Deep Learning for Tumor Localization with Depth Estimation: A Minimally Invasive Robotics-Assisted Approach

Ali Talasaz
Department of Electrical & Computer Science,
Florida Atlantic University,
Boca Raton, USA
atalasaz2019@fau.edu

Abstract—This work is aimed at deploying deep learning models to characterize tissue stiffness properties during telerobotic palpation and localization of tissue abnormality while estimating its depth. The method relies on using a minimally invasive probe with a rigidly mounted tactile sensor at the tip to capture the force distribution map and the indentation depth for each tactile element, thereby generating a stiffness map for the palpated tissue. The probe is attached to a Mitsubishi PA10 robot with hybrid impedance control architecture controlled by a haptic device which enables the user to telerobatically palpate the remote tissue and semi-autonomously obtain the required information from the tissue. When data are collected, a convolutional neural network (CNN) is utilized for tumor classification and an artificial neural network (ANN) is used to estimate the depth of the tumor within the tissue. The proposed method is verified by classifying tumor phantoms using CNN with an accuracy of 95.24% for stiffness and using ANN with an accuracy of 82.14% for depth.

Index Terms—Tactile sensing, tumor localization, deep learning, convolutional neural network, robotic-assisted minimally invasive surgery

I. Introduction

Soft-tissue palpation is one of the most important parts of a surgical procedure which is commonly used for tumor localization as the first step taken in cancer treatment. Since the stiffness of a tumor is higher than that of healthy tissue, it can be distinguished as a hard nodule during remote palpation. Considerable work has been done on the problem of providing haptic information for tumor localization in MIS soft-tissue palpation [1]. Among them, tactile feedback is more of interest for this application because of the detailed information it gives about the palpated area [2], [3]. However, it has been shown that tactile sensing alone may not be sufficient for successful tumor localization [4]. In particular, the determination of tumor depth is a key prerequisite for tumor resection in MIS for which tactile information alone would not be sufficient. Talasaz et al. [5] proposed a new approach for integrating force control with tactile sensing and position control to characterize tissue stiffness and to localize tumors while obtaining depth estimation for tumors. However, the main challenging problem is that the detection and localization of a tumor from an obtained image could be very subjective depending on the level of experience of the surgeon.

This paper addresses the problem mentioned above by applying Convolutional Neural Networks (CNN) for tumor classification as well as a multi-layer neural network for depth estimation while utilizing a stiffness map of the palpated tissue integrated with the exploration force during robot-assisted tumor localization.

Deep learning algorithms are widely used in healthcare domain especially for medical image localization, segmentation and classifications [6]–[8]. Oleksyuk et al. [6] applied a CNN model for mechanical property classification of breast tumor with an experimental tactile sensing system developed for breast cancer applications. Deep learning for automatic liver tumor diagnosis has also attracted considerable interest [9]. In the context of tumor depth estimation, Zhou et al. [10] developed several deep learning methods to recognize the depth of hard inclusions in soft tissue using ordinal classification for robotic palpation (not in the context of MIS). Similarly Xiao et al. [11] developed an experimental setup, not suitable for minimally invasive surgery, and proposed a tactile sensing-based deep recurrent neural network (DRNN) with long short-term memory (LSTM) architecture to improve the accuracy of the detection and depth estimation of tumors embedded in soft tissue.

The contribution of this paper is a novel approach for tumor localization in robotics-assisted minimally invasive surgery (RAMIS). This work characterizes tissue properties telerobotically while the tactile sensing instrument (TSI) [12] is inserted into the patient's body in a minimally invasive manner to collect position, tactile, and force data from sensor-tissue interaction without friction at the trocar interfering with the measurements. Then by utilizing deep learning, tissue abnormality is localized using a CNN model. Two key attributes are also established to classify tumor depth in underlying tissue using a multi-layer ANN model.

The outline of this paper is as follows. Section II introduces the setup and the control algorithm used for telerobotic palpation. The CNN model used for tumor classification as well the ANN model used for depth estimation are discussed in Section III. The experimental results are presented and discussed in Section IV. Finally, concluding remarks are given in Section V.

