ANNALS OF THE UNIVERSITY OF CRAIOVA
Series: Automation, Computers, Electronics and Mechatronics, Vol. 14 (41), No. 1, 2017

Overview of Deep Learning in Medical Imaging

Popa(Ciurezu) Didi Liliana*, Faiq Baji, Popa Radu Teodoru*****

**drd at University of Craiova, Faculty of Automation, Computers and Electronics,*
(e-mail: lilianaciurezu@yahoo.com)
**drd at University of Craiova, Faculty of Automation, Computers and Electronics,*
(e-mail: faiq.baji2015@gmail.com)
***dr. Ing at University of Craiova, Faculty of Automation, Computers and Electronics*
(e-mail: tpopa@software.ucv.ro)

Abstract: Deep learning have gained lately popularity by achieving very good results for recognizing objects such as cars, plants, coffee cups in images. Big companies like Facebook, Google, Amazon - are already using these methods to identify faces, recognize voice commands and even enable self-driving cars. Deep learning is based on classical neural networks and represents a method of machine learning and has evolved over the years to become a research field on its own. Deep neural networks are based on different models: Stacked Auto Encoder ,Deep Belief Networks, Deep Boltzmann Machine ,Convolutional Neural Networks, Recurrent Neural Networks. Most deep learning researchers are not programming neural networks directly but, they are using software libraries like: TensorFlow, Caffe2, Theano, Torch, etc. Deep learning is a central method for developing new applications in medical sector. Medical sector has access to vast quantities of patient data and images can be fed in the deep learning neural networks algorithms to learn from. In medical image analysis many types of deep architectures have been applied .In the field of Convolutional Networks there are several architectures. The most common are: LeNet, AlexNet, GoogLeNet, ZfNet, VggNet, ResNet. Today, deep learning networks can execute a lot of tasks in medical field, especially medical imaging. These network can solve problems like: Classification, Regression and Segmentation but they need a lot of data to train deep models and also need powerful hardware to train the deep networks. In this paper are discussed briefly the latest methods for medical imaging currently in research: Blood vessel detection in ultrasound, Classification of skin cancer close to dermatologist level with deep neural networks, Deep CNNs for Diabetic Retinopathy Detection, Deep Learning for large-scale drug screening, Deep Learning Commercial Applications. In conclusion, deep learning has a great potential impact in changing world.

Keywords: Deep learning, convolutional neural networks, medical imaging, machine learning, medical classification

1. INTRODUCTION

In the last years, deep learning have gained popularity by achieving very good results for recognizing objects such as cars, plants, coffee cups in images. Big companies like Facebook, Google, Amazon - are already using these methods to identify faces, recognize voice commands and even enable self-driving cars.

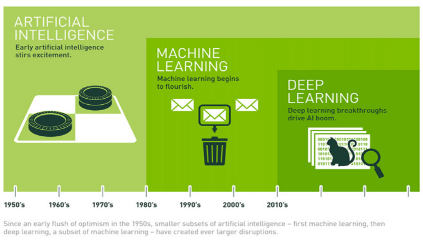


Fig. 1: Evolution of artificial intelligence source: www.nvidia.com/deep-learning-ai

Deep learning is a very interesting domain at the present time and many researchers are working in the direction of developing this domain. It first started as a branch of Machine Learning (ML), which is part of Artificial Intelligence (AI) domain.

The term Deep Learning was introduced to the machine learning community by Rina Dechter in 1986 and in 2000 to Artificial Neural Networks by Igor Aizenberg and colleagues, in the context of Boolean threshold neurons. Industrial applications of deep learning to large-scale speech recognition started around 2010.

Deep Learning works with big data. This means that important resources are necessary, especially for training a network. A new revitalization of the domain is related to the development of games industry, around 2015. The advancements of high-tech central processing units

(CPUs) and graphics processing units (GPUs) for fast and cheap parallel computing. The GPU that was needed to solve the matrix math for rendering graphics, was the necessary type for computations of deep learning.

Deep learning is an improvement of artificial neural networks, consisting of more layers that permit higher levels of abstraction and improved predictions from data (Y. LeCun, 2015)

Deep learning allows the computer to build complex concepts out of simpler concepts. (Ian Goodfellow, 2016)

Even from the very first steps of the development of neuronal networks, the researches have tried to apply it to medical images. In practice, most interpretations of medical images are performed by physicians; however, image interpretation by humans is limited since humans suffer from factors like fatigue and distractions. In order to reduce the misdiagnose putted by a doctor and because in some parts there are no doctors, the interest in the automation of a diagnose grew. The topic -Deep Learning in medical images, is now one of the most important at major conferences and a special issue appeared of IEEE Transaction on Medical Imaging in May 2016 (Greenspan et al. (2016))

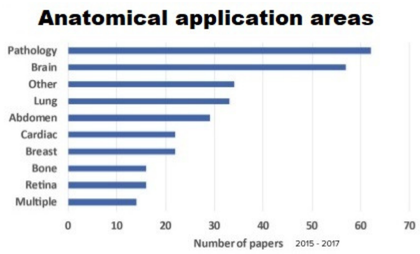


Fig. 2: DL papers published 2015-2017 -graphical representation regarding anatomical application areas - source Litjens at al 2017

2. ARTIFICIAL NEURAL NETWORKS

The neural network are simulating the way that human brain is interacting in the learning process like in Fig. 3.

