USING ENCODER-DECODER CONVOLUTIONAL NETWORKS TO SEGMENT CARBON FIBER CT

Daniel Sammons & William P. Winfree
NASA Langley Research Center
Hampton, VA 23681, USA
{daniel.m.sammons,william.p.winfree}@nasa.gov

ABSTRACT

Materials that exhibit high strength-to-weight ratio, a desirable property for aerospace applications, often present unique inspection challenges. Nondestructive evaluation (NDE) addresses these challenges by utilizing methods, such as x-ray computed tomography (CT), that can capture the internal structure of a material without causing changes to the material. Analyzing the data captured by these methods requires a significant amount of expertise and is costly. Since the data captured by NDE techniques often is structured as images, deep learning can be used to automate initial analysis. This work looks to automate part of this initial analysis by applying the efficient encoder-decoder convolutional network at multiple scales to perform identification and segmentation of defects for NDE.

1 INTRODUCTION

Balancing the strength-to-weight ratio of materials used in aerospace applications is crucial as reductions in weight can lead directly to a reduction in cost to operate a vehicle in which a material is used. However, the complexity of materials that exhibit a high strength-to-weight ratio often present unique inspection challenges. Nondestructive evaluation (NDE) addresses these challenges by utilizing methods, such as x-ray computed tomography (CT), that can capture the internal structure of a material without causing changes to the material. Analyzing the data captured by these methods requires a significant amount of expertise and is costly. Therefore, reducing the time required to perform this analysis would have a significant impact.

For this work, we look to automate part of the NDE analysis process, namely, the identification and segmentation of defects. In particular, we look at automating the initial analysis of carbon fiber reinforced polymer (CFRP) by segmenting a type of defect known as a delamination. We build upon previous work that showed initial success utilizing convolutional networks for dealmination segmentation in CFRP (Sammons et al., 2016). The improvements described in this work result in a significant reduction of processing time and better generalization performance.

2 RELATED WORK

Recently, there has been significant interest in deep learning for image segmentation (Badrinarayanan et al., 2015; Chen et al., 2015; Long et al., 2015; Visin et al., 2015). Of particular interest to this work are applications of deep learning to medical image segmentation.

For example, Ciresan et al. (2012) trained a convolutional network to classify the center pixel of small patches sampled from images of neuronal membrane. More recently, Wang et al. (2015) utilized an encoder-decoder approach for wound segmentation that allowed for end-to-end training from the raw input to the segmentation.

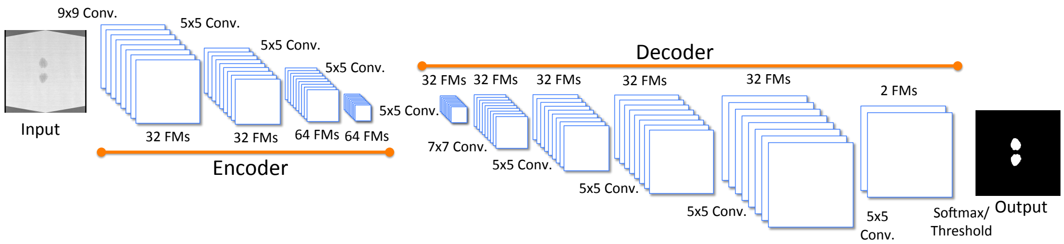

Figure 1: Encoder-decoder convolutional network architecture utilized in this work. Each of the encoder layers also includes an ReLU and max-pooling layer after the convolutional layer. Decoder layers each include a nearest-neighbor upsampling layer before the convolutional layer and an ReLU after. There is no pooling/upsampling in the layer between the encoder/decoder. The final layer of the decoder does not include upsampling/ReLU.

3 Methods

3.1 Encoder-Decoder Convolutional Network

Inspired by previous work in wound segmentation by Wang et al. (2015), we decided to utilize convolutional networks in an encoder-decoder architecture as an efficient method to perform segmentation. An encoder-decoder convolutional network functions much like a convolutional autoencoder except that it is trained to produce a segmentation instead of a reconstruction of the input. In particular, the encoder maps the input to a representation which captures information about the structure in the image. That representation is then mapped back to the original resolution of the image by the decoder, producing a pixel-by-pixel segmentation of the image.

The architecture used in this work (Figure 1) is identical to the architecture utilized by Wang et al. (2015) except that we used nearest-neighbor upsampling instead of unpooling as nearest-neighbor upsampling lead to significantly faster model convergence during training.

3.2 Encoding Multiple Scales

An interesting challenge of utilizing convolutional networks for segmentation is balancing the fine-grained detail required for precise segmentation with the position of the pixel within the larger context of the image. As described by Sammons et al. (2016), this problem manifests itself when using convolutional networks for delamination segmentation in CFRP when pixels located in the center of delaminated regions are labeled as background.

In order to obtain more context to make a prediction for each pixel, we used the pretrained encoder network to encode images at decreasing resolutions. These encodings were then resized to the size of the encoding for the highest resolution image and fed into the decoder as feature maps. The decoder was then trained to combine the information from the different resolution encodings to produce the final segmentation.

