] A. Sahu, F. Saleheen, V. Oleksyuk, Y. Chen, and C. H. Won, "Tactile and hyperspectral imaging sensors for mammary tumor characterization," *Proc. IEEE Sensors*, pp. 1–4, 2013, doi: 10.1109/ICSENS.2013.6688136.
- [12] F. Saleheen, V. Oleksyuk, A. Sahu, and C.-H. Won, "Non-invasive mechanical properties estimation of embedded objects using tactile imaging sensor," in *Smart Biomedical and Physiological Sensor Technology X*, SPIE, May 2013, p. 87190K. doi: 10.1117/12.2015803.
- [13] V. Oleksyuk, R. Rajan, F. Saleheen, D. F. Caroline, S. Pascarella, and C. H. Won, "Risk score based pre-screening of breast tumor using compression induced sensing system," *IEEE Sens. J.*, vol. 18, no. 10, pp. 4038–4045, 2018, doi: 10.1109/JSEN.2018.2817883.
- [14] V. Oleksyuk, N. Rahman, and C. H. Won, "Tactile Sensing System and Convolutional Neural Network for Mechanical Property Classification," *IEEE Sensors Lett.*, vol. 7, no. 10, pp. 1–4, 2023, doi: 10.1109/LSENS.2023.3310356.
- [15] Z. Abderrahmane, G. Ganesh, A. Crosnier, and A. Cherubini, "Haptic Zero-Shot Learning: Recognition of objects never touched before," *Rob. Auton. Syst.*, vol. 105, pp. 11–25, Jul. 2018, doi: 10.1016/j.robot.2018.03.002.
- [16] G. Cao, J. Jiang, D. Bollegala, M. Li, and S. Luo, "Multimodal zero-shot learning for tactile texture recognition," *Rob. Auton. Syst.*, vol. 176, Jun. 2024, doi: 10.1016/j.robot.2024.104688.
- [17] H. Liu, F. Sun, B. Fang, and D. Guo, "Cross-Modal Zero-Shot-Learning for Tactile Object Recognition," *IEEE Trans. Syst. Man. Cybern. Syst.*, vol. 50, no. 7, pp. 2466–2474, Jul. 2020, doi: 10.1109/TSMC.2018.2818184.
- [18] F. Wang, H. Liu, F. Sun and H. Pan, "Fabric recognition using zero-shot learning," in *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 645–653, Dec. 2019, doi: 10.26599/TST.2018.9010095.
- [19] Z. Abderrahmane, G. Ganesh, A. Crosnier and A. Cherubini, "Visuo-Tactile Recognition of Daily-Life Objects Never Seen or Touched Before," *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Singapore, 2018, pp. 1765–1770, doi: 10.1109/ICARCV.2018.8581230.
- [20] M. Y. Lu *et al.*, "Visual Language Pretrained Multiple Instance Zero-Shot Transfer for Histopathology Images," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 2023, pp. 19764–19775, doi: 10.1109/CVPR52729.2023.01893.
- [21] Y. -J. Chen *et al.*, "Zero-Shot Medical Image Artifact Reduction," *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA, 2020, pp. 862–866, doi: 10.1109/ISBI45749.2020.9098566.
- [22] D. Mahapatra, B. Bozorgtabar and Z. Ge, "Medical Image Classification Using Generalized Zero Shot Learning," *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada, 2021, pp. 3337–3346, doi: 10.1109/ICCVW54120.2021.00373.
- [23] J. H. Lee and C. H. Won, "High-resolution tactile imaging sensor using total internal reflection and nonrigid pattern matching algorithm," *IEEE Sens. J.*, vol. 11, no. 9, pp. 2084–2093, 2011, doi: 10.1109/JSEN.2011.2109038.
- [24] A. Goshtasbi *et al.*, "AI-Infused Soft Fluidic Tactile Sensing," *2024 IEEE 7th Int. Conf. Soft Robot.*, pp. 1095–1100, 2024, doi: 10.1109/robosoft60065.2024.10522049.
- [25] D. C. Teichgraeber, M. S. Guirguis, and G. J. Whitman, "Breast cancer staging: Updates in the AJCC cancer staging manual, 8th edition, and current challenges for radiologists, from the AJR special series on cancer staging," *Am. J. Roentgenol.*, vol. 217, no. 2, pp. 278–290, 2021, doi: 10.2214/AJR.20.25223.