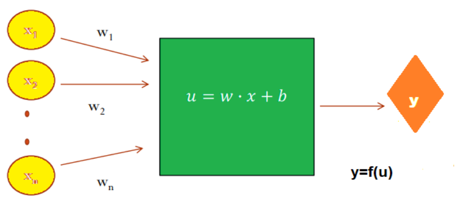


Fig. 3: Artificial neuron

Where the components are:
x – input vector
y – output
w – weights
b – bias
u – linear combination
f – activation function

2.1 Feedforward neural networks (FF)

In a neural network there is a '*hidden*' layer between the input layer and the output layer. The units of the neighbouring layers are fully connected to each other, see Fig 4.

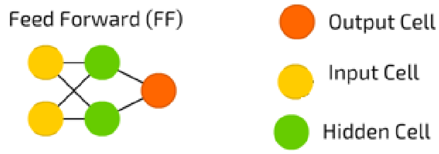


Fig 4: Feedforward Network

A feed forward network approximate the function f. It is defined a mapping $y = f(x;\theta)$ and learns the value of the parameters θ that result in the best function approximation.(Goodfellow-et-al-2016)

3. DEEP NEURAL NETWORKS

When a network has many layers it is often called 'deep' or a deep neural network (DNN).

Regarding deep neural network, there are some basic deep models in literature.

3.1 Stacked Auto Encoder (SAE)

SAE is a type of two-layer neural network that learns a latent or compressed representation of the input, minimizing the reconstruction error between the input and the output values of the network. While a single-layer auto-encoder is limited, when stacking multiple auto-encoders, by taking the activation values of hidden units of an auto-encoder as the input to the following upper auto-encoder, it is possible to improve the representational power greatly. (Bengio Y, 2007).

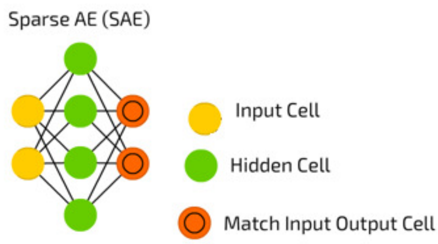


Fig. 5: Stacked Auto Encoder

An important characteristic of SAE is to learn nonlinear patterns. Backpropagation is applied with gradient optimization techniques, first with an aleatory initialization. For a better performance it is applied a greedy layer. This pre training technique is done in an unsupervised manner.

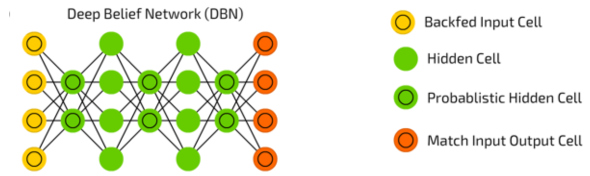


Fig. 6: Deep Belief Network

3.2 Deep Belief Networks (DBN)

In 2010, Hinton described in a paper, "Restricted Boltzmann Machine (RBM) as a type of Markov Random Field (MRF), constituting an input layer or visible layer $x = (x_1, x_2, \ldots, x_N)$ and a hidden layer $h = (h_1, h_2, \ldots, h_M)$ that carries the latent feature representation". The connections between the nodes are bidirectional, given an input vector x one can obtain the latent feature representation h, but also vice versa. As such, the RBM is a generative model, meaning we can sample from it and generate new data points coming from the distribution on which it is trained. (Geert Litjens, 2017)

DBNs are Stacked auto encoders in which there are RBMs instead of Auto encoders layers. (Hinton et al. (2006); Bengio et al. (2007))

The training of the individual layers are done in an unsupervised way. In the final fine-tuning is added a linear classifier at the top layer of the DBN and add a supervised optimization.

3.3 Deep Boltzmann Machine (DBM)

A deep Boltzmann machine, as described by Salakhutdinov et. in 2010, it contains a set of visible units $v \in \{0, 1\}^D$, and a sequence of layers of hidden units $h^1 \in \{0, 1\}^{F1}$, $h^2 \in \{\{0, 1\}^{F2}$,..., $h^L \in \{0, 1\}^L$. There are connections only between hidden units in adjacent layers, as well as between the visible units and the hidden units in the first hidden layer.

All the layers in DBM are unidirectional generative model after stacking Restricted Boltzmann Machine (RBMs), which is a big difference of DBN.

Example of Algorithm 1 Greedy Pretraining Algorithm for a Deep Boltzmann Machine with 3-layers.(Ruslan Salakhutdinov et. al 2010)

1: Make two copies of the visible vector and tie the visible-tohidden weights W^1. Fit W^1 of the first layer RBM to data.

2: Freeze W^1 that defines the first layer of features, and use samples h^1 from $P(h^1 | v, 2W^2)$ as the data for training the next layer RBM with weight vector $2W^2$.

3: Freeze W^2 that defines the second layer of features and use the samples h^2 from $P(h^2 | h^1, 2W^2)$ as the data for training the third layer RBM with weight vector $2W^3$.

4: When learning the top-level RBM, double the number of hidden units and tie the visible-to-hidden weights W^3.

5: Use the weights $\{W^1, W^2, W^3\}$ to compose a Deep Boltzmann Machine.

3.4 Convolutional neural networks CNNs

The most successful models for image analysis are convolutional neural networks (CNNs). Regarding medical images CNNs were applied in 1996 by Sahiner et al. The CNN consisted of an input layer, two hidden layers and an output layer and used backpropagation algorithm. At that time, the training times were huge.

To every neuron, we have inputs on which it performs a dot product and can follow it non-linearity. The entire network expresses a single differentiable score function: from the raw image pixels to class scores. Also they have a loss function (e.g. SVM/Softmax) on the last fully-connected layer.

