

# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://SPIDigitalLibrary.org/conference-proceedings-of-spie)

## Deep learning-based abdominal muscle segmentation on CT images of surgical patient populations

Chaudhary, Usamah, Leitch, Ka'Toria, Chhabra, Avneesh, Kohli, Ajay, Fei, Baowei

Usamah Chaudhary, Ka'Toria N. Leitch, Avneesh Chhabra, Ajay Kohli, Baowei Fei, "Deep learning-based abdominal muscle segmentation on CT images of surgical patient populations," Proc. SPIE 12036, Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging, 120361N (4 April 2022); doi: 10.1117/12.2611773

**SPIE.**

Event: SPIE Medical Imaging, 2022, San Diego, California, United States

# Deep Learning-Based Abdominal Muscle Segmentation on CT Images of Surgical Patient Populations

Usamah Chaudhary<sup>a,b</sup>, Ka'Toria N. Leitch<sup>a</sup>, Avneesh Chhabra<sup>b, c</sup>, Ajay Kohli<sup>b</sup>, and Baowei Fei<sup>a,b</sup>

<sup>a</sup>Department of Bioengineering, University of Texas at Dallas, Richardson, TX

<sup>b</sup>Department of Radiology, UT Southwestern Medical Center, Dallas, TX

<sup>c</sup>Department of Orthopedic Surgery, UT Southwestern Medical Center, Dallas, TX

## ABSTRACT

Computed tomography (CT) is commonly used for the characterization and tracking of abdominal muscle mass in surgical patients for both pre-surgical outcome predictions and post-surgical monitoring of response to therapy. In order to accurately track changes of abdominal muscle mass, radiologists must manually segment CT slices of patients, a time-consuming task with potential for variability. In this work, we combined a fully convolutional neural network (CNN) with high levels of preprocessing to improve segmentation quality. We utilized a CNN based approach to remove patients' arms and fat from each slice and then applied a series of registrations with a diverse set of abdominal muscle segmentations to identify a best fit mask. Using this best fit mask, we were able to remove many parts of the abdominal cavity, such as the liver, kidneys, and intestines. This preprocessing was able to achieve a mean Dice similarity coefficient (DSC) of 0.53 on our validation set and 0.50 on our test set by only using traditional computer vision techniques and no artificial intelligence. The preprocessed images were then fed into a similar CNN previously presented in a hybrid computer vision-artificial intelligence approach and was able to achieve a mean DSC of 0.94 on testing data. The preprocessing and deep learning-based method is able to accurately segment and quantify abdominal muscle mass on CT images.

**Keywords:** Abdominal imaging, machine learning, surgical patient outcomes, CT, convolutional neural network

## 1. INTRODUCTION

Computed tomography (CT) scans are often used for a variety of purposes from diagnosis to monitoring. Analysis of sarcopenia or atrophy of the abdominal muscles can be used in the monitoring of disease and as an outcome prediction biomarker. Several recent studies have correlated pre-surgical muscle mass with both survival and surgical outcomes.<sup>1-5</sup> Semi-automated tissue segmentation from abdominal CT has also been used to predict odds of treatment response in inflammatory bowel disease.<sup>6</sup>

Thus, a robust and reliable automated abdominal muscle segmentation algorithm would aid in analysis of patient sarcopenia in the clinical setting and for longitudinal patient follow ups. Automation of muscle segmentation is a well-studied task given its many uses.<sup>7,8</sup> Other groups have been able to achieve high perform CNNs, but few have looked specifically at surgical patient population CTs. In our previous work, we presented a fully convolutional neural network (CNN) able to take an input abdominal patient CT slice and perform a reliable segmentation of the abdominal muscle mass.<sup>9</sup> However, the model struggles with segmentation in cases where tissue adjacent to the abdominal muscle is of similar texture, such as slices at the level of the liver. To solve this problem, we propose increasing preprocessing techniques to augment and improve our previous model.

In the present work, we outline a series of image preprocessing methods to remove extraneous regions, such as the abdominal cavity in order to reduce the noise fed into the CNN and to improve the segmentation results, especially in previous worst case scenarios. We utilized atlas-based registration, a method commonly used in brain MRI segmentation algorithms.<sup>10,11</sup> First a diverse set of CT slices was generated from the analysis of our training set, then we used this atlas set and registered it with each new input image. This allows us to achieve an initial segmentation of the abdominal muscle region via solely traditional computer vision techniques, and

even achieved results similar to challenging segmentation tasks<sup>12</sup> without the use of deep learning. We then feed these segmented images into a CNN presented in our previous work to achieve the new results.

