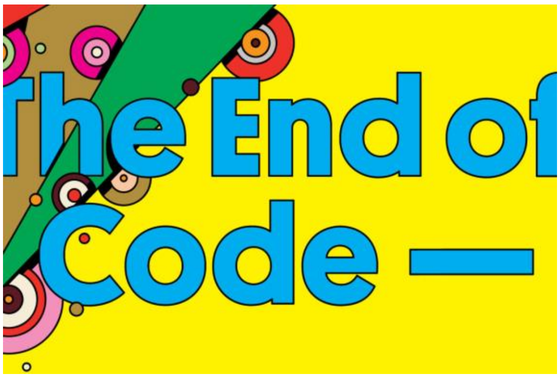
Assistive Computer Vision:
Ten Years Of Technological Transfer And Research With Future Prospects

David Tellez
Radboud University Medical Center
david.tellezmartin@radboudumc.nl

Paradigm Shift

The field of computer vision is living a very profound revolution, a paradigm shift from handcrafted algorithms to smart systems based on machine learning techniques. Recent developments in software, hardware and data availability are enabling scientists and engineers to solve complex vision tasks using more efficient and accurate algorithms, delivering unprecedented results to the scientific community. Through this essay, we will have a look at the reasons of this change, then briefly mention examples of the state-of-art of computer vision, and eventually hypothesize about the near future and new opportunities ahead of us, all of it with a special focus on computer vision applied to healthcare and medical imaging.



Article from Wired in May 2016: "Soon We Won't Program Computers. We'll Train Them Like Dogs" [1]

Deep Learning

If we have to pick the single technology that is making the biggest impact in computer vision nowadays, this would clearly be deep learning. Deep learning is nothing else than very large, densely connected neural networks that are trained end-to-end to automatically learn features from a dataset in order to perform a conventional classification or regression task [2]. In order to fully understand the rise of deep learning as the de facto standard tool to solve challenging computer vision tasks currently, we have to dive deep into the basic ideas that support this technology.

Behind The Scenes

On a more theoretical level, the Holy Grail mechanism that enables the "learning" part of the technique is called backpropagation. Although it is well known since the 70's, it took a few more years to be adopted as the regular algorithm to train neural networks. With backpropagation, networks' parameters can be iteratively adjusted so that the loss function that models the computer vision problem can be minimized. Modern deep networks can be seen as long and complex symbolic mathematical graphs, whose loss function can be seamlessly minimized by propagating the error from the output back through their nodes, assuming that all the operations involved are differentiable so that the chain rule from calculus can be applied [3].

The second pillar of modern deep learning are graphics processing units (GPUs). Although originally their main purpose and target audience was the gaming industry, several reasons such as their low price, high availability and early success accelerating the solution of a variety of scientific problems, quickly impulsed machine learning researchers to try them to train neural networks. Since matrix multiplication is at the core of neural network computations, GPUs can greatly speed up the calculations involved in the training and evaluation of these networks by massively parallelizing their operations. Extremely deep networks that were prohibitively expensive to train once, are now tractable thank to the adoption of GPUs.

The last but not least key idea behind the success of deep learning models is the availability of huge public labeled datasets. One of the most exciting features from deep learning is its ability to automatically and efficiently learn useful features to recognize targets and distinguish among labels. However, this functionality comes at a significant cost: the need for very large labeled training datasets. In order to successfully train modern deep neural networks, models might need to go through hundreds, thousands or millions of labeled images, depending on the use case. It was not until recent years that the scientific community and the industry started to shift from an in-house, private data storage strategy towards a more open attitude where huge datasets were released into the public domain, through either direct publication or organized challenges [4, 5, 6].