II. RAMIS Palpation Setup and Control Algorithm

A. Telerobotic Palpation Setup

Fig. 1 shows the leader-follower teleoperation setup which consists of a Mitsubishi PA10-7C robot as the follower and a 7 Degrees-of-Freedom (DOFs) Haptic Wand [13] as the leader interface. At the robot end effector, a tactile sensing instrument (TSI), shown in Fig. 2, is used

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00087

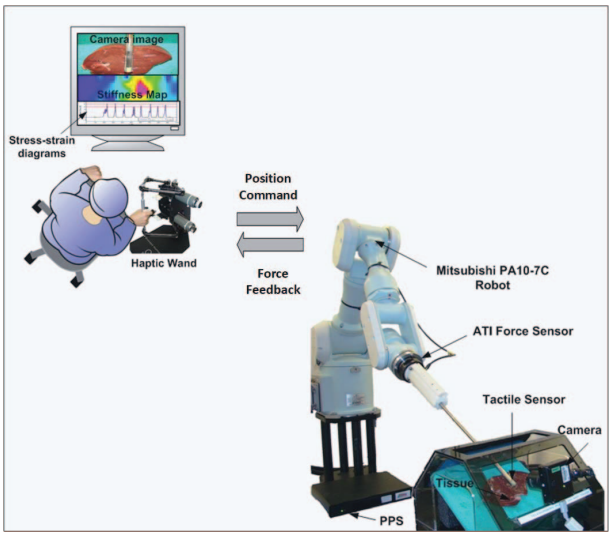

Fig. 1: Leader-follower robotic setup palpating a tissue [4].

to measure the pressure distribution over tissue during tumor localization in soft-tissue palpation. The sensor used in this research is a two-dimensional array (15×4) of pressure sensing capacitive elements in a thin and continuous sheet [14] developed for measuring the tactile pressure distribution between objects in direct physical contact (Fig. 2). Each element is $2\,mm \times 2\,mm$ and the total size of the sensor is $30\,mm \times 8\,mm$. This sensor is attached to a probe, with its shaft length of $385\,mm$ and shaft diameter of $10\,mm$, that is suitable for use in MIS without interfering with trocar-palpator friction. The tactile data obtained from the sensor contains information about the magnitudes, distributions and locations of forces. It also provides information about the contact area and the pressure distribution over it. Moreover, this sensor is also used as the force sensor, by adding up the force measurements of all the elements, to measure the interaction force applied by the TSI on the tissue in the normal direction to the tissue plane (palpation direction).

B. Control Algorithm

As stated earlier, tactile feedback is of interest for localizing tissue abnormality in soft-tissue palpation because of the detailed information it gives about the palpated area. To increase the success rate in tactile sensing tumor

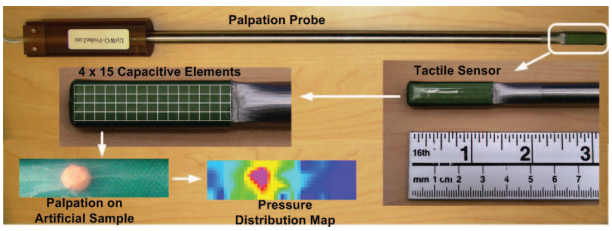

Fig. 2: Palpation probe (TSI).

Algorithm 1 RAMIS Tumor Localization Algorithm

```
Move the TSI to the starting position P_0
while ~TaskDone do
    Switch from position control to force control
    F_d ← F_min
    while F_d > F_max do
        Activate autonomous force control
        if RMSE(F − F_d) < ε then
            Acquire the pressure and stiffness map
            Compute the indentation depth
            F_d ← F_d + δF
        end if
    end while
    Switch from force control to position control
    Move to the next area of interest for palpation
end while
```

localization, Talasaz et al. [5] developed a new approach for integrating force control with tactile sensing and position control to characterize tissue stiffness and to localize tumors while obtaining depth estimation for tumors (The approach has been detailed out in the previous work [5]). In the context of telerobotic palpation, it has been shown that the operator can get some feeling about the indentation depth by the force reflected to their hand through the haptic device. However, it might be difficult for the operator to discern how deep the TSI can palpate tissue based on direct force feedback. Besides, many haptic devices have some limitations on the maximum force that they can reflect to the operator's hand and it might be required to apply higher force if the tissue has higher stiffness and/or if the tumor is located deep in the tissue. The aforementioned problems were looked into in detail in [5] by developing a semi-autonomous force control approach that applies different levels of exploration force consistently on tissue and capturing tissue-sensor interaction data. Algorithm 1 presents the Pseudo-Code of the control algorithm. With this approach, the user controls the probe remotely to scan the tissue surface for tumor localization while the robot controls the normal exploration force autonomously from a starting force to a maximum desirable force with a fixed constant increment.

