Review

Deep Learning for Medical Image Analysis: Applications to Computed Tomography and Magnetic Resonance Imaging

Kyu-Hwan Jung[1], Hyunho Park[1], Woochan Hwang[2]

[1]VUNO Inc., Seoul, Korea
[2]School of Medicine, Imperial College London, London, United Kingdom

Recent advances in deep learning have brought many breakthroughs in medical image analysis by providing more robust and consistent tools for the detection, classification and quantification of patterns in medical images. Specifically, analysis of advanced modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) has benefited most from the data-driven nature of deep learning. This is because the need of knowledge and experience-oriented feature engineering process can be circumvented by automatically deriving representative features from the complex high dimensional medical images with respect to the target tasks. In this paper, we will review recent applications of deep learning in the analysis of CT and MR images in a range of tasks and target organs. While most applications are focused on the enhancement of the productivity and accuracy of current diagnostic analysis, we will also introduce some promising applications which will significantly change the current workflow of medical imaging. We will conclude by discussing opportunities and challenges of applying deep learning to advanced imaging and suggest future directions in this domain.

Key words: Deep Learning; Medical Imaging; Computed Tomography; Magnetic Resonance Imaging

Corresponding Author: Kyu-Hwan Jung
VUNO Inc., 6F, 507, Gangnamdae-ro,
Seocho-gu, Seoul, Korea
Tel: +82-2-515-6646
Fax: +82-2-515-6647
E-mail: kyuhwanjung@gmail.com

INTRODUCTION

Following the recent development in artificial intelligence, where deep learning has become the main methodology, the paradigm of medical image analysis is shifting from the previous clinical experience and knowledge-based feature engineering to the data-driven objective feature analysis of deep learning. Especially, as the application of various techniques developed for natural images to medical images is being accelerated, we are no longer simply adapting the natural image models to medical images but developing new methods, which encompasses the unique characteristics of the medical image domain. Furthermore, as the research on interpretability of decisions made by deep learning models and the way of incorporating clinical knowledge into the model progresses,

we have started to obtain promising results that will allow clinical implementation of deep learning. Among various deep learning models, convolutional neural networks (CNN) have become methodology of choice for visual recognition problems. CNN is a type of feed-forward artificial neural network, which learns hierarchical features by iterating convolution and pooling layers until the output prediction layer is reached. While the convolution layers learn specific patterns in the input or intermediate feature map with locally-connected shared weights, pooling layers reduce the feature map by spatially aggregating activations. In special cases where the output of the model is same as the input or its denoised version, we call the model as convolutional auto-encoder (CAE).

In medical image analysis, machine learning methods have been used in various fields such as detection and classification

of lesions, segmentation of organs, image registration, and similar image retrieval [1]. There were attempts to use CNN for the detection of pulmonary nodules and breast tissue microcalcifications since 1993, when the model had only just been proposed [2,3]. However, due to limitations of available data, model size, learning methodology, and computational resources, such attempts remained at an experimental stage. This was until the recent performance breakthrough of deep learning in image analysis, which led to renewed enthusiasm for applying deep learning in medical imaging. Particularly, after a team using a CNN based model won the ImageNet competition in 2012 by a significant performance gap and machines exceeded humans at an indirect comparison of image recognition task in 2015, the possibility of clinical implementation of deep learning became a major issue [4]. Similarly, in medical imaging, after a CNN based model won the mitotic cell detection task in breast biopsies at the 2012 ICPR (International conference on pattern recognition), recent studies on diabetic retinopathy detection and skin cancer classification demonstrated that deep learning models trained with massive medical images can even surpass the performance of human specialist in diagnostic image analysis [5,6,7].

While early studies focused on 2D medical images, such as chest X-rays, mammograms and histopathological images where deep learning models developed for natural images could be directly applied, recent studies are looking towards applying deep learning on volumetric medical images. Among various volumetric imaging modalities, computed tomography (CT) which uses specialized x-ray equipment to produce cross-sectional images of the body and magnetic resonance (MR) images which uses magnetic field to produce detailed images of soft tissues and organ structures are the most actively studied modalities due to their popularity in diagnostic imaging. However, not alone, the complexity and size of these volumetric images but also their contrast-enhanced or follow up images increase the difficulties of assessing these modalities and have restricted the capabilities of computer-aided systems for medical image analysis. In this regard, recent studies for CT and MR image analysis have shown significant potential of deep learning for the development of clinically useful systems for computer-assisted medical image assessment. In addition, since analysis of these modalities has a direct impact on the final diagnosis and treatment planning, it is expected that deep learning will play a key role in the development of precision

medicine by the prediction of prognosis and survival for each patient. Therefore, efficient and accurate analysis techniques based on deep learning is becoming ever more important.

