# Convolutional Neural Nets with Template-Based Data Augmentation for Functional Lung Imaging Segmentation

Nicholas J. Tustison[1], Zixuan Lin[1], Brian B. Avants[2], Nick Cullen[3], Jaime F. Mata[1], James C. Gee[3], Talissa A. Altes[4], John P. Mugler III[1], and Kun Qing[1]

[1]Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, VA

[2]Biogen, Cambridge, MA

[3]Department of Radiology University of Pennsylvania, Philadelphia, PA

[4]Department of Radiology, University of Missouri, Columbia, MO

Corresponding author:

Nicholas J. Tustison

ntustison@virginia.edu

**Rationale and Objectives:** We propose an automated segmentation pipeline based on deep learning for ventilation-based quantification which improves on previous methods in terms of robustness and computational efficiency. The large data requirements for the proposed framework is made possible by a novel template-based data augmentation strategy.

**Materials and Methods:** Convolutional neural net (i.e., U-net) models were generated using a custom multilabel Dice metric loss function and a novel template-based data augmentation strategy. Development occurred within *ANTsRNet*—a growing open-source repository of well-known deep learning architectures first introduced here which interfaces with the Advanced Normalization Tools package and the R statistical project. Training (including template generation and data augmentation) employed 500 images. Evaluation was performed on the remaining 1?? images through comparison with a previously reported automated segmentation algorithm based on Gaussian mixture modelling with Markov Random field (MRF) spatial priors.

**Results:**

**Conclusions:** The proposed deep learning framework yielded comparable results as the MRF-based algorithm. Such an approach reduces computational time without sacrificing accuracy.

**Key Words:** Advanced Normalization Tools, ANTsRNet, hyperpolarized gas imaging, neural networks, U-net

# INTRODUCTION

Probing lung function under a variety of conditions and/or pathologies has been significantly facilitated by the use of hyperpolarized gas imaging and corresponding quantitative image analysis methodologies. Such developments have provided direction and opportunity for current and future research trends [1]. Computational techniques targeting these imaging technologies permit quantification of spatial ventilation with potential for increased reproducibility, resolution, and robustness over traditional spirometry and radiological readings [2, 3].

In this research context, one of the most frequently used image-based biomarkers for the study of pulmonary development and disease is based on the quantification of regions of limited ventilation, also known as *ventilation defects*. These features have been shown to be particularly salient in a clinical context, for example ventilation defect volume to total lung volume ratio has been shown to outperform other image-based features in discriminating asthmatics vs. non-asthmatics [4]. This has motivated the development of multiple automated (and semi-automated) segmentation algorithms which have been proposed in the literature [5–9] and are currently used in a variety of clinical research investigations (e.g., [10]).

Despite the enormous methodological progress, recent developments in maching learning (specifically "deep learning" [11]) have generated new possibilities for quantification with improved capabilities in terms of accuracy, robustness, and computational efficiency. Deep learning, a term connoting neural network architectures with multiple hidden layers, has gained prominence in recent years due, in large part, to the annual ImageNet Large Scale Visual Recognition Challenge [12]. Specifically, one of the participants of the 2012 ImageNet challenge, a convolutional neural network colloquially known as "AlexNet" [13], significantly surpassed anything that had been proposed previously. The subsequent outgrowth of research, in conjunction with advances in computational hardware, has resulted in significant developments in various image research areas including classification, segmentation, and object localization and has led to co-optation by the medical imaging analysis community [14].

In this work, we develop and evaluate a convolutional neural network segmentation framework, based on the U-net architecture [15], for functional lung imaging using hyperpolarized gas. As part

of this framework we include a deep learning counterpart to earlier work from our group targeting segmentation of proton lung MRI [16]. This is motivated by common use case scenarious in which proton images are used for quantifying corresponding ventilation images (e.g., [5–7]).

