ScienceDirect

Medical Image Analysis
Volume 42, December 2017, Pages 60-88

Survey Paper

A survey on deep learning in medical image analysis

Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez

Show more

Share Cite

https://doi.org/10.1016/j.media.2017.07.005
Get rights and content

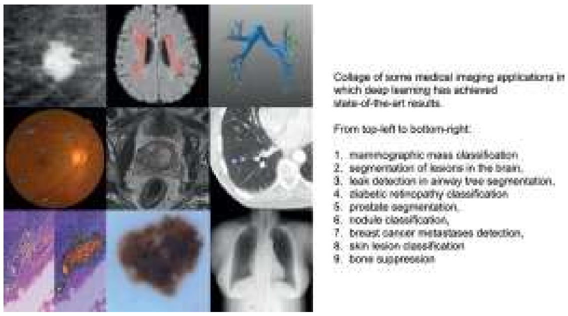
Highlights

- A summary of all deep learning algorithms used in medical image analysis is given.
- The most successful algorithms for key image analysis tasks are identified.
- 300 papers applying deep learning to different applications have been summarized.

Abstract

Deep learning algorithms, in particular convolutional networks, have rapidly become a methodology of choice for analyzing medical images. This paper reviews the major deep learning concepts pertinent to medical image analysis and summarizes over 300 contributions to the field, most of which appeared in the last year. We survey the use of deep learning for image classification, object detection, segmentation, registration, and other tasks. Concise overviews are provided of studies per application area: neuro, retinal, pulmonary, digital pathology, breast, cardiac, abdominal, musculoskeletal. We end with a summary of the current state-of-the-art, a critical discussion of open challenges and directions for future research.

Graphical abstract



Download: Download high-res image (193KB)

Download: Download full-size image

Introduction

As soon as it was possible to scan and load medical images into a computer, researchers have built systems for automated analysis. Initially, from the 1970s to the 1990s, medical image analysis was done with sequential application of low-level pixel processing (edge and line detector filters, region growing) and mathematical modeling (fitting lines, circles and ellipses) to construct compound rule-based systems that solved particular tasks. There is an analogy with expert systems with many if-then-else statements that were popular in artificial intelligence in the same period. These expert systems have been described as GOFAI (good old-fashioned artificial intelligence) (Haugeland, 1985) and were often brittle; similar to rule-based image processing systems.

At the end of the 1990s, supervised techniques, where training data is used to develop a system, were becoming increasingly popular in medical image analysis. Examples include active shape models (for segmentation), atlas methods (where the atlases that are fit to new data form the training data), and the concept of feature extraction and use of statistical classifiers (for computer-aided detection and diagnosis). This pattern recognition or machine learning approach is still very popular and forms the basis of many successful commercially available medical image analysis systems. Thus, we have seen a shift from systems that are completely designed by humans to systems that are trained by computers using example data from which feature vectors are extracted. Computer algorithms determine the optimal decision boundary in the high-dimensional feature space. A crucial step in the design of such systems is the extraction of discriminant features from the images. This process is still done by human researchers and, as such, one speaks of systems with *handcrafted* features.

A logical next step is to let computers learn the features that optimally represent the data for the problem at hand. This concept lies at the basis of many deep learning algorithms: models (networks) composed of many layers that transform input data (e.g. images) to outputs (e.g. disease present/absent) while learning increasingly higher level features. The most successful type of models for image analysis to date are convolutional neural networks (CNNs). CNNs contain many layers that transform their input with convolution filters of a small extent. Work on CNNs has been done since the late seventies (Fukushima, 1980) and they were already applied to medical image analysis in 1995 by Lo et al. (1995). They saw their first successful real-world application in LeNet (LeCun et al., 1998) for hand-written digit recognition. Despite these initial successes, the use of CNNs did not gather momentum until various new techniques were developed for efficiently training deep networks, and advances were made in core computing systems. The watershed was the contribution of Krizhevsky et al. (2012) to the ImageNet challenge in December 2012. The proposed CNN, called AlexNet, won that competition by a large margin. In subsequent years, further progress has been made using related but deeper architectures (Russakovsky et al., 2014). In computer vision, deep convolutional networks have now become the technique of choice.

The medical image analysis community has taken notice of these pivotal developments. However, the transition from systems that use handcrafted features to systems that learn features from the data has been gradual. Before the breakthrough of AlexNet, many different techniques to learn features were popular. Bengio et al. (2013) provide a thorough review of these techniques. They include principal component analysis, clustering of image patches, dictionary approaches, and many more. Bengio et al. (2013) introduce CNNs that are trained end-to-end only at the end of their review in a section entitled *Global training of deep models*. In this survey, we focus particularly on such deep models, and do not include the more traditional feature learning approaches that have been applied to medical images. For a broader review on the application of deep learning in health informatics we refer to Ravi et al. (2017), where medical image analysis is briefly touched upon.

Applications of deep learning to medical image analysis first started to appear at workshops and conferences, and then in journals. The number of papers grew rapidly in 2015 and 2016. This is illustrated in Fig. 1. The topic is now dominant at major conferences and a first special issue appeared of IEEE Transaction on Medical Imaging in May 2016 (Greenspan et al., 2016).