To this end, in this paper, we will introduce various use cases of deep learning for analyzing CT and MR images. We will also discuss about the opportunities, future directions and remaining challenges.

LESION DETECTION AND CLASSIFICATION

The most basic yet important task of radiologists is the assessment of exams by the detection and classification of specific patterns or lesions. Therefore, most computer-aided detection (CADe) and computer-aided diagnosis (CADx) systems have focused on improving the accuracy and productivity of the detection and classification tasks for medical images. Below, we will discuss the use of deep learning in the classification tasks of medical imaging in two levels: exam or image-level classification and lesion or region-level detection and classification.

1. Medical Image/Exam Classification

Image level classifications usually aim to determine the presence of disease or a specific pattern. The goal in medicine is to utilize the medical images and reports as inputs and outputs of the deep learning model, respectively, to automatically learn important features and biomarkers. For example, a 3D CNN model to determine the degree of Alzheimer's disease progression from structural brain MR images has been proposed [8]. In the study, they used a transfer learning method where features learned through pre-training a CAE with a small number of source domain images were used to fine tune the target domain data, which was then used to train the actual classification model. Their method achieved a better classification performance than a supervised training method directly using the target domain data.

In another study, a slice-level classification model was proposed to detect interstitial lung diseases (ILD) from chest CT scans [9]. Most of the existing ILD detection models require a patch-level manual annotation to train a patch-based classification model. However, because ILD lesions are inherently a mixture of various patterns and have unclear borders, it is very labor intensive to obtain a large amount of patch-based annotations. Hence, the proposed study used a CNN model pretrained with natural images to create slice level
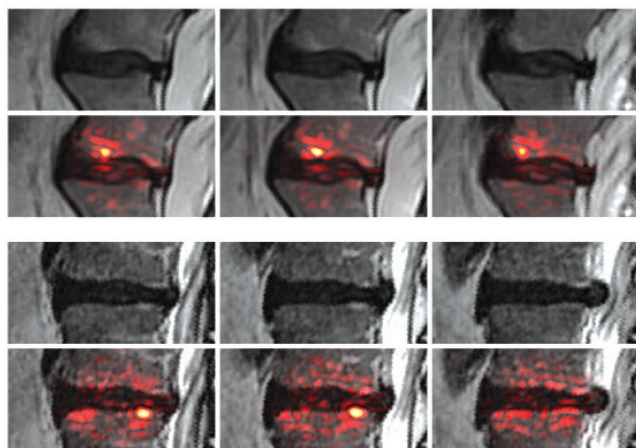
Fig. 1A. Visualization of 'evidence hotspot' in spinal MRI [10]. Adapted with permission

Fig. 1B. HIV patient and normal control   HIV patient and healthy control

predictions for each CT scan, and because multiple ILD subtypes can exist in a single scan, they used a multi-label classification loss which resulted in a high classification accuracy.

Recently, there have been studies focusing on visualizing the evidence of model predictions to overcome the limitations in explainability or interpretability, which has been pointed out as the weakness of deep learning models. Jamaludin et al. trained the multi-task learning model which detects several diseases simultaneously from spinal MR images and visualized salient regions in the image for corresponding predictions as 'evidence hotspots' as seen in Fig.1A [10]. This method visualizes the sensitivity of prediction by the change of input pixel value which can be computed as the partial derivative of model output with respect to every input pixel via back propagation. Although this model can visualize details at the pixel level and is relatively fast as it only requires single backpropagation, the output may not be intuitive and can be hard to understand because the salient pixels tends to be spread over a large area of the input image. To overcome this limitation, a visualization method based on prediction difference analysis which quantifies the difference between the output of the original image and the output when we marginalize out the region of interest, was proposed [11]. In essence, it is similar to the occlusion method in that they both look at the output difference as we apply changes to the region of interest [12]. However, while the naive occlusion method replaces the region of interest with simply zero values, the newly proposed method replaces the region with samples from the area surrounding the region of interest. This method produces a more intuitive and useful visualization of the
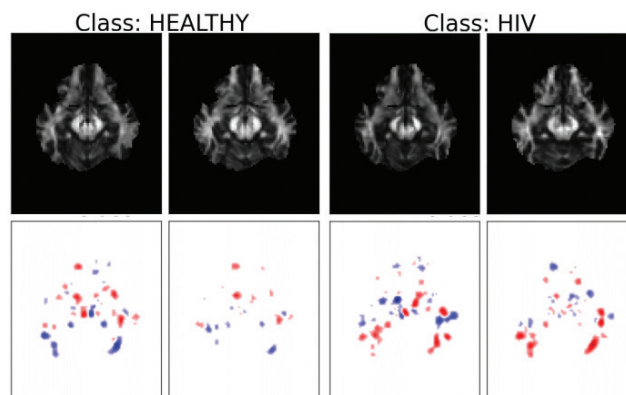
evidence base of a model that classifies HIV positive cases and healthy cases from brain MRIs as seen in Fig.1B.