III. A Deep Learning Algorithm for Tumor Localization

As mentioned earlier, CNN models are very effective tools for medical image classification that can be employed to identify tumors from the tactile images obtained using the TSI during robot-assisted tumor localization. CNN is composed of a set of convolutional layers with corresponding activation functions, max-pooling layers, a flatten layer to provide inputs from a reduced sized image, and fully connected layers with a Softmax function for the final classification decision. Fig. 3 shows the schematic diagram of the CNN model and Table I shows the configuration of the CNN model used for tumor classification in this

TABLE I: The architecture of the CNN model used for tumor classification.

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 15, 4, 64)	256
batch_normalization_9 (BatchNormalization)	(None, 15, 4, 64)	256
activation_9 (Activation)	(None, 15, 4, 64)	0
conv2d_7 (Conv2D)	(None, 15, 4, 64)	16384
batch_normalization_10 (BatchNormalization)	(None, 15, 4, 64)	256
activation_10 (Activation)	(None, 15, 4, 64)	0
max_pooling2d_3 (MaxPooling2D)	(None, 7, 2, 64)	0
flatten_3 (Flatten)	(None, 896)	0
dense_6 (Dense)	(None, 512)	458752
batch_normalization_11 (BatchNormalization)	(None, 512)	2048
activation_11 (Activation)	(None, 512)	0
dropout_3 (Dropout)	(None, 512)	0
dense_7 (Dense)	(None, 2)	1026

Total params: 478978 (1.83 MB)
Trainable params: 477698 (1.82 MB)
Non-trainable params: 1280 (5.00 KB)

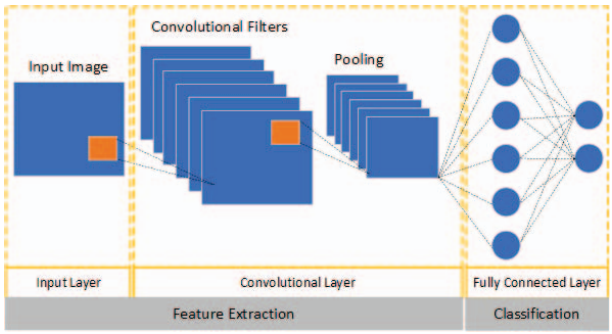

Fig. 3: The schematic diagram of a CNN model with two classes.

work. A CNN sequential model is designed with two convolutional layers, one max-pooling layer and activation functions for the feature learning part.

Each convolutional layer has 64 filters of size 2x2. Adamax optimizer is employed and accuracy is utilized as the metric for the model performance evaluation during training. A Rectified Linear Unit (ReLU) is chosen for the activation function of the convolution layers. Batch normalization is also applied to improve the training performance. The classification part of the model includes a dense layer with a batch normalization layer and ReLU activation functions and a dropout layer of 0.2. The final fully connected layer has a Softmax activation function.

To characterize the depth of a tumor, a multi-layer ANN is employed to classify tumors in three different categories: shallow, deep, and deeper. Fig. 4 presents the schematic

TABLE II: The architecture of the ANN model used for depth estimation.

Layer (type)	Output Shape	Param #
dense_45 (Dense)	(None, 512)	1536
activation_28 (Activation)	(None, 512)	0
dropout_27 (Dropout)	(None, 512)	0
dense_46 (Dense)	(None, 512)	262656
activation_29 (Activation)	(None, 512)	0
dropout_28 (Dropout)	(None, 512)	0
dense_47 (Dense)	(None, 3)	1539

Total params: 265731 (1.01 MB)
Trainable params: 265731 (1.01 MB)
Non-trainable params: 0 (0.00 Byte)