One of the drawbacks to deep learning approaches are the large data requirements for the training process oftentimes necessitating ad hoc strategies for simulating additional data from available data—a process termed *data augmentation* [17]. While common approaches to data augmentation might include the application of randomized simulated linear (e.g., translation, rotation and shearing) or elastic transformations and intensity adjustments (e.g., brightness and contrast), we propose an approach tailored to medical imaging scenarios. In the proposed approach, an optimal shape-based template is constructed from a subset of the available data. Subsequent pairwise image registration between all training data and the resulting template permits a "pseudo-geodesic" transformation of each image to every other image thus converting a data set of size $N$ to an augmented data set of size $N^2$. In this way, transformations are constrainted to the shape space representing the population of interest.

To enhance relevance to the research community, we showcase this work in conjunction with the introduction of *ANTsRNet*—a growing open-source repository of well-known deep learning architectures which interfaces with the Advanced Normalization Tools (ANTs) package [18] and its interface with the R statistical project (i.e., ANTsR) [18]. This permits the public distribution of all code, data, and models which can be found on the GitHub repository corresponding to this manuscript [19]

# MATERIALS AND METHODS

## *Image acquisition*

Hyperpolarized MR image acquisition was performed under an Institutional Review Board (IRB)-approved protocol with written informed consent obtained from each subject. In addition, all imaging was performed under an Food and Drug Administration (FDA)-approved physician's Investigational New Drug application (IND 57866) for hyperpolarized 3He. MRI data were acquired on a 1.5 T whole-body MRI scanner (Siemens Sonata, Siemens Medical Solutions, Malvern, PA) with broad-

band capabilities and a flexible 3He chest radiofrequency coil (RF; IGC Medical Advances, Milwaukee, WI; or Clinical MR Solutions, Brookfield, WI). During a 10–20-second breath-hold following the inhalation of hyperpolarized gas, a set of 19–28 contiguous axial sections were collected. Parameters of the fast low angle shot sequence were as follows: repetition time msec / echo time msec, 7/3; flip angle 10°; matrix, 80 × 128; field of view, 26 × 42 cm; section thickness, 10 mm; and intersection gap, none. Total acquisition time varies between 5-8 seconds depending on the size of the subjects.

A three-dimensional (3D) proton gradient-echo sequence (repetition time [TR]:1.80 ms, echo time [TE] 0.78 ms, flip angle 10□, bandwidth per pixel 1090 Hz/Pixel, partial Fourier: phase direction 6/8, slice direction 6/8) was used to acquire multiple images sets from multiple subjects at varying inflation levels. Acquisition time was 4 sec per image set. All imaging studies were performed under a physician's Investigational New Drug application for Xe129 MRI using a protocol approved by Institutional Review Board of our institute. All subjects provided written informed consent and the data were deidentified prior to analysis.

## *Image Analysis*

### Preprocessing

## Template-based data augmentation

In addition to these software contributions, a significant methodological contribution we have made is the design of a template-based data augmentation strategy. The need for large training data sets is a well-known limitation associated with deep learning algorithms. Whereas the architectures developed for such tasks as the ImageNet competition have access to millions of annotated images, such data access is not always is available and such is typically the case in medical imaging. In order to achieve data set sizes necessary for learning functional models, various data augmentation strategies have been employed. These include application of intensity transformations, such as brightening and enhanced contrast, and simple spatial transformations, such as arbitrary rotations and translations. Regarding the latter, such transformations are not ideal as they might not reflect what is typically seen in medical images and might not sufficiently sample the shape-space of the population