2. Lesion Detection and Classification

Many methods have been proposed to overcome the fact that although we have an enormous repository of medical images, most of them lack expert annotations, which are very expensive and time consuming to produce. The most basic approach is to fine tune with a small set of labeled data after extracting relevant features from unlabeled data via unsupervised learning. Cheng et al. obtained a performance better than the current CADx system by pretraining a stacked denoising autoencoder (SDAE) to determine the malignancy of a lesion from chest CT images [13].

Another way to address the lack of annotations in medical images is to use CNN models pre-trained for natural images. Shin et al. tested several well-known CNN architectures and showed that such transfer learning approaches improve performance in CT patch based thoraco-abdominal lymph node detection and ILD classification [14].

In lung nodule detection, which is a major target for CADe systems, the task is divided into candidate detection and false positive reduction. While it is possible to achieve high sensitivity in the candidate detection step due to the well-known features in chest CT images, false positive reduction tends to be more challenging due to the variations in the nodule morphology and size. Hence, the false positive reduction stage is considered the determining stage of the final CADe performance. Dou et al. achieved a false positive rate
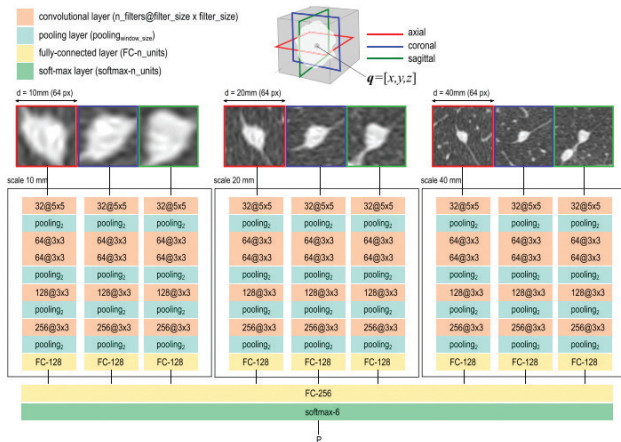
Fig. 2. Multi-stream, multi-scale CNN architecture for pulmonary nodule classification [16]. Adapted with permission.

lower than that of the conventional method through a multi-level contextual 3D CNN that discriminates nodules by fusing extracted features from multi-scale input patches [15].

Ciompi et al. proposed a single system that goes a step further than a CADe system by classifying nodules based on the morphology (solid, non-solid, part-solid, calcified, perifissural and spiculated) for automatic Lung-RADS reporting and malignancy estimation [16]. As seen in Fig 2, they performed data augmentation by extracting multiple view 2D patches from 3D nodule candidate voxels and used a multi-stream CNN architecture that fuses features extracted from multi-scale patches. The study showed a classifier performance better than that of support vector machines (SVM) based on hand-crafted features or features from unsupervised learning models. The result was within the inter-observer variability range of experts, hence providing a possibility of developing a Lung-RADS based automatic pulmonary nodule management system.

In a study by Ghafoorian et al., a CNN model was trained to detect lacunes from T1 and FLAIR brain MRI patches that are closely related to neurodegenerative disorders [17]. By changing the fully-connected layers of the trained patch-based model to convolution layers, they could effectively create a lacunes probability map of the whole brain MRIs. False positive reduction was then performed by training a multi-scale 3D CNN on the extracted candidate voxels. At this stage, they added the distance between the voxel and brain landmarks as contextual information to further enhance the false positive reduction and confirmed that the proposed CADe system improves diagnostic accuracy through an observer study.

LESION AND ANATOMICAL STRUCTURE SEGMENTATION

Segmentation in medical imaging is a crucial step in measuring the length or volume and assessing the morphology of an organ or lesion. It is not only an important research topic in itself but also an important process used for determination of region of interest and false positive reduction.

1. Anatomical Structure Segmentation

The goal of anatomical structure segmentation is to determine the presence or progression of disease through quantitative analysis of the volume and length data of segmented organs and organ substructures. For example, the size of the infant hippocampus is an important index for early brain development, however it is very difficult to accurately segment infant brains using the conventional predefined features used for adults because of poor tissue contrast on infant brain MR images. Guo et al. proposed a segmentation method where sparse patch matching was implemented based on features learned from the complementary information of T1 and T2 weighted MR images by a stacked auto-encoder [18].

