Pattern Recognition Lab
Department Informatik
Universität Erlangen-Nürnberg
Prof. Dr.-Ing. habil. Andreas Maier
Telefon: +49 9131 85 27775
Fax: +49 9131 303811
info@i5.cs.fau.de
www5.cs.fau.de

Tutorial: Deep Learning – Advancing the State-of-the-Art in Medical Image Analysis

Vincent Christlein, Florin C. Ghesu, Tobias Würfl, Andreas Maier,
Fabian Isensee, Simon Kohl, Peter Neher, Klaus Maier-Hein

To cite this version:

Christlein, V.; Ghesu, F.-C.; Würfl, T.; Maier, A.; Isensee, F.; Johl, S.; Neher, P.; Maier-Hein, K. "Tutorial: Deep Learning – Advancing the State-of-the-Art in Medical Image Analysis" in Proceedings of Bildverarbeitung für die Medizin 2017, pp. 6-7, March 12–14, 2017, Heidelberg

DOI: 10.1007/978-3-662-54345-0

Tutorial: Deep Learning Advancing the State-of-the-Art in Medical Image Analysis

Vincent Christlein[1], Florin C. Ghesu[1], Tobias Würfl[1], Andreas Maier[1], Fabian Isensee[2], Simon Kohl[2], Peter Neher[2], Klaus Maier-Hein[2]

[1] Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

[2] Junior Group Medical Image Computing, German Cancer Research Center, Heidelberg, Germany

vincent.christlein@fau.de

Deep Learning (DL) represents a key technological innovation in the field of machine learning. Recent advancements have attracted much attention by showing substantial improvements in a wide range of applications such as image recognition, speech recognition, natural language processing and artificial intelligence. In some cases the performance even surpasses human accuracy, which motivated the introduction of a series of DL-based software products and automatization solutions (for example Apple Siri, Google Now, Google Autonomous Driving etc.). The same success also echoes in the research efforts of the medical imaging community. However, in this case several constraints such as data-availability, inherent data noise or lack of labeled data directly affect the pace of advancements. We will start the tutorial by introducing the core element of deep learning – the deep neural network, with its distinguishing capability of automatically learning hierarchies of complex features directly from the raw training data without the need for 'manual' feature engineering. Using this knowledge-base we will discuss several state-of-the-art deep architectures. In this context we will also present the algorithms required to train these models as well as a practical analysis of regularization / data-normalization techniques which prove to be essential in achieving high performance. This technology also addresses inherent limitations faced in typical medical image analysis problems. The most important of these is the task of object recognition / classification which is an important prerequisite for many clinical applications, for example image-to-image registration, advanced biophysical simulations and cell detection or classification problems for cancer diagnosis. To complement the aforementioned learning techniques for classification problems we will dedicate the third part of our tutorial to applications of deep learning to segmentation problems. Here we will outline the two predominant strategies, namely patch-based and single pass segmentation approaches. For both strategies we will present specific systems from current literature as well as examples of our own work in the areas of DL-based volumetric image parsing. The last part of the tutorial is focused on the inherent constraints and obstacles we face when using DL methods in the context of medical image analysis and how they can be addressed. Specifically, we will discuss recent developments in reducing and training large models, such as model compression and semi-supervised learning. We will also discuss solutions that address the computational limitations of DL models for 3D data based on our own research. The tutorial will be concluded with an open discussion related to unresolved problems and the future of deep learning in the medical imaging community.

References

[1] V. Christlein, D. Bernecker, A. Maier, and E. Angelopoulou, "Offline writer identification using convolutional neural network activation features", in *Pattern Recognition: 37th German Conference, GCPR 2015, Aachen, Germany, October 7-10, 2015, Proceedings*, J. Gall, P. Gehler, and B. Leibe, Eds. Cham: Springer International Publishing, 2015, pp. 540–552.

[2] F. C. Ghesu, E. Krubasik, B. Georgescu, V. Singh, Y. Zheng, J. Hornegger, and D. Comaniciu, "Marginal space deep learning: Efficient architecture for volumetric image parsing", *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1217–1228, May 2016.

[3] F. C. Ghesu, B. Georgescu, and J. Hornegger, "Efficient medical image parsing", in *Deep Learning for Medical Image Analysis*, vol. 1, Elsevier, 2017, pp. 55–81.

[4] F. C. Ghesu, B. Georgescu, T. Mansi, D. Neumann, J. Hornegger, and D. Comaniciu, "An artificial agent for anatomical landmark detection in medical images", in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part III*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds. Cham: Springer International Publishing, 2016, pp. 229–237.