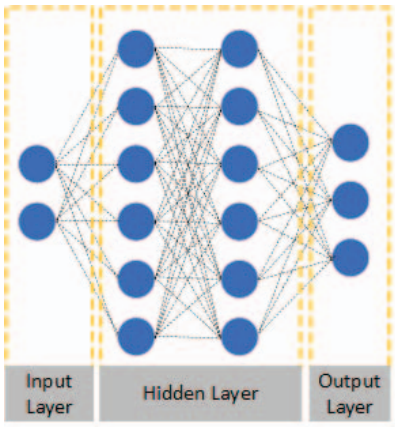

Fig. 4: The schematic diagram of an ANN model with two inputs and three classes

diagram of a multi-layer ANN model with two inputs and three classes. Assuming that the diseased tissue is accessible for the clinician to palpate directly, the way the clinician detects a tumor is to put their finger on the tissue and to apply some force on the tissue to deform the tissue sufficiently to be sensitive to the underlying tumor. If the tumor is located near the surface of the tissue (shallow), they can detect it by their sense of touch (tactile feedback) while a small force is applied on the tissue, for a tumor in the middle of the tissue (deep), more force is required to deform the tissue to let the clinician feel the tumor, and for the tumor near the bottom of the tissue (deeper), a significant amount of force is required for the clinician to detect the tumor. Therefore, in direct palpation, the clinician can get some feeling about the depth of the tumor using the amount of force being applied to the tissue (force feedback), their sense of touch (tactile feedback), and the amount of deformation of the tissue (position data).

Therefore, the exploration force can be selected as the main attribute (feature) for the depth estimator network. As mentioned earlier the tactile feedback can also be incorporated with the tissue deformation information to

compose a stiffness map. The second attribute can be derived from the stiffness images to help in classifying the tumor depth appropriately. To do so, the maximum stiffness of each column of the TSI array ($\max[\sigma_1 ... \sigma_{15}]$) is first found and then the second feature is defined as the average of the four maximum values ($\sum_{n=1}^{4} \sigma_{max_i}/4$). These two features can be utilized to classify the three categories of the tumor depth and thereby estimate how deep the underlying tumor is in the tissue. An ANN sequential model with the configuration shown in Table II is designed to classify tumors in "shallow", "deep" and "deeper" categories. The ANN architecture consists of two hidden layers and the output layer. Two dropout layers of 0.3 as well as L2 regularization are also employed to avoid over-fitting on the training data. Each hidden layer has 512 neurons with ReLU activation functions. Softmax was chosen as the activation function of the output layer.

IV. Experiments

An experimental evaluation was performed to explore the performance of deep learning integrated with tactile imaging in RAMIS tumor localization. In this section, the tissue models used for the experiments are first given. Then, the results and discussion are presented.

A. Experimental Conditions

The tissue used for the experiments was made of silicone gel (Ecoflex0030 with Silicone thinner) with elastic modulus 20 KPa that was experimentally measured by conducting several palpations on tumor-free areas and recording the force data and the amount of indentation. The spherical tumors used for the experiments were made of silicon gel (SORTA-Clear40) and were eight times harder than the tissue phantom. The diameter of the tumors was chosen to be 8 mm - equal to the width of the TSI. For evaluating the performance of the approach for depth estimation, a rectangular shaped tissue phantom with a flat surface was chosen with three tumors embedded in the tissue at different depths: 2 mm(shallow) , 7 mm(deep), and 10 mm(deeper). Fig. 5 shows the dimensions of the tissue phantom, the exact locations of the tumors and the depths at which the tumors were embedded

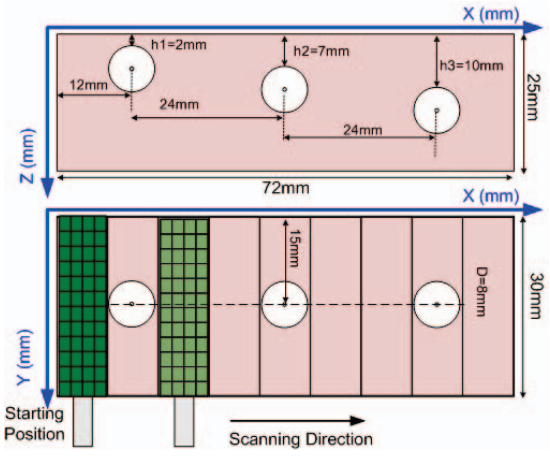

Fig. 5: Tissue model used for the experiments.

