SAND2019-2912C

Segmentation Certainty through Uncertainty: Uncertainty-Refined Binary Volumetric Segmentation under Multifactor Domain Shift

Carianne Martinez, Kevin M. Potter, Matthew D. Smith, Emily A. Donahue, Lincoln Collins,
John P. Korbin, Scott A. Roberts
Sandia National Laboratories*
Albuquerque, NM, USA
{cmarti5, kmpotte, mdsmith, eadonah, lcolli, jpkorbi, sarober}@sandia.gov

Abstract

Deep learning segmentation models are known to be sensitive to the scale, contrast, and distribution of pixel values when applied to Computed Tomography (CT) images. For material samples, scans are often obtained from a variety of scanning equipment and resolutions resulting in domain shift. The ability of segmentation models to generalize to examples from these shifted domains relies on how well the distribution of the training data represents the overall distribution of the target data. We present a method to overcome the challenges presented by domain shifts. Our results indicate that we can leverage a deep learning model trained on one domain to accurately segment similar materials at different resolutions by refining binary predictions using uncertainty quantification (UQ). We apply this technique to a set of unlabeled CT scans of woven composite materials with clear qualitative improvement of binary segmentations over the original deep learning predictions. In contrast to prior work, our technique enables refined segmentations without the expense of the additional training time and parameters associated with additional deep learning models used for postprocessing.

1. Introduction

Advances in non-destructive 3D imaging methods have allowed scientists to study previously hidden features of the natural world. X-ray computed tomography (CT), magnetic resonance imaging (MRI), and other modern diagnostic methods are capable of generating rich data sets, but these methods produce images plagued by noise and scanning artifacts. While it is possible in most cases for a human to interpret imaging data, these interpretations are often expensive, irreproducible, and unreliable. Automated image segmentation is critical in many fields such as medicine, manufacturing, and materials science, where interpretation of data must be done quickly and consistently.

Existing automated segmentation methods such as deep learning models have achieved high accuracy in many image domains, but often fail to generalize when applied to image data from a shifted domain. Our method removes the need for a significant amount of image processing typically used to reduce qualitative and statistical differences between the input domain and target domain.

2. Related Work

The task of semantic segmentation has seen significant improvement after the publication of the Fully Convolutional Network [5] and the encoder-decoder networks that followed. One such architecture, the U-net [8], employed an encoder, decoder, and skip connections to achieve state of the art results on 2D biomedical segmentation. The V-net [7] extended these results to 3D volumes with similar success. Our work enhances the V-net architecture with dropout layers for UQ and focuses on binary segmentation.

While semantic segmentation models have seen further innovation [2], few generalize well if there is a domain gap between the training and testing images. This problem, known as domain shift, has been tackled with adversarial learning [4], co-training [11], or domain statistic alignment approaches [9]. While most approach the problem at the pixel [1], or feature space level [6], Tsai et al [10] address the problem in output space - after inference has occurred. Our work is similarly performed in output space, but rather than using a separate deep learning model to modify the outputs, we leverage the uncertainty in the model's predictions quantified by using dropout at inference time [3].

*Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA0003525. This project was funded through Sandia's Laboratory Directed Research and Development Project 213016.

3. Method

With the goal of automatically segmenting a diverse set of volumetric CT scans of woven composite materials, we train a 3D convolutional neural network (CNN) using a manually labeled training set. Woven composites are materials consisting of yarns comprised of thousands of individual filaments woven into a fabric, assembled into layers, and filled with an epoxy resin. We use label 0 for voxels where no material of interest is present and label 1 to identify voxels that represent the fiber bundles. To quantify the model's uncertainty on a per voxel basis, we employ a dropout technique [3] both during training and inference. When inferring on examples from a domain that is shifted from the training example in resolution, pixel histogram value, or in scan artifacts tied to a specific machine, we run inference multiple times with active dropout layers and generate an uncertainty map for each voxel. The value of the uncertainty at each voxel location is calculated as the standard deviation in the values from the final softmax layer of the CNN over multiple inference runs.