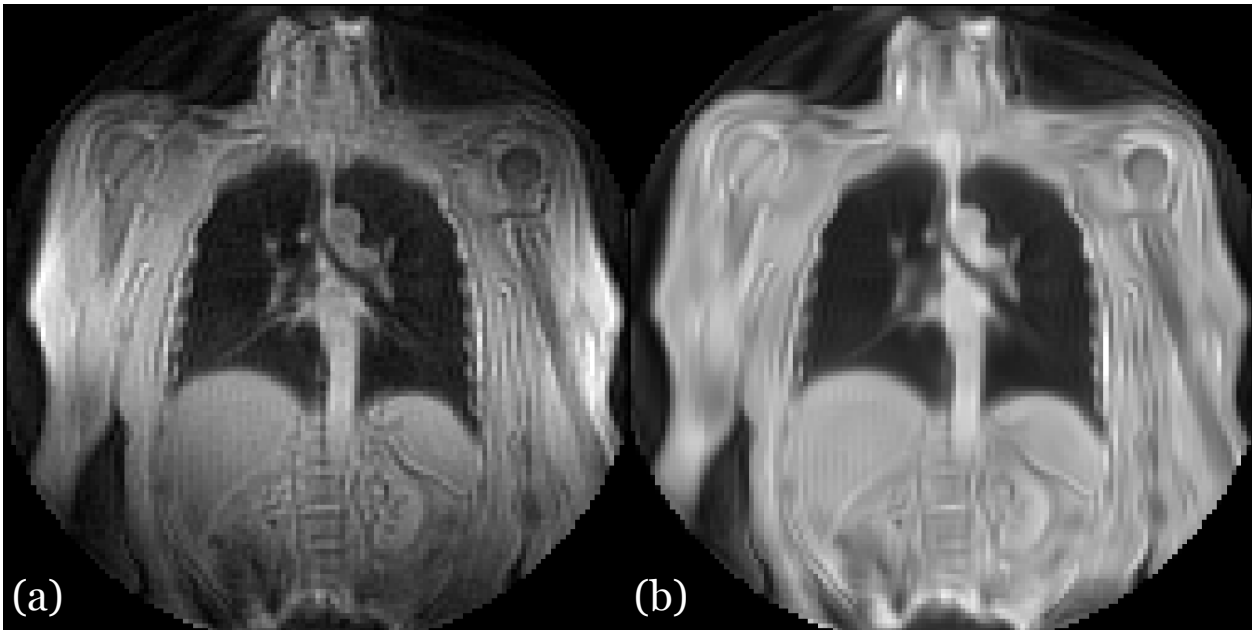


**Figure 1:** Example side-by-side image comparison showing the effects of preprocessing on the proton MRI. (a) Uncorrected image showing MR field inhomogeneity and noise. (b) Corresponding corrected image in which the bias and noise effects have been ameliorated.

currently being studied.

We currently use a template-based approach whereby image data sampled from the population is used to construct a representative template that is optimal in terms of both shape and intensity [20]. In addition to the representative template, this template-building process yields the transformations to/from each individual image to the template space. This permits a propagation of the training data to the space of each individual image. In the simplest case, the training data is used to construct the template and then each individual training data is propagated to the space of every other individual training data. In this way, a training data set of size $N$ can be expanded to a data set of size $N^2$ (cf Figure 1). A more complicated use case could build a template from $M$ data sets (where $M > N$). Transformations between the training data and the template could then be used to propagate the training data to the spaces of the individual members of the template-generating data for an augmented data set size of $M \times N$.