To mimic a real palpation task in MIS, a tissue phantom with an uneven surface was also constructed in order to study how the proposed approach works in this scenario (Fig. 6). The average surface height along the Z-axis was 25 mm with a variation of 5 mm. In this case, three tumors were embedded at the same depth (2 mm from the top flat surface). Considering the 5 mm bump, the first two tumors should be characterized as "shallow" and the third tumor as "deep". For ease of use and to provide a wider range of motion during the experiments, the tissues were placed on a table and palpated in the left to right direction. The operator received some visual cues on a monitor connected to a camera overlooking the tissue but it was not possible to discern the location of the lump in the tissue from the camera image. The operator was asked to palpate the tissue using the TSI through the leader-follower teleoperation setup. The operator was then asked to palpate the tissue in a discontinuous mode in different steps; palpating the first area, raising the TSI off the tissue, moving to the next area and repeating this pattern. When the probe is close enough to the tissue surface, the operator is asked to press a switch requesting autonomous force control through the robot in the palpation direction. This ensures the application of a consistent exploration force during palpation which will vary discretely from 1N to 7N of force with the increments of 1N. When autonomous force control was completed for an area, a flag was set informing the operator about the next area for palpation. During the experiments, the real-time stiffness map was computed by measuring the force distribution map of the palpated area along with the indentation depth of the TSI. The total amount of force applied during palpation (the exploration force) can be computed by adding up the force measurements of all the capacitive elements in the TSI.

B. Tumor Classification Results

With the seven different exploration forces applied to the phantom tissue in Fig. 5, a total of 63 images were acquired for the stiffness of the palpated tissue during the first experiment (63x15x4). To create more images for the training of the CNN model, a data augmentation

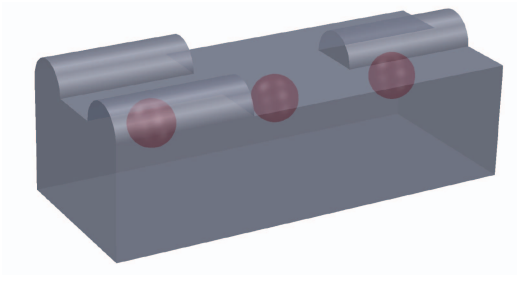

Fig. 6: Tissue model with an uneven surface used for the experiments.

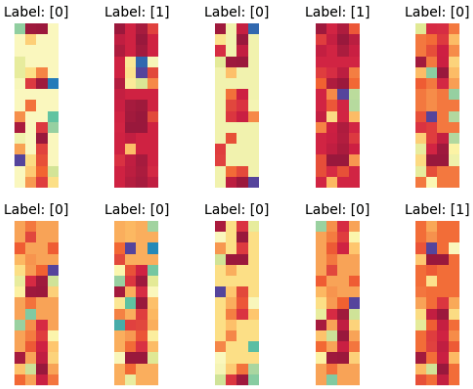

Fig. 7: Sample images used for the CNN model training for tumor classification.

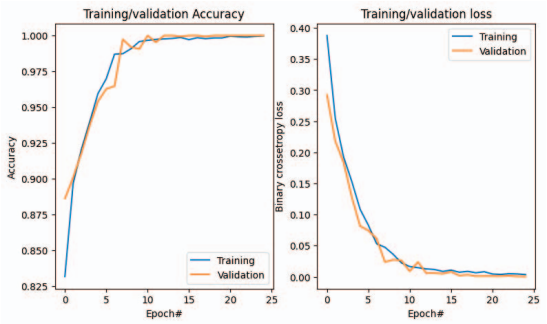

Fig. 8: Training/Validation accuracy and loss for CNN model for tumor classification.

technique is also employed by dividing the TSI array into five 3x4 sub-arrays and considered all possible combinations of the five sub-arrays resulting in a total of 7560 stiffness images (7560x15x4). Among these images the original images (63x15x4) were saved as test data, and the remaining were used for training the network. The images acquired by palpation of the uneven phantom tissue were also used in the test data set to better evaluate the performance of the trained network on a more clinically relevant dataset.

Fig 7 shows 10 random images used for training the CNN model labeled 0 for no tumor and 1 for tumor. Fig 8 also demonstrates the training and validation accuracy as well as the training and validation loss for the 25 epochs used for training the CNN model. As the results show, the CNN model trained perfectly on the augmented dataset. The model was then evaluated on the test data which includes the data for both the phantom tissues shown in Figs. 5 and 6 (which consisted of a total of 126 sample images). The accuracy achieved for the testing data was 95.24% which confirms that the proposed method using the stiffness map was successful in distinguishing a tumor from its surrounding tissue, and that the CNN model developed in this work was also effective in classifying the tissue with tumors with a high degree of accuracy (A

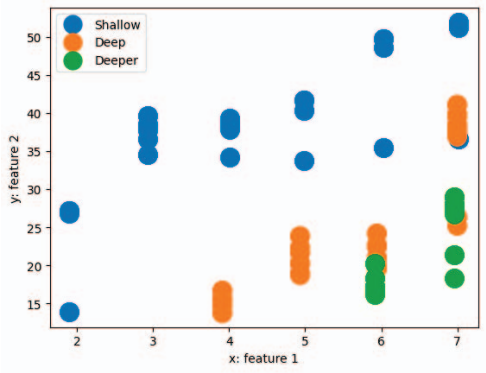

Fig. 9: The feature space for the traing data for depth estimation.