One dedicated review on application of deep learning to medical image analysis was published by Shen et al. (2017). Although they cover a substantial amount of work, we feel that important areas of the field were not represented. To give an example, no work on retinal image analysis was covered. The motivation for our review was to offer a comprehensive overview of (almost) all fields in medical imaging, both from an application and a methodology driven perspective. This also includes overview tables of all publications which readers can use to quickly assess the field. Last, we leveraged our own experience with the application of deep learning methods to medical image analysis to provide readers with a dedicated discussion section covering the state-of-the-art, open challenges and overview of research directions and technologies that will become important in the future.

This survey includes over 300 papers, most of them recent, on a wide variety of applications of deep learning in medical image analysis. To identify relevant contributions PubMed was queried for papers containing ("convolutional" OR "deep learning") in title or abstract. ArXiv was searched for papers mentioning one of a set of terms related to medical imaging. Additionally, conference proceedings for MICCAI (including workshops), SPIE, ISBI and EMBC were searched based on titles of papers. We checked references in all selected papers and consulted colleagues. We excluded papers that did not report results on medical image data or only used standard feed-forward neural networks with handcrafted features. When overlapping work had been reported in multiple publications, only the publication(s) deemed most important were included. We expect the search terms used to cover most, if not all, of the work incorporating deep learning methods. The last update to the included papers was on February 1, 2017. The appendix describes the search process in more detail.

Summarizing, with this survey we aim to:

- show that deep learning techniques have permeated the entire field of medical image analysis;
- identify the challenges for successful application of deep learning to medical imaging tasks;
- highlight specific contributions which solve or circumvent these challenges.

The rest of this survey is structured as followed. In Section 2, we introduce the main deep learning techniques that have been used for medical image analysis and that are referred to throughout the survey. Section 3 describes the contributions of deep learning to canonical tasks in medical image analysis: classification, detection, segmentation, registration, retrieval, image generation and enhancement. Section 4 discusses obtained results and open challenges in different application areas: neuro, ophthalmic, pulmonary, digital pathology and cell imaging, breast, cardiac, abdominal, musculoskeletal, and remaining miscellaneous applications. We end with a summary, a critical discussion and an outlook for future research.

Section snippets

Overview of deep learning methods

The goal of this section is to provide a formal introduction and definition of the deep learning concepts, techniques and architectures that we found in the medical image analysis papers surveyed in this work. …

Image/exam classification

Image or exam classification was one of the first areas in which deep learning made a major contribution to medical image analysis. In exam classification, one typically has one or multiple images (an exam) as input with a single diagnostic variable as output (e.g., disease present or not). In such a setting, every diagnostic exam is a sample and dataset sizes are typically small compared to those in computer vision (e.g., hundreds/thousands vs. millions of samples). The popularity of transfer …

Anatomical application areas

This section presents an overview of deep learning contributions to the various application areas in medical imaging. We highlight some key contributions and discuss performance of systems on large data sets and on public challenge data sets (Fig.3). All these challenges are listed on http://www.grand-challenge.org ↗. …

Overview

From the 308 papers reviewed in this survey, it is evident that deep learning has pervaded every aspect of medical image analysis. This has happened extremely quickly: the vast majority of contributions, 242 papers, were published in 2016 or the first month of 2017. A large diversity of deep architectures are covered. The earliest studies used pre-trained CNNs as feature extractors. The fact that these pre-trained networks could simply be downloaded and directly applied to any medical image …

Acknowledgments

The authors would like to thank members of the Diagnostic Image Analysis Group for discussions and suggestions. This research was funded by grants KUN 2012-5577, KUN 2014-7032, and KUN 2015-7970 of the Dutch Cancer Society. …

Recommended articles

References (350)

G. Apou *et al.*
Detection of lobular structures in normal breast tissue.
Comput. Biol. Med. (2016)

J. Arevalo *et al.*
Representation learning for mammography mass lesion classification with convolutional neural networks.
Comput. Methods Program. Biomed. (2016)

M. Avendi *et al.*
A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI
Med. Image Anal. (2016)

CaiY. *et al.*
Multi-modal vertebrae recognition using transformed deep convolution network.
Comput. Med. Imaging Graph (2016)

J. Charbonnier *et al.*
Improving airway segmentation in computed tomography using leak detection with convolutional networks
Med. Image Anal. (2017)

ChenH. *et al.*
DCAN: Deep contour-aware networks for accurate gland segmentation
Med. Image Anal. (2017)

ChoiH. *et al.*
Fast and robust segmentation of the striatum using deep convolutional neural networks
J. Neurosci. Methods (2016)

D.C. Cireşan *et al.*
Mitosis detection in breast cancer histology images with deep neural networks
Proceedings of the Medical Image Computing and Computer-Assisted Intervention(2013)

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M.,...

M.D. Abràmoff *et al.*
Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning
Invest. Ophthalmol. Vis. Sci. (2016)

View more references

Cited by (11328)

Review of deep learning: concepts, CNN architectures, challenges, applications, future directions ↗

2021, Journal of Big Data

Show abstract

nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation ↗

2021, Nature Methods

UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation ↗

2020, IEEE Transactions on Medical Imaging

A survey on Image Data Augmentation for Deep Learning ↗

2019, Journal of Big Data

A guide to deep learning in healthcare ↗

2019, Nature Medicine

Artificial intelligence in radiology ↗

2018, Nature Reviews Cancer

View all citing articles on Scopus ↗

View full text