Airway segmentation in thorax CT images can be used to support diagnosis by detection and quantification of bronchial wall thickening and changing lumen diameter. Airway segmentation can also increase the accuracy of the segmentation of other structures within the thorax and lower the false positive rate in lung nodule detection. Though there have been many proposals for airway segmentation, most studies have focused on achieving high sensitivity of detecting airways, which led to a high false positive rate [19]. To tackle this problem, Charbonnier et al. trained a CNN to classify leak candidates into airways and leaks from multiple initial airway segmentation results by changing the parameters of the segmentation algorithm [20]. By combining the classified results at the end, they could propose a method that successfully reduced false positives while maintaining airway detection sensitivity.

Organ segmentation in abdominal CT and MR images is also an active area of research. Hu et al. trained a 3D CNN model to accurately segment livers with large variations in shapes or those that have fuzzy borders with adjacent organs or lesions [21]. They proposed a method where probability map produced by the 3D CNN model is used as a shape prior to a global energy function.

For kidney segmentation in CT angiograms, Thong et al. trained a 2D patch based CNN to determine if the center pixel of the patch is part of the kidney [22]. By creating and interleaving pooling layers of various offsets, they could compensate the low-resolution feature maps of deep pooling layers and create a high-resolution probability map of the kidney.

Roth et al. proposed a two-stage method for the segmentation task of the pancreas [23]. During the detection stage, a 3D bounding box is created by the holistically-nested convolutional network on the axial, sagittal and coronal views of CT volumes. In the second stage, a more accurate final pancreas segmentation result is obtained by integrating the mid-level information of CNNs trained to segment the interior and boundary of the pancreas within the bounding box proposed in the first stage.

Yu et al. proposed a 3D volumetric CNN for prostate segmentation on MR images where they expanded the U-net, widely used for 2D biomedical image segmentation, into 3D and added residual connections to combine multiple-scale information [24, 25].

Lastly, Poudel et al. proposed using a recurrent full-convolutional network to segment the left ventricle from cardiac MR images [26]. Recurrent architectures model the spatial dependency between adjacent 2D short-axis slices, and this study showed improved segmentation results by sharing the high level global feature, which connects the encoder and decoder in a U-net structure, between these slices.

2. Lesion Segmentation

Lesion segmentation is an important step for planning treatment and predicting prognosis. The major challenge in lesion segmentation rises because multiple lesions of various shapes can be located anywhere within the organ. Furthermore, because most of the organ is a non-lesion space, the inequality of class distribution makes the segmentation task even more challenging. To overcome such challenges, various methods have been proposed.

Brosch et al. proposed a 3D encoder network for multiple sclerosis lesion segmentation in brain MR images [27]. They could accurately segment the lesion by fine tuning a U-net shaped encoder-decoder network based on an encoder pretrained by a stacked restricted Boltzmann machine.

In a study done by Ghafoorian et al., a network architecture based on multi-scale T1 and FLAIR MR patches was proposed to segment and quantify white matter hyperintensities, which are related to various brain disorders [28]. When they compared various structures that combine data from multiple scales, the multi-scale late fusion with weight sharing (MSWS) structure, which shares the CNN feature extraction weights and fuses the features for the final segmentation, showed the best performance. They also used the 3D coordinates and distance to major anatomical landmarks of each patch to provide anatomical prior knowledge. They showed that the use of such explicit spatial location features in addition to the contextual features of the patches can achieve a more accurate lesion segmentation.

Vivanti et al. proposed a method of liver tumor segmentation in follow up CT scans, which is composed of ROI (region of interest) selection based on deformable registration between the baseline and follow-up CT scans and a CNN trained to classify the voxels within the ROI as tumors and non-tumors. Finally, by removing segmentation leaks and holes, they could successfully segment tumors in follow up CT scans from baseline CT scans with high accuracy.

The recently trending deep learning method known as the generative adversarial network (GAN) has also been actively applied in the medical domain. GAN is a form of artificial neural network which consist of two sub-networks such as generator and discriminator. Two sub-networks are trained in adversarial manner such that the fake examples generated by generator are indistinguishable from the real examples while the discriminator tries to maximize its discrimination performance. While GAN is mostly used for the creation of synthetic images in the field of natural image processing, it has shown promising results in segmentation and conversion of medical images.
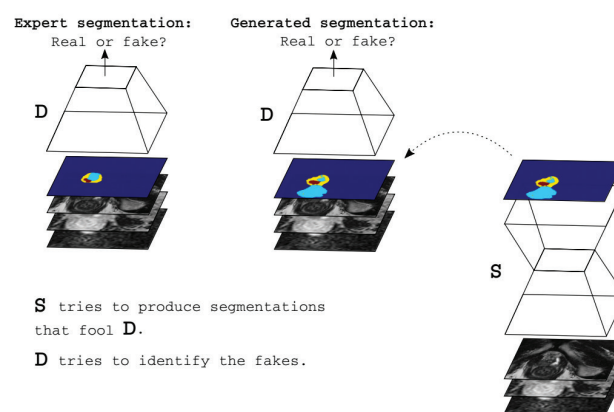

Expert segmentation: Real or fake? Generated segmentation: Real or fake?