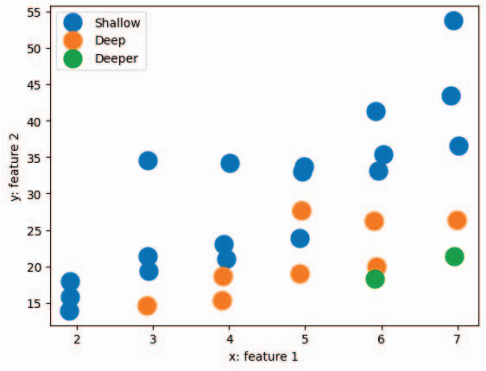

Fig. 10: The feature space for the testing data for depth estimation.

total of 4 false negatives and 2 false positives were also noted).

C. Depth Estimation Results

For the ANN model, the data are first required to be pre-processed and the images showing a tumor corresponding to a certain exploration force to be extracted. The minimum exploration force required to detect a shallow tumor, a deep tumor and a deeper tumor for the phantom tissue used in this work was experimentally measured to be $2N$, $4N$ and $6N$ respectively. After sorting the images from the dataset, a total of 1428 images (out of 7560 images) were obtained for training the ANN model (20% of these were used for validation and the rest for training), and 20 images (out of 126 images) were obtained from the test dataset for testing. Figs. 9 and 10 show the feature space for the training data and testing data respectively. Feature 1 is the exploration force and feature 2 is a representative of the maximum stiffness of the TSI array during palpation as described in Section III. As can be seen from the feature space for the training data, with the defined features, the tumors can be separated into three classes. Due to their different range, data normalization was also applied on both the training and test data to reduce the effect of different scales on the accuracy of the ANN model.

Fig 11 shows 10 random images used for training the ANN model labeled 0 for “shallow” , 1 for “deep” and 2

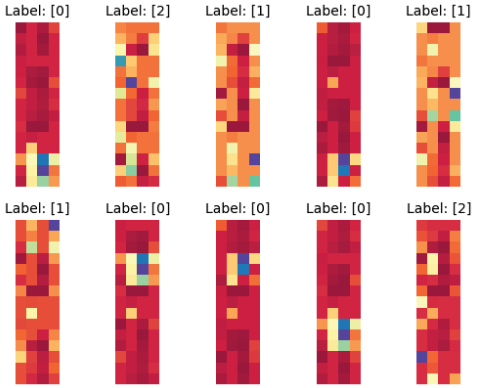

Fig. 11: Sample images used for ANN model training for depth estimation.

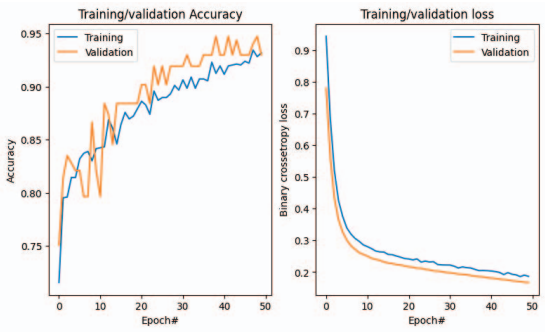

Fig. 12: Training/validation accuracy and loss for the ANN model for depth estimation.

for "deeper". Fig 12 illustrates the training and validation accuracy as well as the training and validation loss for the 50 epochs used for training the ANN model. The model was then evaluated on the test data which includes the data for both phantom tissues shown in Figs. 5 and 6. The accuracy achieved on the testing data was 82.14% showing that the ANN model achieved good performance for characterizing the depth of the tumors embedded in the tissue regardless of the unevenness of the tissue surface (Fig. 6). Sensitivity of 83.33% and specificity of 88.89% were obtained.