We observe that the CNN segmentations of domain-shifted CT scans consistently predict more material than is present in the images, but that uncertainty in the regions of the model's overzealous material classification is higher than in the regions where the segmentation qualitatively appears to be accurate, an example of which is shown in Figure 3. We flip the labels from 1 to 0 for any voxels in these regions of relatively high uncertainty. To automate this refinement process, we choose an uncertainty value threshold to inform our modifications by optimizing Equation 1 for the best threshold t:

$$\max_t \left| \overline{\mathtt{V}_1(t)} - \overline{\mathtt{V}_0(t)} \right| \tag{1}$$

$$\text{where } \mathtt{V}_1(t) = \{v_k | l_k = 1 \cap u_k \leq t\},$$
$$\mathtt{V}_0(t) = \{v_k | l_k = 0 \cup u_k > t\},$$

v_k is the intensity of voxel k with CNN label l_k and uncertainty value u_k, and t is the uncertainty value threshold. Intuitively, $\overline{\mathtt{V}_i(t)}$ is the average value of voxels labeled i after refining segmentation with uncertainty threshold t. By maximizing Equation 1, we are creating the largest separation between the modes of voxel intensity.

4. Experiment

We used one manually labeled volumetric CT scan of a woven composite material as training data for a modified version of the VNet 3D CNN architecture [7] with dropout layers [3] for UQ. To effectively train the model, we divided the entire volume into a set of 48 sub-volumes for training and sets of 8 sub-volumes for validation and testing. We normalized all sub-volumes to have pixel values with zero mean and unit variance.

To obtain the uncertainty values for each voxel, we run inference 5 times with dropout enabled, and take the standard deviation over the values of the models final softmax layer for each voxel.

Our average accuracy over the 8 held-out test sub-volumes from the training domain was 99.1% to the human labels, and the uncertainty maps reflect higher variance in predictions along the boundary between classes. An example of a 2D slice of a CT scan from the test set in the training domain is shown in Figure 1 along with the human label, and CNN label with uncertainty map associated with the slice overlaid.

We then applied our trained model to infer segmentations for a set of CT scans of similarly structured materials, but with varying chemical compositions, resolutions, and from different scanners. The histograms of the normalized pixel values for these domain-shifted scans are significantly different than that of the training data, as shown in Figure 2. By naively applying the previously trained model to these volumes, the direct results were unusable, as the model consistently overestimates the amount of material present, as shown in the second column 2 of Figure 4.

We apply the method described in Section 3 to refine the model's segmentation by flipping the prediction for any voxel labeled 1 whose uncertainty value is higher than the threshold as automatically determined by the metric shown in Equation 1.

5. Results

Figure 4 shows examples of results of segmentation both before and after our uncertainty-driven refinement method is applied. There is a clear qualitative improvement in the segmentation for these examples. We qualitatively compared the results of our method to those obtained by a simple thresholding approach based on pixel value. Experts in the material confirmed that the thresholding approach resulted in noisier segmentations when compared with our refined segmentation over a small set of example slices. Also, thresholding ignores the feature lengthscale necessary for characterizing the material at the desired resolution.

6. Conclusions and Future Work

We have presented preliminary results that introduce a novel method of uncertainty-driven refinement of binary segmentation over shifted domains. While our results are promising, we acknowledge that future work is needed to quantify the improvements enabled by our model. We require more scans of similar materials to understand how this work generalizes, which kinds of features the uncertainty map can consistently identify, and to identify more precisely the qualities we may vary in scans while maintaining segmentation accuracy.

References

[1] Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan. Unsupervised pixel-level domain adaptation with generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3722–3731, 2017.

[2] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 801–818, 2018.

[3] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *International Conference on Machine Learning*, pages 1050–1059, 2016.

[4] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei A Efros, and Trevor Darrell. Cycada: Cycle-consistent adversarial domain adaptation. *arXiv preprint arXiv:1711.03213*, 2017.

[5] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[6] Mingsheng Long, Han Zhu, Jianmin Wang, and Michael I Jordan. Unsupervised domain adaptation with residual transfer networks. In *Advances in Neural Information Processing Systems*, pages 136–144, 2016.