S tries to produce segmentations that fool **D**.
D tries to identify the fakes.

Fig. 3. Segmentation of aggressive prostate cancer lesion using GAN [29]. Adapted with permission.

Kohl et al. proposed the GAN based segmentation method for the detection of aggressive prostate cancer in MR images [29]. As seen in Fig 3, a model is trained to distinguish between expert segmentations and model-generated segmentations. The results are then fed back to the training process of the generator. Through this process, the segmentation generator was trained to mimic the segmentations of experts and achieved a significant performance enhancement compared to the conventional segmentation methods.

IMAGE REGISTRATION

Another important area in medical image processing is image registration. Possible applications in the clinical environment include, but are not restricted to, multi-temporal analysis of the various phases of a contrast CT, multi-temporal analysis between follow up images and multimodal analysis in PET/CT images. The large size of these images and the fact that clinical implementation requires a very strict standard on accuracy make it difficult for us to utilize the conventional methods of image registration on high dimensional medical images. With the recent developments in deep learning, there has been various efforts to apply these learning methods in the field of image registration.

There are two main approaches to applying deep learning in image registration. One is using deep learning to estimate the similarity metric, which is then used to drive an iterative optimization strategy, as seen in Cheng et al. and Simonovsky et al. [30, 31]. In the study by Simonovsky et al., the problem is designed as a classification task, where a CNN is set to discriminate between alignment and misalignment of the two superimposed MRI brain images (T1 and T2 weighting of neonatal brains) [31]. The study of Cheng et al. is similar to this method in many ways, however, they used an autoencoder to pre-train the network [30].

Another approach is to use a deep regression network to directly predict transformation parameters between images. Miao et al. used such a method to directly predict the parameters of transformation between 3D CT images to 2D X-rays [32]. This method showed a significant improvement compared to the conventional intensity based method, where a digitally reconstructed radiograph is derived from the 3D image.

Lastly, there are also studies which focus on improving the speed of the conventional intensity based methods which tend to be very computationally burdening [33]. This study proposed a deep encoder-decoder network that jointly uses the similarity measure and the relationship between patches and deformation parameters. When their network was applied to the LDDMM (Large Deformation Diffeomorphic Metric Mapping) method for registration of brain MRI images, they could greatly decrease the computational burden while maintaining the mathematical properties of the LDDMM model. The proposed model was up to 36 times faster than the conventional optimization based methods.

The development of these various approaches to implement deep learning in the field of medical image registration shows a promising future. With better multi-temporal and multimodal image registration, we will be able to streamline the task of radiologists and clinicians by providing a direct and intuitive way of comparing medical images. Furthermore, accurate image registration will provide another class of medical data that can be utilized for various purposes, including conversion of images to different modalities and survival prediction.

IMAGE ENHANCEMENT AND SYNTHESIS

The enhancement or conversion of medical images are used to improve the accuracy of radiological reading or to utilize the information of different imaging modalities. As the image-to-image translation in natural images develops, there are increased efforts to implement such technologies to medical imaging.

1. Enhancement of Images

The quality of medical images is a very important factor in medical image analysis. However, to obtain high quality images, patients have to be exposed to high radiation doses or use extra expenses and time. That is why research in enhancing low dose CT images to the quality of normal dose CT images and low quality MR images to high quality images are being actively pursued.

Chen et al. proposed adding a residual skip connection to a U-net form encoder-decoder network to enhance the quality of low-dose CT scans to that of normal dose CT scans [34]. Here, the encoder reduces noise and artifacts and the decoder restores the structural information within the CT image. Also, the residual skip connection supplements the details lost while passing through multiple convolution and deconvolution layers to eventually generate an enhanced CT image, much better

than that of a conventional enhancement method.

In a study on enhancing 3T MR images to 7T MR images, Bahrami et al. proposed a method of using both appearance (intensity) and anatomical (brain tissue label) features [35]. They performed tissue segmentation on 3T MR images, and then they trained a 3D CNN that uses the intensity and segmentation labels from 3T MR images as the input and center voxel intensity of 7T MR image voxels as the output. This method not only produced an output very similar to 7T MRI quality but also worked robustly on images collected from different scanners.

Oktay et al. proposed a super resolution method to reconstruct high quality 3D cardiac images from 2D cardiac MR slices [36]. In the study, they used a CNN structure that uses multiple view stacks from both short axis and long axis as the input. Instead of directly reconstructing high resolution MR images, they allowed the network to efficiently learn the difference between low resolution and high resolution images through residuals, thereby producing a cardiac MR image of much higher quality than that of the conventional method.