Convolutional Neural Networks take into consideration that the input consists of images and they limit the architecture in a more subtle way. Therefor, not like a regular Neural Network, the layers of a CNN have neurons arranged in 3 dimensions: width, height, depth. (where depth - refers to RGB).The neurons in a layer are only connected to a small region of the layer before it, unlike all of the neurons in a fully-connected layer.

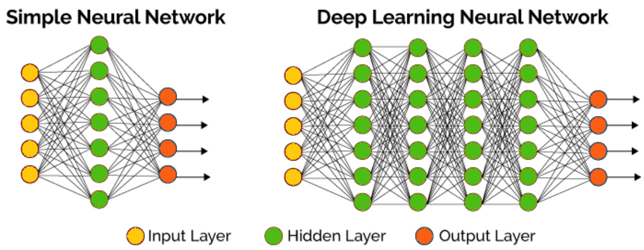


Figure 7: Simple neural network and deep learning neural network

There are some concepts regading CNNs : convolutional layer,pooling layer, Relu Layer

3.4.1 Convolutional Layer

summarize the Convolutional Layer by A. Karpathy in 2017

- Accepts a volume of size $W1 \times H1 \times D1$
- Needs four hyper-parameters:
  - Number of filters K,
  - their spatial extent F,
  - the stride S,
  - the amount of zero padding P.
- Creates a volume of size $W2 \times H2 \times D2$ where:

- $W2 = (W1 - F + 2P)/S + 1$
  $H2 = (H1 - F + 2P)/S + 1$ (i.e. width and height are computed equally by symmetry)
- $D2 = K$ With parameter sharing, it introduces □F□F□D1 weights per filter, for a total of (F□F□D1)□K weights and K biases.
- In the output volume, the depth slice (of size $W2 \times H2$) is the result of performing a valid convolution of the filter over the input volume with a stride of S, and then offset by bias.

A common setting of the hyper-parameters is $F = 3, S = 1, P = 1$

3.4.2 ReLU layer

ReLU is the abbreviation of Rectified Linear Units. This layer applies the non-saturating activation function f(x) = max (0,x). In this way the properties of non-linearity of the decision function is increases without affecting the receptive fields of the convolution layer.

There are other functions that are used to increase nonlinearity, like the sigmoid function $f(x) = (1 + e^{-x})^{-1}$.

ReLU is often preferred to other functions, because it trains the neural network several times faster without a significant penalty to generalisation accuracy according to Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012)

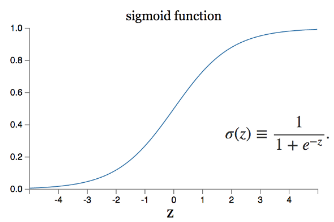


Figure 8: Sigmoid function

3.4.3 Pooling layer

Pooling layers means that the values of neighborhoods pixels are aggregated using functions like maxim .The pooling layer, as summarized by A. Karpathy in 2017:

- Accepts a volume of size $W1 \times H1 \times D1$
- Requires two hyperparameters:
  - their spatial extent F,
  - the stride S,
- Produces a volume of size $W2 \times H2 \times D2$ where:
  - $W2 = (W1 - F)/S + 1$
  - $H2 = (H1 - F)/S + 1$
  - $D2 = D1$
- Introduces no parameters since it computes a fixed function of the input
- in practice: A pooling layer may have $F = 3, S = 2$ (also called overlapping pooling), or $F = 2, S = 2$. Pooling sizes with bigger receptive fields are not used.
-

3.4.4 Fully connected layer

The high-level reasoning in the neural network is done via fully connected layers. Neurons have connections to all activation in the previous layer, just like in a regular neural networks.

3.5 Recurrent Neural Network

Recurrent neural network (RNN) is a class of neural network applied in temporal data including natural language processing, speech recognition, handwriting recognition, and generation tasks. The connections between units form a cycle with a one way direction. In this way it is obtained a dynamic temporal behaviour.

While a convolutional network is specialized for processing a grid of values X like an image and have structures for a feed-forward network, a recurrent neural network is a neural network specialized in processing a sequence of values x(1) , . . . , x(n). Convolutional networks can easily scale to images with large width and height, RNN can do the job that would be difficult for networks without sequence-based specialization. (Goodfellow-et-al-2016)

Fig. 9: Recurrent Neural Network

4. DEEP LEARNING- OPEN SOURCE FRAMEWORKS

In order to have access to hardware acceleration: CPUs, GPUs most deep learning researchers are not programming deep neural networks directly but, they are using software libraries (such as PyTorch or TensorFlow) that handle this. CUDA and OpenCL are the two main ways for programming GPUs. CUDA is proprietary language created by Nvidia, so it can't be used by GPUs from other companies. In order to use the libraries, you need access to the right type of GPU, which in most of cases means having access to GPU from the company Nvidia.

Deep learning is a relatively young field, and the libraries and tools are changing quickly.

Python is by far the most commonly used language for deep learning. Same of the available deep learning libraries are:

1. Tensorflow - is one of the most appreciated framework, made by the researchers at Google. According to tehnsorflow.org, TensorFlow™ is an open source software library for numerical computation using data

flow graphs. It is a low level library. It supports Python and C++, along to allow computing distribution among CPU, GPU (many simultaneous) and even horizontal scaling using open source remote procedure call (gRPC).

2. Theano - The interface is in Python, it has integrated Numpy and allows automatic function gradient computations, and this is why in its beginning was one of the most used for general purpose Deep Learning. Theano is low-level library and doesn't have multi-GPU support nor horizontal capabilities.