The use of a hybrid computer vision-artificial intelligence approach for certain tasks has been noted in literature.<sup>13</sup> Both Hu et al. and Kiranmayee et al. were able to utilize various computer vision and image processing techniques in preprocessing to improve or speed up their machine learning algorithms.<sup>14, 15</sup>

## 2. METHODS

### 2.1 Dataset

CT imaging data from the L2 level to the lesser trochanters was retrospectively obtained from 33 patients (age ranging from 18 to 75 years) with a slice thickness of 5 mm. Abdominal muscles were manually segmented by trained readers under supervision of a fellowship-trained musculoskeletal radiologist, example shown in Figure 1. The Hounsfield window for each slice was restricted to the range -128 to 127 to optimize textural analysis of the abdominal muscle by the CNN. An abdominal slice at the L3 level of 39 additional patients was added to the training set to increase generalizability of the model.

### 2.2 CNN for Arm and Fat Removal

Using the expert abdominal muscle segmentation, we applied a filling transformation to achieve an expert-guided complete abdominal region segmentation (shown in Figure 1). We fed these segmentations along with the original slices to create a CNN capable of removing patients' arms and external fat from individual CT slices. We utilized a Unet architecture with Root mean square propagation (RMSProp) as the optimizer, Dice Similarity Coefficient (DSC) defined as  $DSC = \frac{2(A \cap B)}{(A + B)}$  where A represents the ground truth segmentation and B represents the CNN generated segmentation as the loss function, a learning rate of  $10^4$ , and 25 epochs. We divide our dataset into training (64 patients including 39 single patients slices), validation (3 patients), and testing (5 patients).

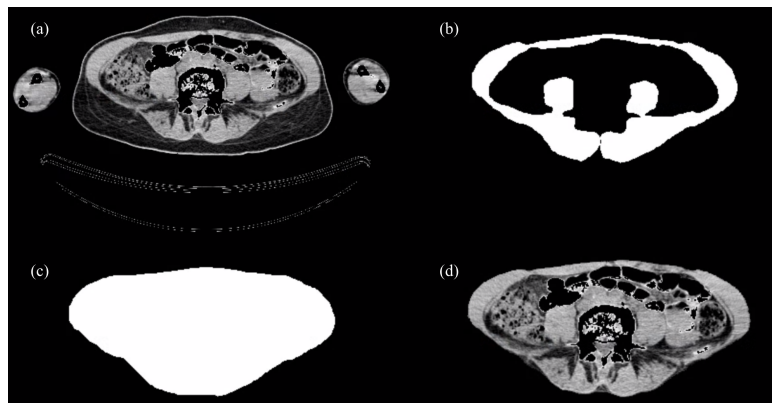


Figure 1: Arm Removal CNN. (a) shows an original CT slice with Hounsfield range of [-128, 127]. (b) shows an expert radiologist segmentation of the abdominal muscle. (c) shows algorithmic filling of image (b) to fully fill the abdominal cavity. (d) shows our machine learning model prediction run on image (a).

Removal of arms and fat on each CT slice is necessary to properly apply registrations between abdominal regions, as will be seen in Subsection 2.3, while also aiding in the goal of reducing extraneous regions of each slice to be fed into the CNN. Depending on scan site and CT machine, arms also may or may not be included in scans, thus removal is extremely necessary to standardize the images for registration purposes.

### 2.3 Atlas Registration

Our methodology for using registration algorithms as a means of segmentation or augmenting segmentation involves creating a diverse set of atlas masks to which an input abdominal CT slice can register closely. To do so, for each mask  $n$  we performed rigid intensity based registrations between all other expert abdominal muscle segmentation masks in our training set, thus performing  $n \times n$  registrations. To prevent over-fitting to

our training set, we ensured that registrations between slices from the same patient were removed. For these registrations, we calculated a DSC to populate an  $n \times n$  matrix. This matrix allowed us to not only pick the best fit mask for each mask based on DSC, but also select the top  $m$  best fit masks (thus giving us a  $n \times m$  matrix). A distribution of the best fit DSC for each slice is shown in Figure 2.

