Learning image transformations via convolutional neural networks: a review

Nicholas J. Tustison[1], Brian B. Avants[1], James C. Gee[2],

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

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

Corresponding author:

Nicholas J. Tustison

ntustison@virginia.edu

Abstract

Recent methodological innovations in deep learning and associated advancements in computational hardware have significantly impacted the various core subfields of quantitative medical image analysis. The generalizability, computational efficiency, and open-source availability of deep learning algorithms, particularly those utlizing convolutional neural networks, have produced paradigm shifts within the field. This impact is evident from topical prevalance in the literature, conference and workshop themes, and winning methodologies in relevant competitions. In this work, we review the various state-of-the-art, fully convolutional network approaches to learning and predicting image transformations. Although of primary importance within the quantitative imaging domain, image registration algorithmic development, in the context of these deep learning strategies, has received comparatively less attention than its counterparts (e.g., image segmentation). Nevertheless, significant inroads have been made and presented in various research venues. We contextualize these contributions within the broader scope of deep learning advancements and, in so doing, attempt to facilitate the leveraging and further development of such techniques within the medical imaging research community.

Key words: deep learning, diffeomorphisms, image registration, spatial normalization

Introduction

Determining the spatial correspondence between imaging domains is frequently a critical component in quantitative image analysis workflows. The evolution of image registration theoretical and technological development has led to increasingly high quality transformational mappings that have significantly improved performance in related processing tasks (e.g., image segmentation via joint label fusion [1]) and imaging-based statistical analysis strategies (e.g., sparse canonical correlation analysis [2]). Several reviews [3–8] have charted this chronology and provided insight into related issues such as algorithmic classification, available implementations, evaluation strategies, and speculation concerning future directions of the field. While prescient in many respects, speculation vis-à-vis deep learning was somewhat limited due to deep learning's relatively recent and sudden explosion in popularity and research focus.

The foundational concepts that form the basis for contemporary deep learning studies date back decades [9, 10]. From this historically seminal work, major developmental milestones include the *Neocognitron*, an early neural network architecture for character recognition [11], and convolutional neural networks (CNN or ConvNets) utilized in speech [12] and visual signal processing [13], largely inspired by the visual cell types of the feline visual cortex [14]. The common approach to gradient-based optimization of CNNs using backpropagation was first performed in [13]. A key event in the widespread adoption of CNNs was the 2012 ImageNet Large Scale Visual Recognition Challenge for object classification [15]. The winning entry, an architecture colloquially known as *AlexNet* [16], reduced the error rate by almost half over other entries. Subsequent years' competitions were dominated by CNN variants such as VGG [17] and GoogLeNet [18].

- Local connections
- shared weights
- pooling
- use of many layers

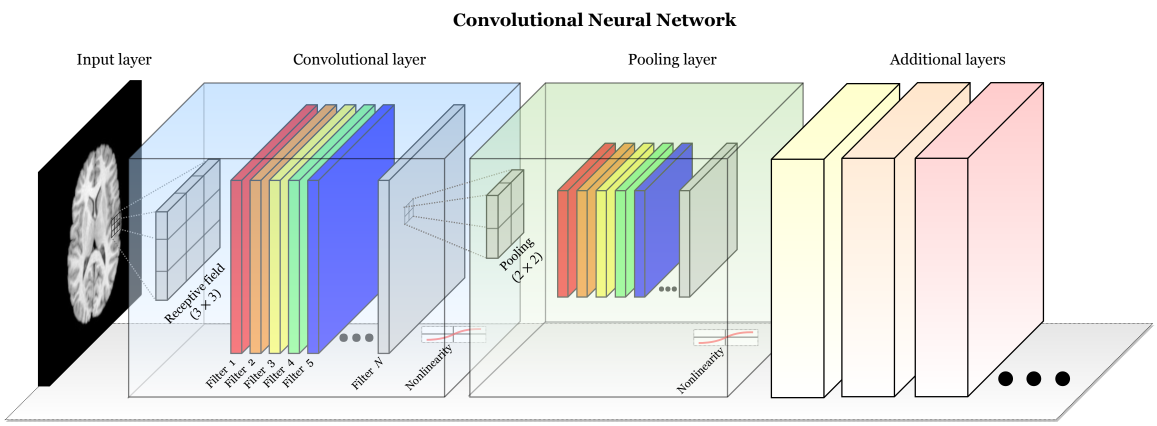

Figure 1: Caption here.

References

1. Iglesias, J. E. and Sabuncu, M. R. "**Multi-Atlas Segmentation of Biomedical Images: A Survey**" *Med Image Anal* 24, no. 1 (2015): 205–219. doi:10.1016/j.media.2015.06.012

2. Avants, B. B., Cook, P. A., Ungar, L., Gee, J. C., and Grossman, M. "**Dementia Induces Correlated Reductions in White Matter Integrity and Cortical Thickness: A Multivariate Neuroimaging Study with Sparse Canonical Correlation Analysis**" *Neuroimage* 50, no. 3 (2010): 1004–16. doi:10.1016/j.neuroimage.2010.01.041

3. Brown, L. G. "**A Survey of Image Registration Techniques**" *ACM Comput. Surv.* 24, no. 4 (1992): 325–376. doi:10.1145/146370.146374, Available at http://doi.acm.org/10.1145/146370.146374

4. Maintz, J. B. and Viergever, M. A. "**A Survey of Medical Image Registration**" *Med Image Anal* 2, no. 1 (1998): 1–36.

5. Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. "**Mutual-Information-Based Registration of Medical Images: A Survey**" *IEEE Trans Med Imaging* 22, no. 8 (2003): 986–1004. doi:10.1109/TMI.2003.815867

6. Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., and Gopinath, K. "**Brain Functional Localization: A Survey of Image Registration Techniques**" *IEEE Trans Med Imaging* 26, no. 4 (2007): 427–51. doi:10.1109/TMI.2007.892508

7. Viergever, M. A., Maintz, J. B. A., Klein, S., Murphy, K., Staring, M., and Pluim, J. P. W. "**A Survey of Medical Image Registration - Under Review**" *Med Image Anal* 33, (2016): 140–144. doi:10.1016/j.media.2016.06.030

8. Keszei, A. P., Berkels, B., and Deserno, T. M. "**Survey of Non-Rigid Registration Tools in Medicine**" *J Digit Imaging* 30, no. 1 (2017): 102–116. doi:10.1007/s10278-016-9915-8

9. LeCun, Y., Bengio, Y., and Hinton, G. "**Deep Learning**" *Nature* 521, no. 7553 (2015): 436–44. doi:10.1038/nature14539

10. Ivakhnenko, A. G. "**Polynomial Theory of Complex Systems**" *IEEE Transactions on Sys-*

tems, Man, and Cybernetics SMC-1, no. 4 (1971): 364–378.

11. Fukushima, K. "**Neocognitron: A Self Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position**" *Biol Cybern* 36, no. 4 (1980): 193–202.

12. Waibel, A. "**Phoneme Recognition Using Time-Delay Neural Networks**" *Meeting of the institute of electrical, information and communication engineers (ieice).* (1987):

13. LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jacke, L. D. "**Backpropagation Applied to Handwritten Zip Code Recognition**" *Neural Computation* 1, no. 4 (1989): 541–551.

14. Hubel, D. H. and Wiesel, T. N. "**Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex**" *J Physiol* 160, (1962): 106–54.

15. 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.

16. 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

17. 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

18. 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