[7] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571. IEEE, 2016.

[8] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[9] Baochen Sun and Kate Saenko. Deep coral: Correlation alignment for deep domain adaptation. In *European Conference on Computer Vision*, pages 443–450. Springer, 2016.

[10] Yi-Hsuan Tsai, Wei-Chih Hung, Samuel Schulter, Kihyuk Sohn, Ming-Hsuan Yang, and Manmohan Chandraker. Learning to adapt structured output space for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7472–7481, 2018.

[11] Zhi-Hua Zhou and Ming Li. Tri-training: Exploiting unlabeled data using three classifiers. *IEEE Transactions on Knowledge & Data Engineering*, (11):1529–1541, 2005.

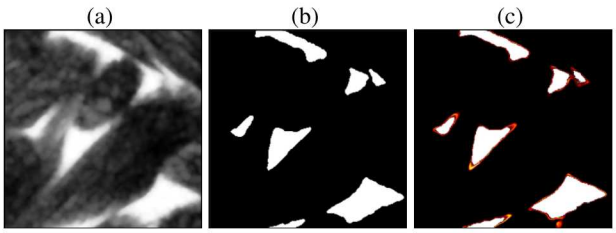

Figure 1. Binary segmentation results on held-out test examples from training domain. (a) Slice of CT scan of woven composite material. (b) Human binary labels for slice. (c) 3D CNN predicted binary labels for slice with uncertainty overlaid.

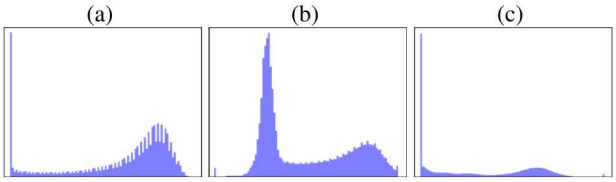

Figure 2. Histogram shapes of normalized pixel values from various domains: (a) CT scan from training domain. (b) and (c) CT scans outside of the training domain.

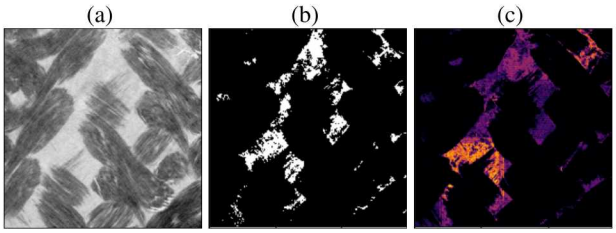

Figure 3. Binary segmentation results on domain-shifted CT scan from CNN without refinement. (a) Slice of CT scan from shifted domain. (b) 3D CNN predicted binary labels for slice. (c) Uncertainty maps for each voxel with brighter pixels representing more uncertainty.

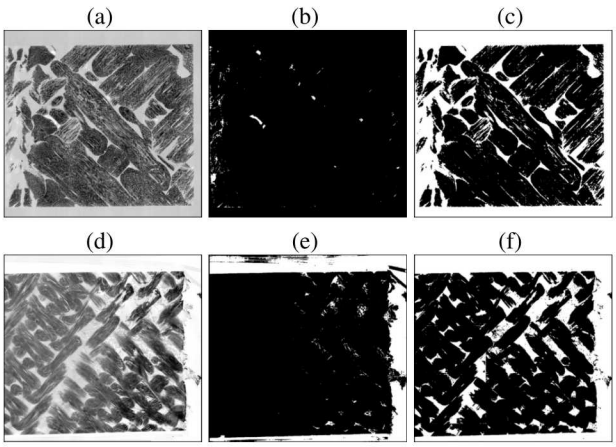

Figure 4. Results from our method applied to CT scans of woven composite materials with composition, resolution, and greyscale histogram different from the training set. (a) and (d) Slices of CT scan to be segmented. (b) and (e) Predicted binary labels for the CT slice from the trained CNN without uncertainty-guided refinement. (c) and (f) Resulting binary labels after applying our method of uncertainty-guided refinement.