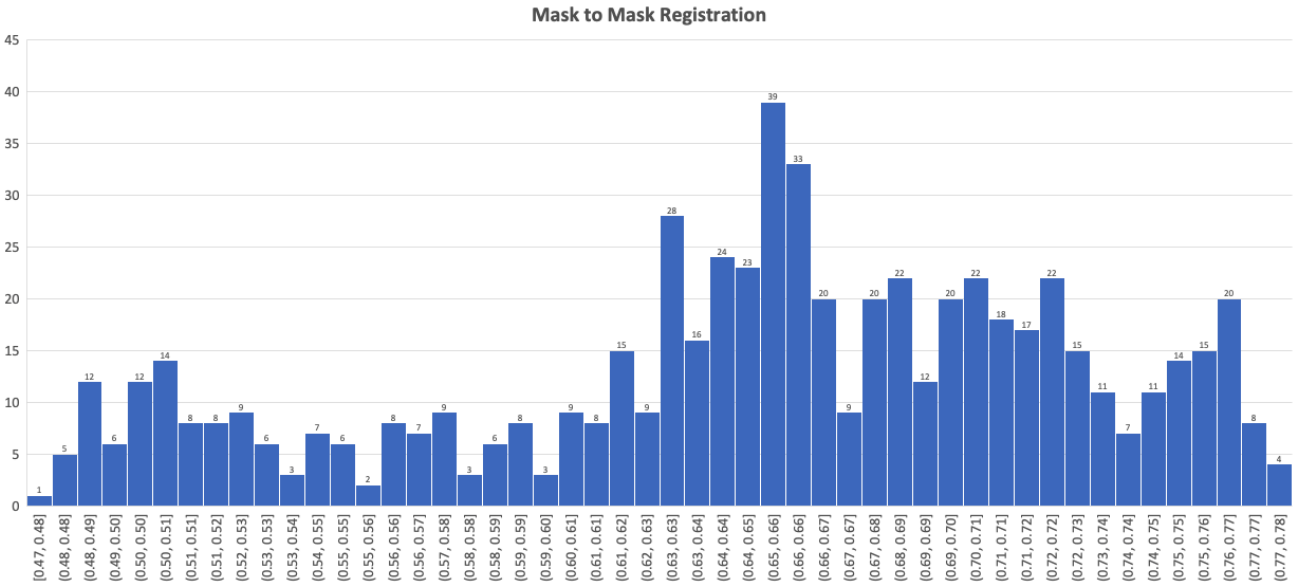


Figure 2: Dice Similarity Coefficient distribution for expert mask to mask registrations.

To create our diverse atlas mask set, we used the  $n \times m$  matrix of best fit mask registrations and determined the most commonly occurring mask. This mask is the current most versatile mask as it registers in the top  $m$  for the most images. We added this mask to our atlas set and removed all masks registering to it in  $n \times m$ , allowing us to calculate the next most versatile mask. We repeated this process until an adequate percent of masks was covered. For our dataset, we used the top 20 masks which achieved a coverage of 88%, meaning that these 20 masks were in the top 20 registration DSC for all  $n$  masks.