3. Caffe2 was developed by Facebook and improves Caffe in a series of directions like mobile deployment, new hardware support (in addition to CPU and CUDA), first-class support for large-scale distributed training and flexibility for future directions such as quantized computation. Caffe2 is built to excel at utilizing both multiple GPUs on a single-host and multiple hosts with GPUs.The difference between PyTorch and Caffe2 is that Pytorch is great for research, experimentation and trying out exotic neural networks, while Caffe2 is more oriented towards supporting more industrial-strength applications with a heavy focus on mobile, according to caffe2.ai.

4. Keras it's a very high-level library. According to keras.io it is capable of running on top of TensorFlow, Microsoft Cognitive Toolkit, Theano, or MXNet .In Keras you can build a Neural Network in just a few lines of code. The interface is in Python.

5. Lasagne wants to be at the same level as Keras. It is a library to build and train neural networks in Theano. Its mission was to abstract a bit the complex computation underlying to Deep Learning algorithms and to have a friendlier interface (in Python too). Keras has better documentation and is more complete.

6. Dsstne is a framework developed by Amazon, not meant for research but only for production, using GPU. It is not very popular.

7. Torch is was a specially known framework because it's used in Facebook Research and in DeepMind before being acquired by Google (after that, they migrate to Tensorflow). According to torch.ch, you can build arbitrary graphs of neural networks, and paralleled them over CPUs and GPUs in an efficient manner. Lua is the programming language .This is a disadvantage because most the most of Deep Learning is focused on Python.

8. PyTorch is an open source machine learning library for Python, based upon Torch, used for applications such as natural language processing. It was released in oct 2016, and was primarily developed by Facebook's and Uber. It has strong GPU acceleration and PyTorch provides two high-level features: tensor computation (like numpy) with strong GPU acceleration and deep Neural Networks built on a tape-based autograph system. (pytorch.org)

9. Microsoft Cognitive Toolkit, previously known as CNTK is a deep learning framework developed by Microsoft Research. It describes neural networks as a series of computational steps via a directed graph and it is written in C++.

10. PaddlePaddle - is a framework developed by Baidu used in computer vision,natural language understanding and deep learning.

11. Other used frameworks are MxNet (University of Washington, adapted by Amazon); DeepLearning4j (Skymind), Nnabla (Sony).

One distinction among deep learning libraries is whether they use dynamic or static computations (TensorFlow, allow for both). Dynamic computation mean that the program is executed in the order you wrote it, making the debugging easier. Static computation means that you build a structure for your neural network in advance, and then execute operations on it. PyTorch uses dynamic computation.

5. DEEP LEARNING ARCHITECTURES

Many types of deep architectures have been applied to medical image analysis.

In the field of Convolutional Networks there are several architectures. The most common are: LeNet, AlexNet, GoogLeNet, ZfNet, VggNet, ResNet.

5.1 LeNet-5

It was described by LeCun et al., in 1998. Convoution filters were 5x and were applied at stride 1. Subsampling (Pooling) layers were 2x2 and applied at stride 2 for example: CONV-POOL-CONV-POOL-FC-FC as described in Fei-Fei,2017.

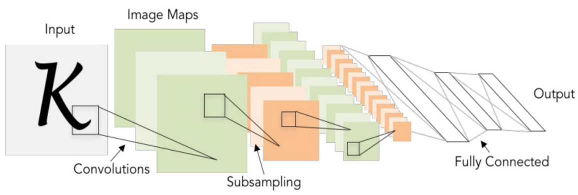


Fig. 10: LeNet-5 architecture

5.2 AlexNet

AlexNet is the first work that popularized CNN in Computer Vision.It was developed in 2012 by Alex Krizhevsky, Ilya Sutskever and Geoff Hinton. The Network was similar to LeNet architecture, with the difference that it has deeper, bigger, and featured Convolutional Layers connected on top of each other, while previously it was common to only have a single convolutional CONV layer always immediately followed by a POOL layer.

It had 8 layers: conv-max pool-norm-conv-max pool-norm-conv-conv-conv-max pool-fc-fc-fc

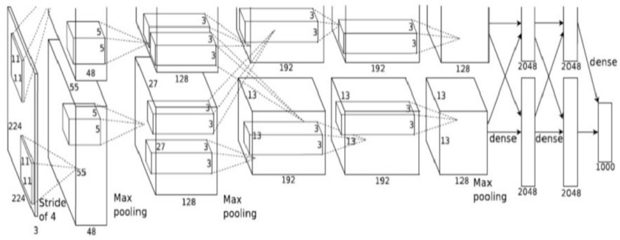


Fig. 11: AlexNet architecture (A.Krizhevsky 2012- AlexNet architecture)

<u>First layer of convolution</u> 96 11x11 filters applied at stride 4.For first layer it had 35000 Parameters $(11*11*3)*96$.

Layer 1 is a Convolution Layer, Input Image size is – $224 \times 224 \times 3$

Number of filters – 96 Filter size – $11 \times 11 \times 3$ Stride – 4 Layer 1 Output 2 $24/4 \times 224/4 \times 96 = 55 \times 55 \times 96$ (because of stride 4)

<u>Second layer of pooling</u>: 3x3 filters applied at stride 2

Details: - first use of ReLU - used Normalization layers - heavy data augmentation - dropout 0.5 - batch size 128 - SGD Momentum 0.9 - Learning rate 1e-2, - L2 weight decay 5e-4.Layer 2 is a Max Pooling Followed by Convolution

Input – $55 \times 55 \times 96$; Max pooling – $55/2 \times 55/2 \times 96 = 27 \times 27 \times 96$; Number of filters – 256

Filter size – $5 \times 5 \times 48$; Layer 2 Output $27 \times 27 \times 256$; Split across 2 GPUs – So $27 \times 27 \times 128$ for each GPU

Pooling is a sub-sampling in a 2×2 window (usually). Max pooling is max of the 4 values in 2×2 window. The intuition behind pooling is that it reduces computation & controls overfitting.