[5] F. C. Ghesu, B. Georgescu, Y. Zheng, J. Hornegger, and D. Comaniciu, "Marginal space deep learning: Efficient architecture for detection in volumetric image data", in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 710–718.

[6] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier, "Deep learning computed tomography", in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part III*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds. Cham: Springer International Publishing, 2016, pp. 432–440.

Tutorial Literature

[1] M. Arjovsky *et al.*, "Wasserstein GAN", *ArXiv preprint arXiv:1701.07875*, 2017.

[2] J. L. Ba, J. R. Kiros, and G. E. Hinton, “Layer normalization”, *ArXiv preprint arXiv:1607.06450*, 2016.
[3] K. Bousmalis *et al.*, “Unsupervised Pixel-Level Domain Adaptation with Generative Adversarial Networks”, *ArXiv preprint arXiv:1612.05424*, 2016.
[4] C. F. Cadieu, H. Hong, D. L. K. Yamins, N. Pinto, D. Ardila, E. A. Solomon, N. J. Majaj, and J. J. DiCarlo, “Deep Neural Networks Rival the Representation of Primate IT Cortex for Core Visual Object Recognition”, *PLOS Computational Biology*, vol. 10, no. 12, pp. 1–18, Dec. 2014.
[5] A. Canziani, A. Paszke, and E. Culurciello, “An Analysis of Deep Neural Network Models for Practical Applications”, *ArXiv preprint arXiv:1605.07678*, 2016.
[6] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs”, *ArXiv preprint arXiv:1606.00915*, 2016.
[7] A. Choromanska, M. Henaff, M. Mathieu, G. B. Arous, and Y. LeCun, “The loss surfaces of multilayer networks.”, in *AISTATS*, 2015.
[8] Ö. Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger, “3d u-net: Learning dense volumetric segmentation from sparse annotation”, in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, L. Joskowicz, Ed., ser. LNCS, (available on arXiv:1606.06650 [cs.CV]), vol. 9901, Springer, Oct. 2016, pp. 424–432.
[9] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3d u-net: Learning dense volumetric segmentation from sparse annotation”, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2016, pp. 424–432.
[10] ——, “3d u-net: Learning dense volumetric segmentation from sparse annotation”, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2016, pp. 424–432.
[11] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, “Deep neural networks segment neuronal membranes in electron microscopy images”, in *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc., 2012, pp. 2843–2851.
[12] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, “Identifying and attacking the saddle point problem in high-dimensional non-convex optimization”, in *Advances in neural information processing systems*, 2014, pp. 2933–2941.
[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd ed. New York: Wiley-Interscience, Nov. 2000.
[14] H. Z. *et al.*, “StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks”, *ArXiv preprint arXiv:1612.03242*, 2016.
[15] I. G. *et al.*, “Generative Adversarial Nets”, in *NIPS*, 2014, pp. 2672–2680.
[16] Y. L. *et al.*, “Detecting Cancer Metastases on Gigapixel Pathology Images”, *ArXiv preprint arXiv:1703.02442*, 2017.

[17] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling", *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1915–1929, 2013.

[18] Y. Ganin *et al.*, "Unsupervised Domain Adaptation by Backpropagation", *ArXiv preprint arXiv:1409.7495*, 2014.

[19] F. C. Ghesu, E. Krubasik, B. Georgescu, V. Singh, Y. Zheng, J. Hornegger, and D. Comaniciu, "Marginal space deep learning: Efficient architecture for volumetric image parsing", *IEEE TMI*, vol. 35, no. 5, pp. 1217–1228, 2016.

[20] I. Goodfellow, "Nips 2016 Tutorial: Generative Adversarial Networks", *ArXiv preprint arXiv:1701.00160*, 2016.

[21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[22] K. Greff, R. K. Srivastava, and J. Schmidhuber, "Highway and Residual Networks learn Unrolled Iterative Estimation", *ArXiv preprint arXiv:1612.07771*, 2016.

[23] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural network with pruning, trained quantization and huffman coding", *CoRR*, vol. abs/1510.00149, 2015.

[24] S. Hawkins, H. He, G. Williams, and R. Baxter, "Outlier detection using replicator neural networks", in *International Conference on Data Warehousing and Knowledge Discovery*, Springer, 2002, pp. 170–180.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[26] ——, "Identity Mappings in Deep Residual Networks", in *European Conference on Computer Vision (ECCV)*, Springer, 2016, pp. 630–645.

[27] M. H. Herzog and A. M. Clarke, "Why Vision is not Both Hierarchical and Feedforward", *Frontiers in Computational Neuroscience*, vol. 8, p. 135, 2014.

[28] G. Huang, Y. Sun, Z. Liu, D. Sedra, and K. Q. Weinberger, "Deep Networks with Stochastic Depth", in *European Conference on Computer Vision (ECCV)*, Springer, 2016, pp. 646–661.