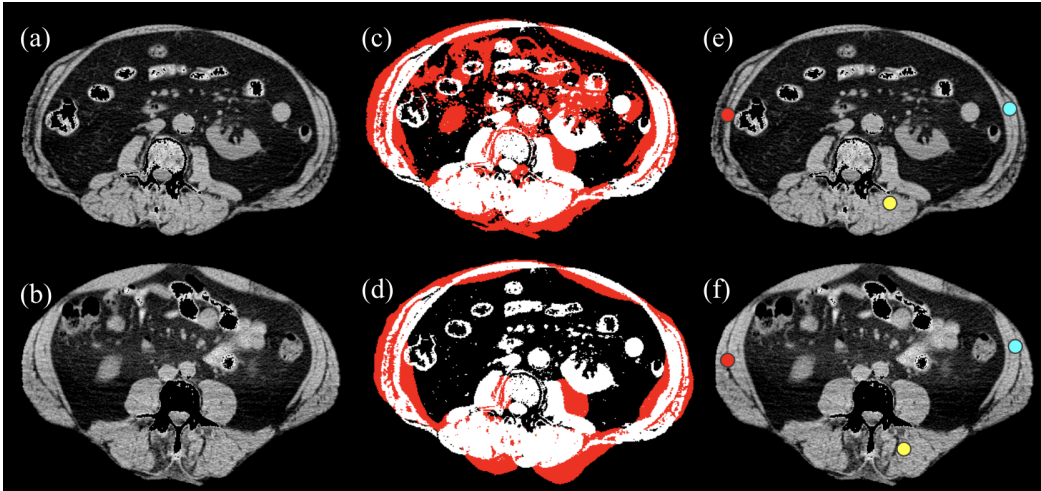


Figure 3: Intensity-Based Registration and Dense Correspondence. (a) and (b) show two CT abdominal slices after arm and fat removal CNN, with (a) serving as the input image and (b) the atlas image. (c) shows the intensity-based registration between (a), shown in white, and (b), shown in red. (d) depicts the expert binary mask, shown in red, placed in the registration position of (b). (e) and (f) illustrate 3-point correspondences (red-red, blue-blue, and yellow-yellow) between the two registered images.

Using our atlas mask set, we were then able to perform our registration step in order to improve inputs into the segmentation CNN. For an input image, we registered it to all images  $m$  within the atlas set (aside from any images in the atlas set that may be from same patient as the input slice) and calculated a DSC (calculated between arm and fat-removed CT slices). We then picked the registration with the highest DSC and in that position, multiplied the binary expert mask of the atlas image to the input image. This allowed us to remove much of the internal abdominal contents that can often interfere with segmentation CNNs. This process is based on the fact that a quality registration allows us to create a dense correspondence between the two images. A dense correspondence means that for each pixel in the atlas image, we can align to a pixel in the input image, thus allowing us to map out the corresponding abdominal muscle in the input image. Dense correspondence is a commonly used method in facial recognition and is extremely effective at matching specific facial regions of interest to one another.<sup>16–18</sup> An example of this correspondence is shown in Figure 3 (e) and (f), for every pixel in (f), the atlas image, we can map to a specific pixel in (e), the input image. This correspondence ensures that we segmented out the abdominal wall when we apply the binary mask of the expert segmentation to the input image. We also applied 3 series of dilations to the mask (shown in Figure 4 (c)) to ensure coverage of abdominal muscle even if over segmenting to ensure maximum sensitivity. An example of this process is shown in Figure 4.



Figure 4: Registration Based Segmentation. (a) shows an original CT abdominal slice after arm and fat removal CNN. (b) shows the best fit mask from the mask atlas set for (a). (c) shows a series of dilations applied to the mask and (d) shows the result after applying the binary mask (c) applied to image (a).

## 2.4 Segmentation CNN

After performing our methods of preprocessing and atlas registration, we were then able to feed all images into a segmentation CNN. We divided our dataset in the same manner as the arm and fat CNN: training (64 patients including 39 single patients slices), validation (3 patients), and testing (5 patients). We are using the same network presented in our previous work<sup>9</sup> and in arm and fat CNN (Subsection 2.2). This network utilizes a Unet architecture with Root mean square propagation (RMSProp) as the optimizer, Dice Similarity Coefficient (DSC) as the loss function, a learning rate of  $10^{-4}$ , and 25 epochs. For post-processing, objects less than 200 pixels were removed to reduce the number of false positives.

## 3. RESULTS

Our CNN for removal of arm and fat utilizes the same training, validation, and test split for our abdominal muscle segmentation CNN (64 training patients, 3 validation patients, and 5 testing patients). For removal of arms and fat, our segmentation CNN was able to achieve a mean DSC of 0.99 and a minimum DSC of 0.98.

The preprocessing methods with a diverse atlas mask set was able to achieve a DSC of 0.53 on validation data and 0.50 on testing data. These images were then fed into a segmentation CNN, described in 2.4. Prior to

feeding into the CNN, we applied a threshold on the DSC of the preprocessing to ensure that the mask was not applied where there was not a quality registration. We applied thresholds at 0.4, 0.5, and 0.6 DSC (summarized in Table 1) and compared the results to our previous model. Cases in which DSC was below these thresholds, the segmentation CNN was applied to just the arm and fat removed image. This ensured that the atlas registration did not diminish segmentation quality for cases where they may not be a good enough fit mask in the atlas dataset.