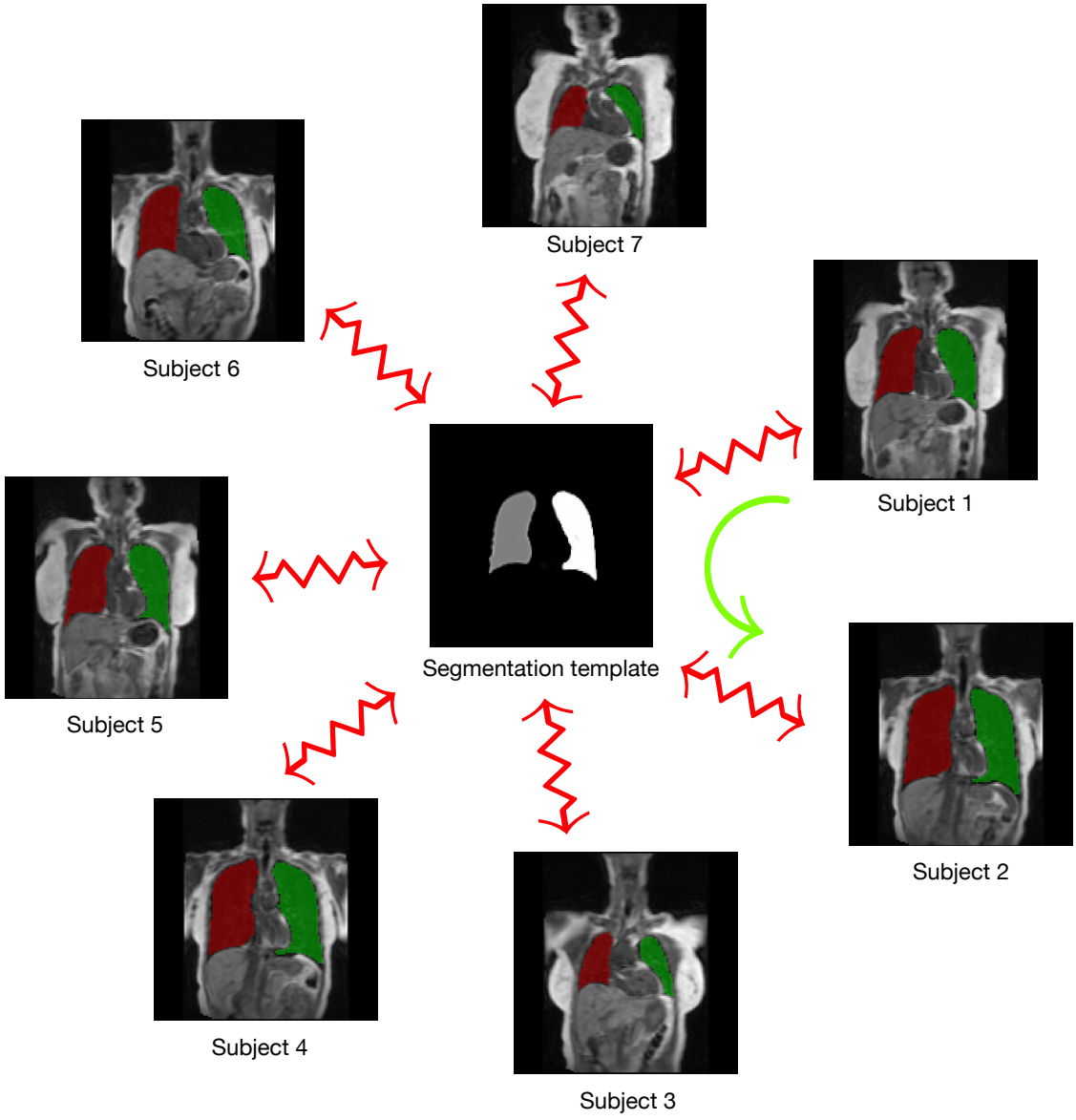


**Figure 2:** We introduce a novel data augmentation strategy for medical images using ANTs-based template construction. Shown here is the 2-D U-net example where we create a template from the training data segmentation images where the foreground designates the left and right lungs. This avoids the lack of internal correspondence while generating plausible global shape variations when mapping between individual training data. We used 60+ images to create such a template permitting $60^2$ = 3600 possible deformable shapes which can be further augmented by more conventional strategies (e.g., brightness transformations, translations, etc.).

## ANTsRNet

The recent interest in deep learning techniques and the associated successes with respect to a variety of applications has motivated adoption of such techniques within the medical imaging research community. Basic image operations such as classification, object identification, and segmentation (as well as more focused techniques) has significant potential for facilitating basic medical research. In light of these new developments, and in order to better meet the modern needs of the community, we have modified this specific aim for ITK-Lung to include the implementation and dissemination of open-source deep learning architectures relevant to the use cases of our partner investigators.