2. Conversion to Different Modalities

With the objective of reducing the time and resource spent on extra exams and increasing accuracy of diagnosis and treatment plans, there have been various approaches to utilize deep learning in conversion between different image modalities.

Li et al. showed improvement in diagnostic accuracy of Alzheimer's using positron emission tomography (PET) images estimated from MR images [37]. They trained a 3D CNN model to estimate the corresponding PET voxel form MRI voxels and observed that the diagnostic accuracy was higher when the estimated PET images were used in addition to the MR images.

GAN can also be used in translation or conversion between images of different modalities, and Nie et al. used GAN to estimate CT images from MR images as seen in Fig. 4 [38]. Because the output CT image is blurry when we use a simple
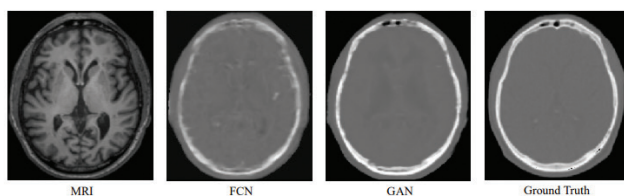
GAN loss to estimate CT images, they added a gradient difference loss function in the training process to maintain the intensity gradient between pixels in MRI. As a result, they obtained an image much sharper and closer to a real CT image than that of the conventional method.

SURVIVAL ANALYSIS

Survival analysis is another field of research where a deep learning based approach can predominate compared to traditional approaches, which were based on handcrafted features and limited sets of selected imaging modalities.

Nie et al. presented a model for survival prediction of high-grade glioma, applying a multi-channel 3D CNN model to automatically extract features from fMRI, DTI and T1 MRI images and a support vector machine for integrating non-image clinical features [39]. Although deeply-learned features showed a better result in predicting overall survival compared to hand crafted features, the best result was achieved by adopting both features. Researchers also proposed that features from fMRI and DTI images have more impact in functional, neurological and oncological applications.

Van der Burgh et al. incorporated clinical characteristics, structural connectivity and brain morphology based on high-resolution diffusion-weighted and T1-weighted MRI images to build a highly accurate model for survival analysis in amyotrophic lateral sclerosis (ALS) patients as shown in Fig. 5 [40].
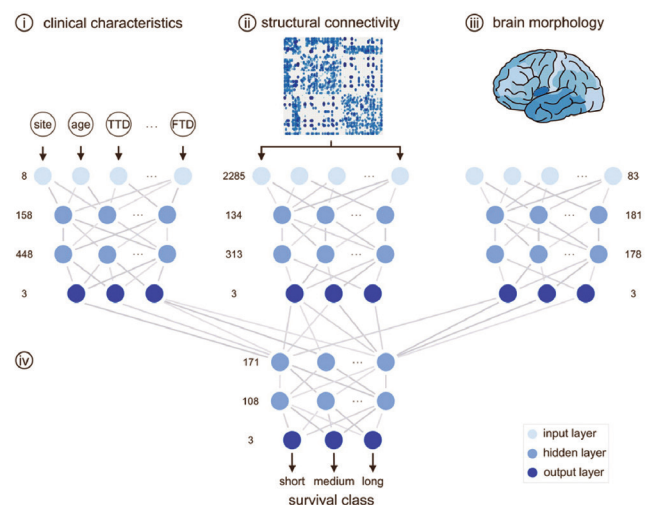

Fig. 4. Conversion of MRI to CT using GAN [38]. Adapted with permission

Fig. 5. Example of a model adopting deep learning for survival analysis [40]. Adapted with permission.

However, most of the earlier works present survival prediction into roughly divided groups. Long or short overall survival time, short, medium or long survivors, and future disease activity within two years [39, 40, 41]. These works have their limitation from its retroscopic nature and impractical categories of survival prediction.

Recently, Oak den-Rayner et al. implemented conceptual experiments on survival analysis to prove the concept of using features extracted from the medical image as a biomarker, especially from routine CT images [42]. This research suggests that a set of images could be solely manipulated and applied into a CNN to predict mortality, without any restrictions on specific disease entity or organ region. Even with small datasets and methods without novelty, the results show promising potential for future research on radiomics adopting deep learning.

CONCLUSION AND PERSPECTIVES

In this paper, we have reviewed examples of the deep learning based CT and MRI image analysis for various purposes. Beside these topics, there are many other examples: a methodology to estimate the uncertainty of the prediction that artificial intelligence model generates; principled way to integrate clinical or medical knowledge into the model training; development of content-based case retrieval system that searches images of similar diseases or conditions; efficient analysis of higher-dimensional medical image, such as contrast enhanced images or follow-up images; and research on the privacy and security related to medical images when training and implementing artificial intelligence models [28, 43, 44, 45, 46]. These topics have not been thoroughly studied, and larger-scale studies are required. In addition, considering that it is relatively difficult and time-consuming to generate and collect lesion-level annotation for high dimensional images, development of computer-assisted annotation tools and standardized protocols for lesion labeling are the most urgent topics to be studied.