3.3 Regularization with Reconstruction

Previous work has determined that pretraining a network using autoencoding acts as a form of regularization on a deep neural network (Erhan et al., 2010). Recently, Zhao et al. (2015) suggested that the regularization effect of pretraining with an autoencoder is limited. Instead, they suggested that better regularization could be provided by training a classifier from the encoded representation of an autoencoder while concurrently training the autoencoder.

Inspired by this approach, we decided to experiment with regularizing encoder-decoder convolutional networks by simultaneously training two decoders to decode the same representation, training one decoder for segmentation and training the other decoder to reconstruct the original input. This joint training forced the encoder network to encode a more robust representation of the data, preventing the segmentation encoder-decoder from developing a trivial solution. We found this method

Table 1: Results from testing the encoder-decoder models on simulated datasets.

Network	Pixel Error	Pixel Precision	Pixel Recall	Mean IoU
Single-scale	98.8%	73.3%	70.9%	56.4%
Multi-scale	98.9%	89.0%	61.5%	58.3%

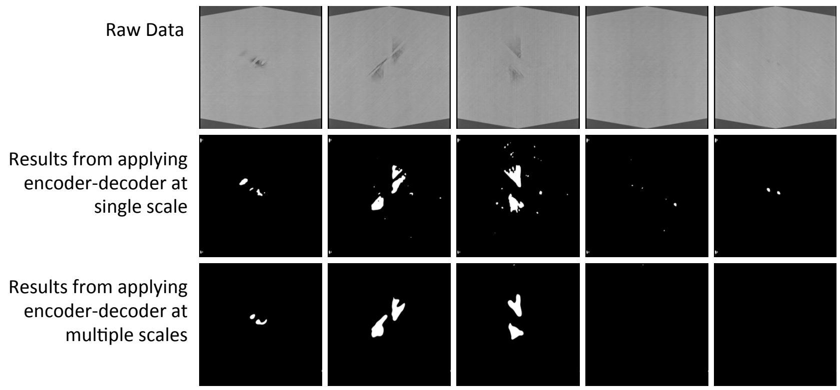

Figure 2: Results from classifying real data with models trained on simulated data.

of regularization crucial for training the encoder-decoder network with entire images as training without it would result in severe overfitting of the training set.

4 TRAINING

Labeling data of real samples of NDE data is extremely expensive because of the expertise required to perform the labeling. As such, we did not have a significant set of real data to use for training. Instead, we chose to train on a set of simulated data that was designed to mimic many of the key characteristics of delaminations in CFRP.

Training was accomplished using stochastic gradient descent. When reconstruction regularization was employed, training was accomplished by performing a step of stochastic gradient descent on the same input with each decoder. Challenges with training stemming from the severe class imbalance which was present in the high-resolution images of CFRP were addressed using the "snowball" training method (Wang & Jean, 1993).

5 RESULTS AND ANALYSIS

Table 1 provides quantitative results from segmenting a test set of simulated data. Since we did not have a labeled real dataset, we only provided qualitative results for real images in Figure 2.

The multi-scale encoder-decoder convolutional network provides the most consistent results on the real data. Further, the results are quite remarkable considering the models were only trained with simulated data. We are not sure whether to attribute this success to the ability of the simulated data to reflect the characteristics of the real data or to the generalization abilities of the models.

6 CONCLUSIONS AND FUTURE WORK

In this work, we showed that applying an encoder-decoder convolutional network at multiple scales is an effective method for performing segmentation of delaminations in CFRP. In the future, we hope to perform better quantitative analysis on real datasets and would like to compare our results with other methods for image segmentation.

REFERENCES

Vijay Badrinarayanan, Ankur Handa, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *arXiv preprint arXiv:1505.07293*, 2015.

Liang-Chieh Chen, Yi Yang, Jiang Wang, Wei Xu, and Alan L Yuille. Attention to scale: Scale-aware semantic image segmentation. *arXiv preprint arXiv:1511.03339*, 2015.

Dan Ciresan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pp. 2843–2851, 2012.

Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, and Samy Bengio. Why does unsupervised pre-training help deep learning? *The Journal of Machine Learning Research*, 11:625–660, 2010.

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*, pp. 3431–3440, 2015.

Daniel Sammons, William P. Winfree, Eric Burke, and Shuiwang Ji. Segmenting delaminations in carbon fiber reinforced polymer composite ct using convolutional neural networks. *AIP Conference Proceedings*, 1706:110014, 2016. doi: http://dx.doi.org/10.1063/1.4940585.

Francesco Visin, Kyle Kastner, Aaron Courville, Yoshua Bengio, Matteo Matteucci, and Kyunghyun Cho. Reseg: A recurrent neural network for object segmentation. *arXiv preprint arXiv:1511.07053*, 2015.

Changhan Wang, Xinchen Yan, Max Smith, Kanika Kochhar, Marcie Rubin, Stephen M Warren, James Wrobel, and Honglak Lee. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pp. 2415–2418. IEEE, 2015.

Jin Wang and Jack Jean. Resolving multifont character confusion with neural networks. *Pattern Recognition*, 26(1):175–187, 1993.

Junbo Zhao, Michael Mathieu, Ross Goroshin, and Yann Lecun. Stacked what-where auto-encoders. *arXiv preprint arXiv:1506.02351*, 2015.