Towards this end, we have created *ANTsRNet*–a collection of well-known deep learning architectures ported to the R language. ANTsRNet is built using the Keras neural network library (available through R) and is highly integrated with the ANTsR package, the R interface of the ANTs toolkit. Consistent with our other software offerings, ongoing development is currently carried out on GitHub using a well-commented coding style, thorough documentation, and self-contained working examples.

It should be noted that various implementations of different deep learning architectures exist and are largely available to the public. However, we feel that this work fills an unmet need. Based on our own search, many publicly available implementations, while functional, are not developed with large-scale distribution and application as end goals. There is little, if any, coding consistency between the various implementations leading to non-standardized APIs and difficulties in code navigation for debugging and/or didactic reasons. In addition, the vast majority employ the Python language which is understandable given its widespread usage by data scientists. However, this work makes these powerful new developments available through a major platform heavily used by statisticians and data scientists alike. In addition, the R-based interface to the ANTs toolkit allows for preprocessing and data augmentation strategies specific to medical imaging. As a result of these current efforts, we were recently awarded a Titan XP GPU from the NVIDIA corporation for facilitating ongoing development.

Although much work remains to be completed, we have made significant progress. As noted below, several architectures have been implemented for both 2-D and 3-D images spanning the broad ap-

plication areas of image classification, object detection, and image segmentation. It should be noted that most reporting in the literature has dealt exclusively with 2-D implementations. This is understandable due to memory and computational speed constraints limiting practical 3-D application on current hardware. However, given the importance that 3-D data has for medical imaging and the rapid progress in hardware, we feel it worth the investment in implementing corresponding 3-D architectures. Each architecture is accompanied by one or more self-contained examples for testing and illustrative purposes. In addtion, we have made novel data augmentation strategies available to the user and illustrated them with Keras-specific batch generators. These contributions are outlined below.

# RESULTS

# DISCUSSION

**Table 1:** Current ANTsRNet capabilities comprising architectures for applications in image segmentation, image classification, and object localization. Self-contained examples with data are also provided to denstrate usage for each of the architectures. Although the majority of neural network architectures are originally described for 2-D images, we generalized the work to 3-D implementations where possible.

| **ANTsRNet** | | |
|---|---|---|
| **Image Segmentation** | | |
| U-net [**?**] | (2-D) | Extends fully convolutional neural networks by including an upsampling decoding path with skip connections linking corresponding encoding/decoding layers. |
| V-net [**?**] | (3-D) | 3-D extension of U-net which incorporates a customized Dice loss function. |
| **Image Classification** | | |
| AlexNet [**?**] | (2-D, 3-D) | Convolutional neural network that precipitated renewed interest in neural networks. |
| VGG16/VGG19 [**?**] | (2-D, 3-D) | Also known as 'OxfordNet'. VGG architectures are much deeper than AlexNet. Two popular styles are implemented. |
| GoogLeNet [**?**] | (2-D) | A 22-layer network formed from *inception blocks* meant to reduce the number of parameters. |
| ResNet [**?**] | (2-D, 3-D) | Characterized by specialized *residualized blocks* (and skip connections. |
| ResNeXt [**?**] | (2-D, 3-D) | A variant of ResNet distinguished by a hyper-parameter called *cardinality* definining the number of independent paths. |
| DenseNet [**?**] | (2-D, 3-D) | Based on the observation that performance is typically enhanced with shorter connections between the layers and the input. |
| **Object Localization** | | |
| SSD300/SSD512 [**?**] | (2-D, 3-D) | The Multibox Single-Shot Detection (SSD) algorithm for determining bounding boxes around objects of interest. |
| SSD7 [**?**] | (2-D, 3-D) | Lightweight SSD variant which increases speed by slightly sacrificing accuracy. Training size requirements are smaller. |