However, even widely used off-the-shelf technologies can be combined with a well-defined problem and high-quality data to bring the artificial intelligence based medical image analysis technologies into clinical practice. In the near future, artificial intelligence technology will not only handle medical images but also integrate and analyze various patient health information and genome information to achieve reduction of medical expenditure and improve the quality of life of patients through early detection of disease and prediction of prognosis and survival. Therefore, assuming the gatherings and collaborations among hospitals, companies, and clinical and artificial intelligence researchers will become more common, the implementation of data-driven precision medicine and its dissemination is highly likely to take place.

REFERENCES

1. Litjens G, Kooi T, Bejnordi BE, et al. A Survey on Deep Learning in Medical Image Analysis. arXiv [Internet] 2017;1702.05747:1–34. Available from: http://arxiv.org/abs/1702.05747/.

2. Lot SB, Lint JJ, Freedmant MT, Munt SK. Computer-Assisted Diagnosis of Lung Nodule Detection Using Artificial Convolution Neural Network. Med Imaging 1993;1898:859–69.

3. Chan HP, Lo SC, Sahiner B, Lam KL, Helvie M a. Computer-aided detection of mammographic microcalcifications: pattern recognition with an artificial neural network. Med Phys 1995;22:1555–67.

4. Imagenet Challenge. Available from: http://image-net.org/challenges/LSVRC/2012/results.html

5. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis Detection in Breast Cancer Histology Images using Deep Neural Networks. Proc Med Image Comput Comput Assist Interv 2013;411–8.

6. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 2016;304:649–56.

7. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8.

8. Hosseini-Asl E, Keynton R, El-Baz A. Alzheimer's disease diagnostics by adaptation of 3D convolutional network. ICIP 2016;126–30.

9. Gao M, Xu Z, Lu L, Harrison AP, Summers RM, Mollura DJ. Holistic Interstitial Lung Disease Detection using Deep Convolutional Neural Networks: Multi-label Learning and Unordered Pooling. arXiv [Internet] 2017;9352:1–9. Available from: http://arxiv.org/abs/1701.05616/.

10. Jamaludin A, Kadir T, Zisserman A. SpineNet: Automatically pinpointing classification evidence in spinal MRIs. Vol. 9901 of Lect Notes Comput Sci 2016:166-75.

11. Zintgraf L, Cohen T, Adel T, Welling M. Visualizing Deep Neural Network Decisions: Prediction Difference Analysis. arXiv [Internet] 2016;1511.06488:1–12. Available from: http://arxiv.org/abs/1511.06488/.

12. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. Comput Vis ECCV 2014:8689:818–33.

13. Cheng J-Z, Ni D, Chou Y-H, et al. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. Sci Rep 2016;6:24454.

14. Shin HC, Roth HR, Gao M, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging 2016;35:1285–98.

15. Dou Q, Chen H, Yu L, Qin J, Heng PA. Multi-level Contextual 3D CNNs for False Positive Reduction in Pulmonary Nodule Detection. IEEE Trans Biomed Eng 2016;99.

16. Ciompi F, Chung K, van Riel SJ, et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci Rep 2017;7:46479.

17. Ghafoorian M, Karssemeijer N, Heskes T, et al. Deep multi-scale location-aware 3D convolutional neural networks for automated detection of lacunes of presumed vascular origin. NeuroImage Clin 2017;14:391–9.

18. Guo Y, Wu G, Commander LA, et al. Segmenting hippocampus from infant brains by sparse patch matching with deep-learned features. Vol. 8674 of Lect Notes Comput Sci 2014:308-315.

19. Lo P, Van Ginneken B, Reinhardt JM, et al. Extraction of airways from CT (EXACT'09). IEEE Trans Med Imaging 2012;31:2093–107.

20. Charbonnier JP, Rikxoort EM va., Setio AAA, Schaefer-Prokop CM, Ginneken B van, Ciompi F. Improving airway segmentation in computed tomography using leak detection with convolutional networks. Med Image Anal 2017;36:52–60.

21. Hu P, Wu F, Peng J, Liang P, Kong D. Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. Phys Med Biol 2016;61(24):8676–98.

22. Thong W, Kadoury S, Piché N, Pal CJ. Convolutional networks for kidney segmentation in contrast-enhanced CT scans. Comput Methods Biomech Biomed Eng Imaging Vis 2016;1163:1–6.