V. Conclusion

This work proposed a new semi-autonomous palpation-based method for tumor localization via tactile sensing and utilized deep learning (a CNN model) to successfully classify tissue abnormality with a success rate of 95.24%. Appropriate attributes were established for classifying tumor depths for three categories (shallow, deep, and deeper) using a multi-layer ANN model. The results also demonstrated good performance of the depth estimator network with an accuracy of 82.14% for characterizing the depth of the tumors embedded in the tissue. The proposed RAMIS tumor localization approach showed robustness in the performance even in case of palpating an uneven tissue surface while maintaining a good accuracy.

VI. Acknowledgements

The Tactile Sensing Instrument and the RAMIS setup used in the experimental work for this paper were designed and developed at CSTAR (Canadian Surgical Technologies & Advanced Robotics) under supervision of Dr. Rajni Patel in a project on exploring the effect of haptics in RAMIS tumor localization. The previously collected data at CSTAR were used for the work in this paper. Dr. Patel's support and guidance with the theoretical and experimental development of the RAMIS-based palpation are acknowledged and greatly appreciated.

References

[1] A. M. Okamura, "Haptic feedback in robot-assisted minimally invasive surgery," Current opinion in urology, vol. 19, no. 1, pp. 102–107, 2009.
[2] M. E. H. Eltaib and J. R. Hewit, "Tactile sensing technology for minimal access surgery – a review," Mechatronics, vol. 13, pp. 1163–1177, 2003.
[3] A. Talasaz, R. V. Patel, and M. D. Naish, "Haptics-enabled teleoperation for robot-assisted tumor localization," in IEEE International Conference on Robotics and Automation, 2010, pp. 5340 – 5345.
[4] A. Talasaz and R. Patel, "Integration of force reflection with tactile sensing for minimally invasive robotics-assisted tumor localization," IEEE Transactions on Haptics, DOI: 10.1109/TOH.2012.64, 2012.
[5] ——, "Telerobotic palpation for tumor localization with depth estimation," IEEE/RSJ International Conference on Intelligent Robots and Systemss, pp. 463–468, 2013.
[6] V. Oleksyuk, N. Rahman, and C. H. Won, "Tactile sensing system and convolutional neural network for mechanical property classification," IEEE Sensors Letters, vol. 7, no. 10, pp. 1–4, 2023.
[7] J. Zhao, D. Li, X. X. anf F. Accorsi, H. Marshall, T. Cossetto, D. Kim, D. McCarthy, C. Dawson, S. Knezevic, B. Chen, and S. Li, "United adversarial learning for liver tumor segmentation and detection of multi-modality non-contrast mri," Medical Image Analysis, vol. 73, pp. 1–14, 2021.
[8] C. Hamm, C. Wang, L. Savic, M. Ferrante, I. Schobert, T. Schlachter, M. Lin, J. Duncan, J. Weinreb, J. Chapiro, and B. Letzen, "Deep learning for liver tumor diagnosis part i: development of a convolutional neural network classifier for multi-phasic mri," European Radiology, vol. 29, no. 7, pp. 3338–3347, 2019.
[9] S. Gul, M. Khan, A. Bibi, A. Khandakar, M. Ayari, and M. Chowdhury, "Deep learning techniques for liver and liver tumor segmentation: A review," Computers in Biology and Medicine, vol. 147, pp. 1–16, 2022.
[10] Z. Zhou, B. Huang, R. Zhang, M. Yin, C. Liu, Y. Liu, Z. Yi, and X. Wu, "Methods to recognize depth of hard inclusions in soft tissue using ordinal classification for robotic palpation," IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, vol. 71, p. 2516512, 2022.
[11] B. Xiao, W. Xu, J. Guo, H. Lam, G. Jia, W. Hong, and H. Ren, "Depth estimation of hard inclusions in soft tissue by autonomous robotic palpation using deep recurrent neural network," IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, vol. 17, no. 4, pp. 1791–1799, 2020.
[12] M. T. Perri, A. L. Trejos, M. D. Naish, R. V. Patel, and R. Malthaner, "Initial evaluation of a tactile/kinesthetic force feedback system for minimally invasive tumor localization," IEEE/ASME Transaction on Mechatronics, vol. 15, no. 6, pp. 925–931, Dec 2010.
[13] Quanser. [Online]. Available: http://www.quanser.com
[14] http://www.pressureprofile.com/.