Layers 3, 4 & 5 follow on similar lines, as described in (Fei-Fei, 2017) .The fully-connected layers have 4096 neurons each.(A.Krizhevsky 2012)

For some year, was the first choice in implementing a CNN- for its fastness.

5.3 ZFNet

Zeiler and Fergus in 2013 describes why is superior to AlexNet by improving AlexNet with following: CONV1: change from (11x11 stride 4) to (7x7 stride 2) CONV3, 4,5: instead of 384, 384, 256 filters use 512, 1024, 512

It works the same as AlexNet but: CONV1: change from (11x11 stride 4) to (7x7 stride 2) CONV3, 4,5: instead of 384, 384, 256 filters use 512, 1024, 512

5.4 VGG Network

Was described by Simonyan and Zisserman in 2014. Stack of three 3x3 conv (stride 1) layers has same effective receptive field as one 7x7 conv layer.

Only 3x3 CONV stride 1, pad 1 and 2x2 MAX POOL stride 2.

Fig. 12:VGG16 Architecture

Details of the VGG Network are: -it was second in classification ILSVRC'14 and first in localization and had similar training procedure as Krizhevsky 2012 with No Local Response Normalisation (LRN) and used VGG16 or VGG19 (VGG19 only slightly better, more memory) and used ensembles for best results with FC7 features generalize well to other tasks, as described in (Fei-Fei,2017)

5.5 GoogLeNet

Implemented by [Szegedy et al., 2014];Has 22 layers and introduce for the first time the Efficient "Inception" module -.Also there are Full connected layers with only 5 million parameters, which is 12 times less than AlexNet - This CNN was ILSVRC'14 classification winner (6.7% top 5 error)

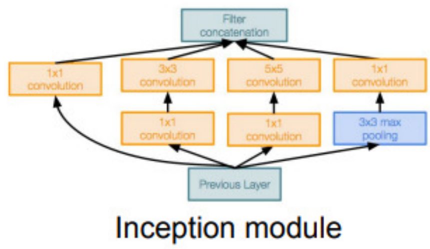


Fig. 13: The building block of GoogleNet

It was based on "Inception module" which meant designing a good local network topology (network within a network) and then stack these modules on top of each other.

Apply parallel filter operations on the input from previous layer: - Multiple receptive field sizes for convolution (1x1, 3x3, 5x5) - Pooling operation (3x3) Concatenate all filter outputs together depth-wise.Apply parallel filter operations on the input from previous layer: - Multiple receptive field sizes for convolution (1x1, 3x3, 5x5) - Pooling operation (3x3) Concatenate all filter outputs together depth-wise,as described in (Fei-Fei,2017)

5.6 ResNet

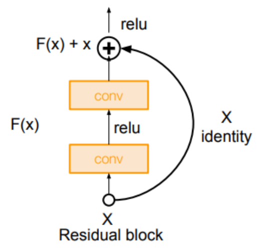


Fig. 14: The residual connection in ResNet

Is a very deep networks using residual connections .First implemented by He et al., 2015. It has 152-layer model for ImageNet and it was ILSVRC 2015 winner in classification (3.57% top 5 error).

VGG, GoogLeNet, ResNet all in wide use and are available to researcher ResNet current best default and there is a trend towards extremely deep networks

6. DEEP LEARNING TEHNIQUES FOR MEDICAL IMAGING

Deep learning techniques have recently been expamded for medical image analysis.In segmentation and registration, in different applications have been obtained interesting results.

Medical images can be obtained using ultrasound, Electrocardiogram (EKG), Computer Tomograf (CT), Magnetic Resonance Imaging (MRI), Mammography. The format for the medical software is: DICOM, EKG, and HT7. OpenCL, OpenGL,VTK, ITK, DicomTK are some of the libraries used. For segmentation there are different techniques like - Region Growing, Watershed, Levelset, Optical flow. Extraction of a blood vase from an image, means using convolution and gradients. For example, dates from an EKG can be extracted using Fourier analyze, Wavelet transform.

In the following section will be discussed briefly the latest methods for medical imaging currently in research.

6.1 Blood vessel detection in ultrasound

Smistad in 2017 describes a new method using a deep convolutional neural network for detecting blood vessels in echographic ultrasound images.They used a blood vessel segmentation method which performs an ellipse fitting at each pixel in the image using a graphic processing unit (GPU).

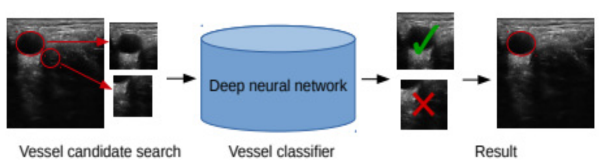


Fig. 15: Blood vessel Segmentation method (Smistad, 2017)

But their method cannot distinguish blood vessels from non-vessels and was only made to detect a single vessel for each image.In their method the output of blood vessel segmentation is passed on to a deep neural network (AlexNet)classifier which determines if the region contains a vessel or not and provides position and size..They used Caffe framework for training and testing of a deep convolutional neural network classifier. The AlexNet network had two convolution-pooling stages with 9 and 32 convolutions respectively. This is followed by three fully connected (FC) layers with dropout to reduce overfitting. A local response normalization (LRN) is performed after the first convolution. Rectified linear units (ReLU) are used as non-linear activation units in all stages.