**ACKNOWLEDGMENTS**

*Add ITK Lung grant here.*

We also gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

# REFERENCES


1. Liu, Z., Araki, T., Okajima, Y., Albert, M., and Hatabu, H. "**Pulmonary Hyperpolarized Noble Gas Mri: Recent Advances and Perspectives in Clinical Application**" *Eur J Radiol* 83, no. 7 (2014): 1282–1291. doi:10.1016/j.ejrad.2014.04.014

2. Roos, J. E., McAdams, H. P., Kaushik, S. S., and Driehuys, B. "**Hyperpolarized Gas Mr Imaging: Technique and Applications**" *Magn Reson Imaging Clin N Am* 23, no. 2 (2015): 217–29. doi:10.1016/j.mric.2015.01.003

3. Adamson, E. B., Ludwig, K. D., Mummy, D. G., and Fain, S. B. "**Magnetic Resonance Imaging with Hyperpolarized Agents: Methods and Applications**" *Phys Med Biol* 62, no. 13 (2017): R81–R123. doi:10.1088/1361-6560/aa6be8

4. Tustison, N. J., Altes, T. A., Song, G., Lange, E. E. de, Mugler, J. P., 3rd, and Gee, J. C. "**Feature Analysis of Hyperpolarized Helium-3 Pulmonary Mri: A Study of Asthmatics Versus Nonasthmatics**" *Magn Reson Med* 63, no. 6 (2010): 1448–55. doi:10.1002/mrm.22390

5. Tustison, N. J., Avants, B. B., Flors, L., Altes, T. A., Lange, E. E. de, Mugler, J. P., 3rd, and Gee, J. C. "**Ventilation-Based Segmentation of the Lungs Using Hyperpolarized (3)He Mri**" *J Magn Reson Imaging* 34, no. 4 (2011): 831–41. doi:10.1002/jmri.22738

6. Kirby, M., Heydarian, M., Svenningsen, S., Wheatley, A., McCormack, D. G., Etemad-Rezai, R., and Parraga, G. "**Hyperpolarized 3He Magnetic Resonance Functional Imaging Semiautomated Segmentation**" *Acad Radiol* 19, no. 2 (2012): 141–52. doi:10.1016/j.acra.2011.10.007

7. He, M., Kaushik, S. S., Robertson, S. H., Freeman, M. S., Virgincar, R. S., McAdams, H. P., and Driehuys, B. "**Extending Semiautomatic Ventilation Defect Analysis for Hyperpolarized (129)Xe Ventilation Mri**" *Acad Radiol* 21, no. 12 (2014): 1530–41. doi:10.1016/j.acra.2014.07.017

8. Zha, W., Niles, D. J., Kruger, S. J., Dardzinski, B. J., Cadman, R. V., Mummy, D. G., Nagle, S. K., and Fain, S. B. "**Semiautomated Ventilation Defect Quantification in Exercise-Induced Bronchoconstriction Using Hyperpolarized Helium-3 Magnetic Resonance Imaging: A Repeatability Study**" *Acad Radiol* 23, no. 9 (2016): 1104–14. doi:10.1016/j.acra.2016.04.005

9. Hughes, P. J. C., Horn, F. C., Collier, G. J., Biancardi, A., Marshall, H., and Wild, J. M. "**Spatial Fuzzy c-Means Thresholding for Semiautomated Calculation of Percentage Lung Ventilated Volume from Hyperpolarized Gas and 1 H Mri**" *J Magn Reson Imaging* 47, no. 3 (2018): 640–646. doi:10.1002/jmri.25804

10. Trivedi, A., Hall, C., Hoffman, E. A., Woods, J. C., Gierada, D. S., and Castro, M. "**Using Imaging as a Biomarker for Asthma**" *J Allergy Clin Immunol* 139, no. 1 (2017): 1–10. doi:10.1016/j.jaci.2016.11.009

11. LeCun, Y., Bengio, Y., and Hinton, G. "**Deep Learning**" *Nature* 521, (2015): 436–444.

12. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. "**ImageNet Large Scale Visual Recognition Challenge**" *International Journal of Computer Vision* 115, no. 3 (2015): 211–252.