	Edwards et al.	No Threshold	Threshold at 0.4	Threshold at 0.5	Threshold at 0.6
Validation Mean DSC	0.91	0.90	0.88	0.92	<b>0.93</b>
Validation Minimum DSC	0.74	0.65	0.67	0.80	0.83
Test Mean DSC	0.92	0.89	0.84	0.91	<b>0.94</b>
Test Minimum DSC	0.83	0.58	0.46	0.71	0.87

Table 1: Segmentation CNN Results. Comparison of hybrid computer vision-artificial intelligence segmentation CNN at varying thresholds shown in comparison to previous model Edwards et al.<sup>9</sup>

As seen in Table 1, our comprehensive preprocessing at thresholding of 0.6 DSC was able to increase mean and minimum DSC results in both our validation and independent test sets. In our testing dataset, our preprocessing methods improved the mean DSC by 2.26% and the minimum DSC by 4.13%.

## 4. DISCUSSION AND CONCLUSION

Our main goal in developing image processing methods for increased preprocessing is to improve difficult segmentation cases, such as those where abdominal CT slices contain organs adjacent to the abdominal wall like the liver or slices that have very thin abdominal muscle. By developing methods that are able to remove regions outside of the segmentation region of interest, we are able to improve worst case situations from our previously developed CNN.

Our arm and fat removal CNN is extremely robust given the generalizability of the task, yielding a mean DSC of 0.99. This initial segmentation is needed to clean up and standardize the images for the registration task. This CNN provides us the ability to utilize scans from varying sources without the need of developing specialized tools to remove regions information from the scans.

Atlas based registration is a method commonly used in brain MRI imaging. We developed methods for creating a diverse atlas set from our dataset that is able to effectively register that a large percentage of input scans. Then utilizing the concept of dense correspondence commonly implemented in facial recognition, we used the registration for an initial segmentation of the abdominal muscles. In designing these preprocessing methods, we created a non-artificial intelligence approach to a difficult segmentation task and have achieved a DSC of 0.53 on validation and 0.50 on test sets. As noted in the recent literature,<sup>13–15</sup> incorporating traditional computer vision methods alongside machine learning can aid in both segmentation and classification tasks.

Thus, we then utilized these preprocessing methods to improve our overall segmentation CNN. In comparison to our previous work, this new method is more sensitive and specific. As can be seen in Table 1, in our validation set the worst-case segmentation achieved a DSC of 0.74 while with improved preprocessing methods, it increased to 0.83. For the test set, the minimum DSC improved from 0.83 to 0.87. The increased preprocessing in addition, improves the overall segmentation as well, with increased mean DSC results shown in Table 1.

Our future work includes expansion of our dataset alongside expansion of the atlas set, this will increase initial segmentation and generalizability results. Expansion of the atlas data will decrease the number of cases that are below the threshold to be used in the atlas-registration, which will allow for improved overall segmentation. We also hope to incorporate 3D information into the segmentation CNN in order to improve results.

We present a series of image processing and computer vision techniques alongside machine learning methods to improve results from an abdominal muscle segmentation task. Utilizing a CNN to remove fat and arms from each slice along with novel atlas-registration based initial segmentation, we are able to improve segmentation results from a previously presented CNN with minimal preprocessing. Use of registration with a diverse atlas set allows the identification and removal of extraneous regions outside the abdominal muscle. The methods



developed in the present work can greatly improve current clinical workflows by allows for quick and accurate quantification of abdominal muscle mass on CT. This can be an extremely useful tool by allowing clinicians to easily monitor changes in abdominal muscle mass over time and adjust therapies or make predictions on patient outcomes.

## 5. ACKNOWLEDGEMENTS

This research was supported in part by the U.S. National Institutes of Health (NIH) grants (R01CA156775, R01CA204254, R01HL140325, and R21CA231911) and by the Cancer Prevention and Research Institute of Texas (CPRIT) grant RP190588.