Their method was able to find vessels in each image in less than 100 ms to be able to process the ultrasound image stream in real-time. The vessel candidate search, subimage creation and resizing were all implemented on the GPU. Caffe was run in GPU mode.

6.2. Dermatologist-level classification of skin cancer with deep neural networks

In 2017 a research Esteva and his team, demonstrated classification of some skin lesions by using a single CNN, which was trained from images, using pixels and disease labels as inputs. They trained a deep neural convolutional network using a dataset which consisted of 129,450 clinical images consisting of 2,032 different diseases. They compared its performance to 21 board-certified dermatologists on biopsy-tested clinical images with two important binary classification cases: keratinocyte carcinomas vs benign seborrheic keratoses; and malignant melanomas vs benign nevi. They utilized for pattern recognition a GoogleNet Inception v3 CNN architecture that was pre-trained on approx. 1.28 million images (1,000 object categories) from the “ImageNet Large Scale Visual Recognition Challenge 6 " in 2014 and train it on their dataset using transfer learning.

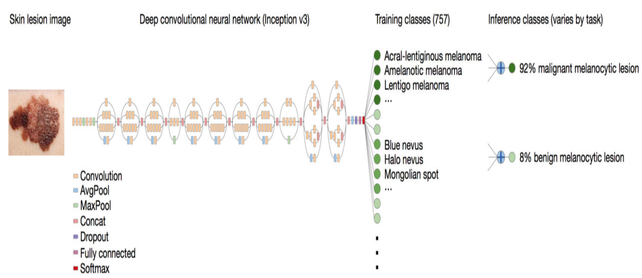


Fig. 16: Inception v3 CNN architecture (Esteva, 2017)

The CNN was trained by using 757 disease classes. Their data-set is composed of "dermatologist- labeled images organized in a structured tree taxonomy of 2,032 diseases, in which the individual diseases make the leaf nodes."

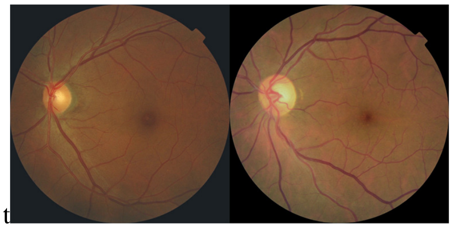
The input images consisted of 18 clinician-acquired, open-access online repositories, also from clinical datasets from Stanford University Medical Center.

They validated the effectiveness of the algorithm by two methods, with nine-fold cross-validation. First, they validated the algorithm using a 3-class disease partition as the first-level nodes of the taxonomy, which represent benign lesions, malignant lesions and non-neoplastic By this validation the CNN achieves a "maximum of 72.1 ± 0.9% (mean+sd) overall accuracy (the average of individual inference class accuracies) and two dermatologists attained 65.56% and 66.0% accuracy on a subset of the validation set.”

Second, they validated the algorithm using a nine-class disease partition—the second-level nodes—so that the diseases of each class have similar medical treatment plans. The CNN achieved 55.4 ± 1.7% overall accuracy where the same two dermatologists attained 53.3% and 55.0% accuracy. A CNN trained on a finer disease partition seems to perform much better .They conclude that" the method is mainly constrained by data input and can classify many visual conditions if sufficient training examples exist." Also they claim that Deep learning doesn’t care about the type of image data that is used and could be adapted to other specialties like otolaryngology, radiology, ophthalmology and pathology.

6.3 Deep CNNs for Diabetic Retinopathy Detection

In a competition organized by Kagle in 2015, were more teams of researchers try to o classify the severity of diabetic retinopathy in the tested images of retina,, Bogucki and his team- a participant in this competition describes the implementation of an CNN for diabetic retinopathy classification.



t

Fig. 17: Ophthalmologic images, left image has no marks of diabetic retinopathy, while the other has severe retinopathy (Bogucki, 2015)

According to (Bogucki, 2015) "Diabetic retinopathy is the leading disease that causes blindness in the active population of the developed world and is estimated to involve 93 million people."

"The contestants were given as input over 35,000 images of retinas, each having a severity rating. There were 5 severity classes, and the distribution of classes was almost imbalanced. Most of the images showed no signs of the disease. Only a few percent had the two most severe ratings. The contest begun in February 2015, and over 650 teams took part in it, fighting for the prize pool of \$100,000." (Bogucki, 2015)

The metric with which the predictions were rated was a quadratic weighted kappa.

Their solution consisted of multiple steps:

1. Image preprocessing (Crop the image, Scale it to 256×256, histogram normalization

2. Training multiple deep CNNs (their best single model consisted of 9 convolutional layers.). All Convolutional layers had 3×3 kernel, stride 1 and padding 1.In all their convolutional layers they follow the convolutional layer by batch normalization layer and ReLu activations. They used max pooling. They trained the CNN using stochastic gradient descent with a momentum and multiclass log-loss as a loss function. Further, the learning rate was adjusted manually a few times during the training. They used a personalized CNN based on Theano and Nvidia with DNN.

3. Eye blending

4. Kappa score optimization

6.4 Deep Learning for large-scale drug screening

For drug discoveries, in 2017 Gilmer, described a method using machine learning. Drug discoveries involve predicting the properties of molecules.