13. Krizhevsky, A., Sutskever, I., and Hinton, G. E. "**ImageNet Classification with Deep Convolutional Neural Networks**" *Commun. ACM* 60, no. 6 (2017): 84–90. doi:10.1145/3065386, Available at http://doi.acm.org/10.1145/3065386

14. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Laak, J. A. W. M. van der, Ginneken, B. van, and Sánchez, C. I. "**A Survey on Deep Learning in Medical Image Analysis**" *Med Image Anal* 42, (2017): 60–88. doi:10.1016/j.media.2017.07.005

15. Ronneberger, O., Fischer, P., and Brox, T. "**U-Net: Convolutional Networks for Biomedical Image Segmentation**" *Proceedings of the international conference on medical image computing and computer-assisted intervention* 9351, (2015): 234–241.

16. Tustison, N. J. and Herrera, J. M. "**Two Luis Miguel Fans Walk into a Bar in Nagoya —> (Yada, Yada, Yada) —> an ITK-Implementation of a Popular Patch-Based Denoising Filter**" *Insight Journal* (2016):

17. Taylor, L. and Nitschke, G. "**Improving Deep Learning Using Generic Data Augmentation**" *CoRR* abs/1708.06020, (2017): Available at http://arxiv.org/abs/1708.06020

18. Available at https://github.com/stnava/ANTsR

19. Available at https://github.com/ntustison/DeepVentNet

20. Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., and Gee, J. C. "**The Optimal Template Effect in Hippocampus Studies of Diseased Populations**" *Neuroimage* 49, no. 3 (2010): 2457–66. doi:10.1016/j.neuroimage.2009.09.062

21. Ronneberger, O., Fischer, P., and Brox, T. "**U-Net: Convolutional Networks for Biomedical Image Segmentation**" *CoRR* abs/1505.04597, (2015): Available at http://arxiv.org/abs/1505.04597

22. Milletari, F., Navab, N., and Ahmadi, S. "**V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation**" *CoRR* abs/1606.04797, (2016): Available at http://arxiv.org/abs/1606.04797

23. Krizhevsky, A., Sutskever, I., and Hinton, G. E. "**ImageNet Classification with Deep Convolutional Neural Networks**" *Proceedings of the 25th international conference on neural information processing systems - volume 1* (2012): 1097–1105. Available at http://dl.acm.org/citation.cfm?id=2999134.2999257

24. Simonyan, K. and Zisserman, A. "**Very Deep Convolutional Networks for Large-Scale Image Recognition**" *CoRR* abs/1409.1556, (2014): Available at http://arxiv.org/abs/1409.1556

25. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. "**Rethinking the Inception Architecture for Computer Vision**" *CoRR* abs/1512.00567, (2015): Available at http://arxiv.org/abs/1512.00567

26. He, K., Zhang, X., Ren, S., and Sun, J. "**Deep Residual Learning for Image Recognition**" *CoRR* abs/1512.03385, (2015): Available at http://arxiv.org/abs/1512.03385

27. Xie, S., Girshick, R. B., Dollár, P., Tu, Z., and He, K. "**Aggregated Residual Transformations for Deep Neural Networks**" *CoRR* abs/1611.05431, (2016): Available at http://arxiv.org/abs/1611.05431

28. Huang, G., Liu, Z., and Weinberger, K. Q. "**Densely Connected Convolutional Networks**" *CoRR* abs/1608.06993, (2016): Available at http://arxiv.org/abs/1608.06993

29. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S. E., Fu, C., and Berg, A. C. "**SSD: Single Shot Multibox Detector**" *CoRR* abs/1512.02325, (2015): Available at http://arxiv.org/abs/1512.

02325

30. Available at https://github.com/pierluigiferrari/ssd_keras