[29] S. Hwang and H.-E. Kim, "Self-transfer learning for weakly supervised lesion localization", in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2016, pp. 239–246.

[30] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and ¡1mb model size", *CoRR*, vol. abs/1602.07360, 2016.

[31] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift", in *Proceedings of The 32nd International Conference on Machine Learning*, 2015, pp. 448–456.

[32] P. Isola *et al.*, "Image-to-Image Translation with Conditional Adversarial Networks", *ArXiv preprint arXiv:1611.07004*, 2016.

[33] A. Jamaludin, T. Kadir, and A. Zisserman, "SpineNet: Automatically Pinpointing Classification Evidence in Spinal MRIs", in *International Con-*

ference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 166–175.

[34] S. Jégou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation", *ArXiv preprint arXiv:1611.09326*, 2016.

[35] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation", *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.

[36] A. Karpathy. (2016). CS231n Convolutional Neural Networks for Visual Recognition, [Online]. Available: http://cs231n.github.io/convolutional-networks (visited on 03/06/2017).

[37] A. Khoreva, R. Benenson, J. Hosang, M. Hein, and B. Schiele, "Simple does it: Weakly supervised instance and semantic segmentation", *ArXiv preprint arXiv:1603.07485*, 2016.

[38] V. Koltun, "Efficient inference in fully connected crfs with gaussian edge potentials", *Adv. Neural Inf. Process. Syst*, vol. 2, no. 3, p. 4, 2011.

[39] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", in *Advances In Neural Information Processing Systems 25*, Curran Associates, Inc., 2012, pp. 1097–1105. arXiv: 1102.0183.

[40] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[41] M. Lin, Q. Chen, and S. Yan, "Network in Network", *ArXiv preprint arXiv:1312.4400*, 2013.

[42] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[43] M. Manassi, B. Sayim, and M. H. Herzog, "When Crowding of Crowding Leads to Uncrowding", *Journal of vision*, vol. 13, no. 13, p. 10, 2013.

[44] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation", in *3D Vision (3DV), 2016 Fourth International Conference on*, IEEE, 2016, pp. 565–571.

[45] D. S. Nathan Silberman and R. Fergus, "Instance segmentation of indoor scenes using a coverage loss", in *ECCV*, 2014.

[46] A. Ng, "Cs294a lecture notes", 2011.

[47] D. Nie *et al.*, "Medical Image Synthesis with Context-Aware Generative Adversarial Networks", *ArXiv preprint arXiv:1612.05362*, 2016.

[48] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation", in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1520–1528.

[49] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Is object localization for free?-weakly-supervised learning with convolutional neural networks", in

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 685–694.

[50] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2015, pp. 234–241.

[51] T. Salimans *et al.*, "Improved Techniques for Training GANs", in *Advances in Neural Information Processing Systems*, 2016, pp. 2226–2234.

[52] T. Salimans and D. P. Kingma, "Weight normalization: A simple reparameterization to accelerate training of deep neural networks", in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds., Curran Associates, Inc., 2016, pp. 901–909.

[53] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps", *ArXiv preprint arXiv:1312.6034*, 2013.

[54] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-scale Image Recognition", *ArXiv preprint arXiv:1409.1556*, 2014.

[55] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting.", *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[56] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks", *ArXiv preprint arXiv:1505.00387*, 2015.

[57] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions", in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1–9.

[58] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with Convolutions", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.

[59] F. Tseng. (2016). Neural Nets, [Online]. Available: `http://frnsys.com/ai_notes/machine_learning/neural_nets.html` (visited on 03/07/2017).

[60] A. Veit, M. J. Wilber, and S. Belongie, "Residual Networks Behave Like Ensembles of Relatively Shallow Networks", in *Advances in Neural Information Processing Systems*, 2016, pp. 550–558.

[61] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion", *Journal of Machine Learning Research*, vol. 11, no. Dec, pp. 3371–3408, 2010.

[62] L. Wan, M. Zeiler, S. Zhang, Y. L. Cun, and R. Fergus, "Regularization of neural networks using dropconnect", in *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, 2013, pp. 1058–1066.

[63] Z. Wu, C. Shen, and A. v. d. Hengel, "Wider or deeper: Revisiting the resnet model for visual recognition", *ArXiv preprint arXiv:1611.10080*, 2016.

[64] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions", *ArXiv preprint arXiv:1511.07122*, 2015.

[65] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features", *IEEE Transactions on Medical Imaging*, vol. 27, no. 11, pp. 1668–1681, 2008.

[66] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2921–2929.