One reason molecular data is so interesting from a machine learning is that one natural representation of a molecule is as a graph with atoms as nodes and bonds as edges. Models that can take into consideration inherent symmetries in data will tend to generalize better — part of the success of convolutional neural networks on images is due to their ability to incorporate prior knowledge. Their paper “Neural Message Passing for Quantum Chemistry" describes a model family of neural networks called Message Passing Neural Networks (MPNNs), which are defined abstractly enough to include many previous neural net models that are invariant to graph symmetries.

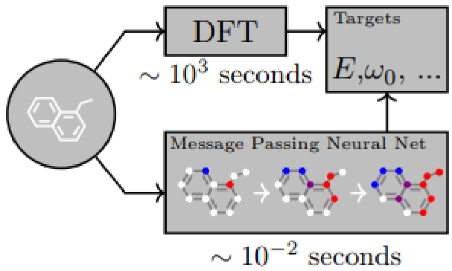


Fig. 18: A Message Passing Neural Network predicts quantum properties of an organic molecule by modelling a computationally expensive DFT calculation. (Gilmer, 2017)

They developed novel variations within the MPNN family which significantly outperform all baseline methods on the QM9 benchmark–a public collection of molecules paired with DFT (Density Functional Theory)-computed electronic, thermodynamic, and vibrational properties with improvements of nearly a factor of four on some targets.

7. DEEP LEARNING COMERCIAL APPLICATIONS FOR MEDICAL IMAGES

In medicine, there are so many variables it is difficult to always arrive at the correct diagnosis for people or machines. Sometimes AI can outperform doctors, especially in case of a rare disease.A doctor can see only some cases during his entire career. AI has the advantage of reviewing hundreds or even thousands of these rare studies from archives to become proficient at reading them and identify a proper diagnosis. Another thing in favor of deep learning applications is the fact that computers don't get tired like the humans. There are several applications for medical images.

Big or small companies from all over the world are developing applications using DL. Here are some examples:

- Enlitic (Australian company) is one of the first start up using DL for tumour detection based on algorithms capable of identifying relevant characteristics of lung tumours using CT and obtaining good accuracy rate.

- Arterys, a DL medical imaging technology company, recently partnered with GE Healthcare to combine its quantification and medical imaging technology with GE Healthcare's magnetic resonance (MR) cardiac solutions. The application is giving the possibility to view and quantify the blood flow inside the heart and putting a diagnosis of cardiovascular disease. (arterys.com)

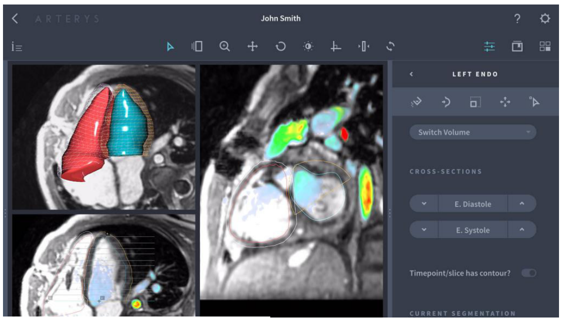


Fig 19: Arterry (artery.com)

The software detects in a fraction of the time of conventional cardiac MR scans.

- Lunit, a South Korean start up established in 2013, uses its DL algorithms to analyse and interpret X-ray and CT images.

- In India, Goolge is working with some hospitals for implementing DL-trained models at scale and helping doctors to detect DR early enough for a proper treatment.

- Microsoft's InnerEye initiative (started in 2010) is presently working on image diagnostic tools. The InnerEye uses algorithms such as Deep Decision Forests (as used already in Kinect and Hololens) as well as Convolutional Neural Networks (as available in CNTK) for the automatic, voxel-wise analysis of medical images, according to Microsoft.

- The "old" and known IBM Watson for detecting Cancer.

8. OPEN RESEARCH ISSUES AND FUTURE DIRECTIONS

The deep learning applications have their limitations regarding explanations. "The black box", as the deep learning are referred sometimes, cannot explain "how" it arrived at its predictions – even when they're correct. In health care, this is quite a problem because on one side are the doctors that wants to know how the computer came to the conclusion; one on the other hand are the patients that are not so willingly to get an operation or start a cure of chemotherapy based only on the decision of the computer taking into consideration their current situation and others image cases.

But sooner or later, the development of deep learning applications will affect every aspect of health care. Even if now some applications are just in the testing or collecting dates, this situation will change in few years. As we seen above, the applications are in different branches of medicine- radiology, cardiology, neurology, ophthalmology, oncology, abdominal image analysis. Machine learning software will serve as a very experienced clinical assistant, helping the doctor and making the medical act more efficient.

There is never an only way to solve a problem regarding machine learning- because there are many ways in preprocess dates and a lot of algorithms from which to decide. Sometimes the best results came from combining different approaches.

A problem that appeared is that in some deep learning applications is the need of data from medical past of patients, and from this is the problem of data securities- who is using them and when, the acceptance from the patients and so on.

Deep learning needs big data and the annotation of those in medical domain needs time, money and different experts.

Another problem that will need to be discussed is society, is that regarding the replacement of some doctors with machine learning, some will not be very happy to lose their jobs in favor of computers applications.

At some of the problems, solutions will come by facilitating close collaboration between doctors and DL researchers, and also between hospitals and researchers.