23. Roth HR, Lu L, Farag A, Sohn A, Summers RM. Spatial aggregation of holistically-nested networks for automated pancreas segmentation. Vol. 9901 of Lect Notes Comput Sci 2016:451-459.

24. Yu L, Yang X, Chen H, Qin J, Heng P-A. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images. Thirty-First AAAI Conference on Artificial Intelligence 2017;66–72.

25. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI 2015:234–41.

26. Poudel RPK, Lamata P, Montana G. Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. Vol. 10129 of Lect Notes Comput Sci 2017:83–94.

27. Brosch T, Tang LYW, Yoo Y, Li DKB, Traboulsee A, Tam R. Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation. IEEE Trans Med Imaging 2016;35(5):1229–39.

28. Ghafoorian M, Karssemeijer N, Heskes T, et al. Location Sensitive Deep Convolutional Neural Networks for Segmentation of White Matter Hyperintensities. Sci Rep 2017;7:5110.

29. Kohl S, Bonekamp D, Schlemmer H-P, Yaqubi K, Hohenfellner M, Hadaschik B, et al. Adversarial Networks for the Detection of Aggressive Prostate Cancer. arXiv [Internet] 2017;1702.08014:1–12. Available from: http://arxiv.org/abs/1702.08014/.

30. Cheng X, Zhang L, Zheng Y. Deep similarity learning for multimodal medical images. Comput Methods Biomech Biomed Eng Imaging Vis 2016;1163:1–5.

31. Simonovsky M, Gutiérrez-Becker B, Mateus D, Navab N, Komodakis N. A Deep Metric for Multimodal Registration. arXiv [Internet] 2016;1609.05396:1–10. Available from: http://arxiv.org/abs/1609.05396/.

32. Miao S, Wang ZJ, Liao R. A CNN Regression Approach for Real-Time 2D/3D Registration. IEEE Trans Med Imaging 2016;35(5):1352–63.

33. Yang X, Kwitt R, Styner M, Niethammer M. Fast Predictive Image Registration. arXiv [Internet] 2017;1703.10902:1–10. Available from: http://arxiv.org/abs/1703.10902/.

34. Chen H, Zhang Y, Kalra MK, et al. Low-Dose CT with a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN). IEEE Trans Med Imaging 2017;1–1.

35. Bahrami K, Shi F, Rekik I, Shen D. Convolutional Neural Network for Reconstruction of 7T-like Images from 3T MRI Using Appearance and Anatomical Features. MICCAI 2016 DL workshop 2016;39–47.

36. Oktay O, Bai W, Lee M, Guerrero R, Kamnitsas K, Caballero J, et al. Multiinput cardiac image super-resolution using convolutional neural networks. Vol. 9902 of Lect Notes Comput Sci 2016:246–254.

37. Li R, Zhang W, Suk H-I, et al. Deep learning based imaging data completion for improved brain disease diagnosis. Med Image Comput Comput Assist Interv 2014;17:305–12.

38. Nie D, Trullo R, Petitjean C, Ruan S, Shen D. Medical Image Synthesis with Context-Aware Generative Adversarial Networks. arXiv [Internet] 2016;1612.05362:1–11. Available from: http://arxiv.org/abs/1612.05362/.

39. Nie D, Zhang H, Adeli E, Liu L, Shen D. 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci 2016:212–220.

40. Van der Burgh HK, Schmidt R, Westeneng HJ, de Reus MA, van den Berg LH, van den Heuvel MP. Deep learning predictions of survival based on

MRI in amyotrophic lateral sclerosis. NeuroImage Clin 2017;13:361–9.

41. Yoo Y, Tang LW, Brosch T, Li DKB, Metz L, Traboulsee A, Tam R. Deep learning of brain lesion patterns for predicting future disease activity in patients with early symptoms of multiple sclerosis. Deep Learn Data Label Med Appl (2016). Vol. 10008 of Lect Notes Comput Sci 2016:86–94.

42. Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley AP, Palmer LJ. Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework. Sci Rep 2017;7:1648.

43. Leibig C, Allken V, Berens P, Wahl S. Leveraging uncertainty information from deep neural networks for disease detection. bioRxiv;2016;084210:1–21. Available from: http://www.biorxiv.org/content/early/2016/10/28/084210/.

44. Burak Akgül C, Rubin DL, Napel S, Beaulieu CF, Greenspan H, Acar B. Content-based image retrieval in radiology: Current status and future directions. J Digit Imaging 2011;24:208–22.

45. Shokri R, Shmatikov V. Privacy-preserving deep learning. 53rd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2015;909–910.

46. Papernot N, Abadi MM, Erlingsson Ú, Goodfellow I, Talwar K. Semi-supervised knowledge transfer for deep learning from private training data. arXiv [Internet] 2016;1610.05755:1–16. Available from: http://arxiv.org/abs/1610.05755/.