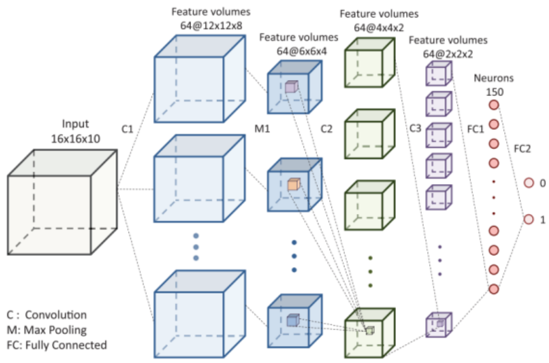
State-Of-The-Art

Machine learning in general, and deep learning in particular, is becoming a core tool in medical image analysis nowadays, beating classic algorithms in an unprecedented pace and pushing the state-of-the-art of the field faster than ever. As recent as in May 2016, the scientific journal Transactions on Medical Imaging (TMI) from IEEE published an entire Special Section on deep learning [7], pointing out the importance of this technology for the development of future solutions for imaging in healthcare.

Computer vision algorithms have typically struggled with image heterogeneity. Classical classifiers relying only on hand-crafted features used to underperform when instances of the same class where too different from each other. Some of the papers published in the TMI deep learning special issue have addressed the problem of heterogeneity by switching to deep neural networks. This problem is particularly severe in histopathological applications where problems like cell nuclei detection and tissue classification are among the most common tasks in the clinic. In this regard, the work of Korsuk Sirinukunwattana reflects the state-of-the-art of detection and classification of cell nuclei in histopathology images of cancerous tissue, where a convolutional neural network based approach outperforms all existing solutions [8]. This piece of work is in line with the thesis of image algorithms revolutionizing how image analysis in healthcare applications is performed and delivered nowadays, achieving exceptional results and potentially freeing valuable personnel in the clinic to work in other tasks with more added value for the patient than “simple” cell nuclei detection.

Another classical problem where computer vision algorithms typically underperform when compared to human agents is the task of image segmentation. Until recently, segmentation algorithms were significantly prone to errors, in part due to the intrinsic complexity of the problem but also as a consequence of the lack of high volume of curated and expertly annotated datasets. Sérgio Pereira demonstrated how powerful deep learning solutions can be when applied to brain tumor segmentation in magnetic resonance imaging (MRI), confirming that the same core technique can be applied to completely different medical imaging modalities such as histopathology and MRI [9].

The field of computer vision is expanding in data dimensionality as well. The idea of exploiting 3D and 4D data is not new in medical image analysis, however, it is now going through a sort of renaissance since deep learning algorithms can painlessly learn underlying target structures in more than two dimensions. Particularly interesting in this regard is the work of Qi Dou who presented in his paper in TMI that convolutional neural networks can be efficiently applied to automatically detect cerebral microbleeds in magnetic resonance (MR) 3D imaging [10]. Furthermore, they claim that their method can be easily generalized to any other detection problem based in volumetric medical data which dives deep into the idea of modern neural networks as a general algorithm for multiple modalities of image data in the medical domain.



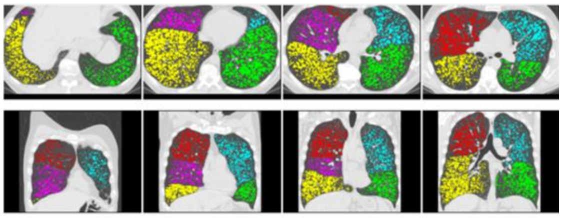
The hierarchical architecture of the 3D CNN model, Qi Dou et al. [10].

Technology Transfer

Assistive computer vision is not only booming among the academic community but also within the startup scene. Many young companies and spin-offs whose main business is based on assistive computer vision have been funded in the recent years all around the globe. Technology transfer plays a key role in this regard, involving high-level university management and individual researchers looking for an opportunity to export and deploy their solutions in daily clinical environments. Motivations to transform an academic prototype or idea into a fully working application are highly varied, and can range from humanitarian reasons to very profitable business opportunities. However, there is a common driving force behind many of the startup ideas that are coping the market nowadays, an idea that every successful startup in the assistive computer vision should be exploiting today: scalability. In the following paragraphs, we are going to discuss a few startup cases where assistive computer vision and scalability are at the core of their visions and strategies.