Many researchers and not only, are asking questions regarding the direction in which this field will evolve. At the recent "International conference on Medical Image Computing and Computer Assisted Interventions (MICCAI 2017)" some observations were made:

- the need of more research on generative methods, decision visualisation techniques, intelligent optimisation methods, semi- and weakly-supervised learning, and network architectures

- it is necessary to go even further and provide self-learning, workflow-aware, and robotic imaging solutions that are capable of imaging the right entity at the right time during an examination or a surgery, without any further assistance.

9. CONCLUSIONS

Deep learning is a central method for developing new applications in medical sector.

Medical sector has access to vast quantities of patient data and images can be fed in the deep learning neural networks algorithms to learn from. Sorting through massive amounts of big data is a major component of how deep neural networks can discover what is important for clinicians, what data elements are related to various diseases, therefor gaining clinical understanding.

At the present time, a lot of funds, from global companies or local companies, are directed throw this sector of medical research. Also the number of researchers in DL has grown exponentially in the last few years. If this current will be maintained, inevitability applications using DL will surround us in the years that will come.

Today, deep learning networks can solve some tasks in medical health research, especially medical imaging. The problems that these network can solve are: Classification, Regression and Segmentation but they need a lot of data to train deep models and also need powerful resources to train the deep networks.

Pre-trained models can achieve a really good performance in classification, regression for new drug discoveries.

Deep learning has a great potential impact in the developing world, is the way that things were done in medicine, education, transport, agriculture, and many other fields.

REFERENCES

Igor Aizenberg, Naum N. Aizenberg, Joos P.L. Vandewalle (2000). "Multi-Valued and Universal Binary Neurons: Theory, Learning and Applications". Springer Science & Business Media

Bengio Y, Lamblin P, Popovici D, Larochelle H. ," Greedy layer-wise training of deep networks." Proceedings of Neural Information Processing Systems (NIPS) 2007:153–160

Robert Bogucki et al,2015, "Diagnosing diabetic retinopathy with deep learning", 2015, blog.deepsense.ai/diagnosing-diabetic-retinopathy-with-deep-learning

Michael Coperland, 2016,"What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning?" https://www.nvidia.co.uk/deep-learning-ai/

Dechter, Rina 1986." Learning While Searching in Constraint-Satisfaction Problems." In Proceedings of the Fifth National Conference on Artificial Intelligence, 178–183. Menlo Park, Calif.: American Association for Artificial Intelligence

Esteva Andre , Kuprel Brett , Novoa, Roberto A., Ko Justin,Swetter Susan M.,Blau Helen , Thrun Sebastian "Dermatologist-level classification of skin cancer with deep neural networks",Nature 2017/01/25/online 542 pages 115-118, "Macmillan Publishers Limited", part of Springer Nature - http://dx.doi.org/10.1038/nature21056

Fei-Fei & Justin Johnson & Serena Yeung,Stanford Course http://cs231n.stanford.edu/slides/2017, 2017 accessed in 03.2018

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, George E. Dahl,"Neural Message Passing for Quantum Chemistry", Corr Journal, April 2017

Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016, www.deeplearningbook.org

H. GREENSPAN, BRAM VAN GINNEKEN, RONALD M. SUMMERS "Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique "IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016

Hinton, G. E., Osindero, S., Teh, Y.-W., 2006." A fast learning algorithm for deep belief nets." Neural Computation NCBI 18, 1527–1554.

Andrej Karpathy, "notes Stanford CS class 231n: Convolutional Neural Networks for Visual Recognition", 2017 http://cs231n.github.io/ accessed in 03.2018

A Krizhevsky, I Sutskever, GE Hinton, "ImageNet classification with deep convolutional neural networks", 2012, Advances in neural information processing systems, 1097-1105

A. Krizhevsky. cuda-convnet. https://code.google.com/p/cuda-convnet/, 2012

Yann. Lecun ; L. Bottou ; Y. Bengio ; P. Haffner,"Gradient-based learning applied to document recognition", Proceedings of the IEEE (Volume: 86, Issue: 11, Nov 1998),Page(s): 2278 - 2324

Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015

Lee, J.-G., Jun, S., Cho, Y.-W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017)." Deep Learning in Medical Imaging: General Overview." Korean Journal of Radiology, 18(4), 570–584. http://doi.org/10.3348/kjr.2017.18.4.570

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, "A survey on deep learning in medical image analysis" Medical Image Analysis,Volume 42,2017,Pages 60-88,ISSN 1361-8415, https://doi.org/ 10.1016/j.media. 2017.07.005.

B. Sahiner et al.,1996, "Classification of mass and normal breasttissue: A convolution neural network classifier with spatial domain and texture images," IEEE Trans. Med. Imag., vol. 15, no. 5, pp. 598–610, Oct. 1996.

K. Simonyan ∗ & A.Zisserman ,"Very deep convolutional networks for large-scale image recognition",Iclr 2015

Erik Smistad and Lasse Løvstakken, "Vessel detection in ultrasound images using deep convolutional neural networks ",Proceedings of 2nd Workshop on Deep Learning in Medical Image Analysis (DLMIA). Athens, 21. Oct. 2016. Pages 30-38

Matthew D Zeiler, Rob Fergus,"Visualizing and Understanding Convolutional Networks",ECCV 2014: Computer Vision – ECCV 2014 pp 818-833

https://www.itnonline.com/article/how-artificial-intelligence-will-change-medical-imaging accessed 03.2018

www.Arterys.com accessed 03.2018

www.Caffe2.ai accessed 03.2018

www.Kera.io accessed 03.2018

www.microsoft.com/en-us/research/project/medical-image-analysis

www.Pytorch.orgaccessed 03.2018

www.tehnsorflow.org accessed 03.2018

www.Torch accessed 03.2018