## REFERENCES

- [1] Gibson, D. J., Burden, S. T., Strauss, B. J., Todd, C., and Lal, S., “The role of computed tomography in evaluating body composition and the influence of reduced muscle mass on clinical outcome in abdominal malignancy: a systematic review,” *European journal of clinical nutrition* **69**(10), 1079–1086 (2015).
- [2] DiMartini, A., Cruz Jr, R. J., Dew, M. A., Myaskovsky, L., Goodpaster, B., Fox, K., Kim, K. H., and Fontes, P., “Muscle mass predicts outcomes following liver transplantation,” *Liver Transplantation* **19**(11), 1172–1180 (2013).
- [3] Sui, K., Okabayashi, T., Iwata, J., Morita, S., Sumiyoshi, T., Iiyama, T., and Shimada, Y., “Correlation between the skeletal muscle index and surgical outcomes of pancreaticoduodenectomy,” *Surgery today* **48**(5), 545–551 (2018).
- [4] Lee, J. H., Paik, Y. H., Lee, J. S., Ryu, K. W., Kim, C. G., Park, S. R., Kim, Y. W., Kook, M. C., Nam, B.-h., and Bae, J.-M., “Abdominal shape of gastric cancer patients influences short-term surgical outcomes,” *Annals of surgical oncology* **14**(4), 1288–1294 (2007).
- [5] Weijs, P. J., Looijaard, W. G., Dekker, I. M., Stapel, S. N., Girbes, A. R., Oudemans-van Straaten, H. M., and Beishuizen, A., “Low skeletal muscle area is a risk factor for mortality in mechanically ventilated critically ill patients,” *Critical care* **18**(1), 1–7 (2014).
- [6] Gu, P., Chhabra, A., Chittajallu, P., Chang, C., Mendez, D., Gilman, A., Fudman, D. I., Xi, Y., and Feagins, L. A., “Visceral adipose tissue volumetrics inform odds of treatment response and risk of subsequent surgery in ibd patients starting antitumor necrosis factor therapy,” *Inflammatory Bowel Diseases* (2021).
- [7] Park, H. J., Shin, Y., Park, J., Kim, H., Lee, I. S., Seo, D.-W., Huh, J., Lee, T. Y., Park, T., Lee, J., et al., “Development and validation of a deep learning system for segmentation of abdominal muscle and fat on computed tomography,” *Korean journal of radiology* **21**(1), 88–100 (2020).
- [8] Weston, A. D., Korfiatis, P., Kline, T. L., Philbrick, K. A., Kostandy, P., Sakinis, T., Sugimoto, M., Takahashi, N., and Erickson, B. J., “Automated abdominal segmentation of ct scans for body composition analysis using deep learning,” *Radiology* **290**(3), 669–679 (2019).
- [9] Edwards, K., Chhabra, A., Dormer, J., Jones, P., Boutin, R. D., Lenchik, L., and Fei, B., “Abdominal muscle segmentation from CT using a convolutional neural network,” in *[Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging]*, Krol, A. and Gimi, B. S., eds., **11317**, 135 – 143, International Society for Optics and Photonics, SPIE (2020).
- [10] Liu, Y., Wei, Y., and Wang, C., “Subcortical brain segmentation based on atlas registration and linearized kernel sparse representative classifier,” *IEEE Access* **7**, 31547–31557 (2019).
- [11] Van der Lijn, F., De Bruijne, M., Klein, S., Den Heijer, T., Hoogendam, Y. Y., Van der Lugt, A., Breteler, M. M., and Niessen, W. J., “Automated brain structure segmentation based on atlas registration and appearance models,” *IEEE transactions on medical imaging* **31**(2), 276–286 (2011).
- [12] Chen, Y., Xing, L., Yu, L., Bagshaw, H. P., Buyyounouski, M. K., and Han, B., “Automatic intraprostatic lesion segmentation in multiparametric magnetic resonance images with proposed multiple branch unet,” *Medical physics* **47**(12), 6421–6429 (2020).
- [13] O’Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., Riordan, D., and Walsh, J., “Deep learning vs. traditional computer vision,” in *[Science and Information Conference]*, 128–144, Springer (2019).

- [14] Hu, Z., Alsadoon, A., Manoranjan, P., Prasad, P., Ali, S., and Elchouemic, A., “Early stage oral cavity cancer detection: Anisotropic pre-processing and fuzzy c-means segmentation,” in [*2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*], 714–719, IEEE (2018).
- [15] Kiranmayee, B., Rajinikanth, T., and Nagini, S., “Enhancement of svm based mri brain image classification using pre-processing techniques,” *Indian Journal of Science and Technology* **9**(29), 1–7 (2016).
- [16] Liu, C., Yuen, J., and Torralba, A., “Sift flow: Dense correspondence across scenes and its applications,” *IEEE transactions on pattern analysis and machine intelligence* **33**(5), 978–994 (2010).
- [17] Grewe, C. M. and Zachow, S., “Fully automated and highly accurate dense correspondence for facial surfaces,” in [*European conference on computer vision*], 552–568, Springer (2016).
- [18] Guo, J., Mei, X., and Tang, K., “Automatic landmark annotation and dense correspondence registration for 3d human facial images,” *BMC bioinformatics* **14**(1), 1–12 (2013).