In the domain of dermatology, SkinVision [11] is a startup funded by a PhD student in 2011 that aims to make skin monitoring an easy routine to reduce deaths from melanoma cancer by detecting tumors in a very early stage. The use case is straightforward, when the user has been exposed to the sun for a long period, or he recognizes a suspicious spot in his skin, a picture of the problematic area is submitted to their system through a smartphone app. Then, an automatic computer vision algorithm examines the image, produces a sort of diagnosis, and recommends an action. This simple idea delivers value to every player in the healthcare system: the patient gets a preliminary diagnose painlessly, insurance institutions cut costs by reducing the number of patients visiting expensive dermatologists, and healthcare staff are released from trivial diagnoses and can focus their attention in more difficult cases. Scalability and modern computer vision technologies are behind this success story of technology transfer.

In the domain of Computed Tomography (CT), we find Thirona [12], a spin-off from the Radboud University Medical Center in the Netherlands that focuses on automatic quantitative analysis of thoracic CT scans, with applications ranging from the detection of lung diseases to the optimization of individual treatment planning. Although their current business model concentrates in installing local workstations in the clinic, their service could be greatly scaled by executing their algorithms in the cloud, delivering results all over the globe. This is a valuable idea because one of the main problems in developing countries is the great lack of human experts to read and interpret medical images such as CT scans. With such a service, local healthcare centers in these regions could upload their CT scans to the system, receiving immediate feedback about the status of the patient, and allowing them to take further action with more expert information in their hands.



Example of automatic quantitative analysis of thoracic CT scans performed by Thirona [12].

Future Prospects

With more data, smarter algorithms, and outstanding results, researchers in the field of assistive computer vision, in particular medical imaging, are realizing that it is time to think big and be ambitious with the goals that they would like to achieve in the next decade.

Doctors' diagnostics are made using multiple sources of information, including several different image modalities such as CT, MRI, or even digital histopathology images. However, the typical use case for computer vision in diagnosis only include one of these modalities, trying to solve a very narrow problem. It is time to leverage all the power of machine learning embracing multiple disciplines within the medical imaging domain. Diagnostics should be understood and treated as a single unit, in a comprehensive manner, where every image subtype contributes to the final outcome, namely the patient diagnosis or prognosis. There is a lot of value to be explored in this regard, giving computer algorithms the opportunity to find patterns and links in a scale that humans struggle with.

References:

[1]. "Soon We Won't Program Computers. We'll Train Them Like Dogs", Wired Magazine, May 2016. http://www.wired.com/2016/05/the-end-of-code/

[2]. http://deeplearning.net/

[3]. How the backpropagation algorithm works. http://neuralnetworksanddeeplearning.com/chap2.html

[4]. The ImageNet database. http://www.image-net.org/

[5]. The Common Objects in Context database. http://mscoco.org/

[6]. Grand Challenges in Biomedical Image Analysis. http://grand-challenge.org/

[7]. IEEE TMI Special Issue on Deep Learning. https://ieee-tmi.org/fast-facts/featured-article.asp?id=14&title=Special-Edition:-Deep-Learning-in-Medical-Imaging,-May-2016

[8]. K. Sirinukunwattana, S. E. A. Raza, Y. W. Tsang, D. R. J. Snead, I. A. Cree and N. M. Rajpoot, "Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images," in IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1196-1206, May 2016.

[9]. S. Pereira, A. Pinto, V. Alves and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," in IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1240-1251, May 2016.

[10]. Q. Dou et al., "Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks," in IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1182-1195, May 2016.

[11]. SkinVision 2016. https://www.skinvision.com/

[12]. Thirona 2016. http://www.thirona.eu/