IJEGEO International Journal of Environment and Geoinformatics 7(3):280-288 (2020)

Research Article

Deep Learning for Medicine and Remote Sensing: A Brief Review

Mehmet Eren Yuksel[1], **Sefa Kucuk**[2], **Seniha Esen Yuksel**[2,*], **Erkut Erdem**[3]

[1] Intensive Care Unit, Ankara Yıldırım Beyazıt University, School of Medicine, Bilkent, Ankara, TR
[2] Deparment of Electrical and Electronics Engineering, Faculty of Engineering, Hacettepe University, Ankara, TR
[3] Department of Computer Engineering, Faculty of Engineering, Hacettepe University, Ankara, TR

* Corresponding author: S.E. YUKSEL
E-mail: eyuksel@ee.hacettepe.edu.tr

Received 30 March 2020
Accepted 15 Sep 2020

How to cite: Yüksel, et al., (2020). Deep Learning for Medicine and Remote Sensing: A Brief Review, *International Journal of Environment and Geoinformatics (IJEGEO)*, 7(3):280-288 . DOI: 10.30897/ijegeo.710913

Abstract

In recent years, deep learning methods have come to the forefront in many areas that require remote sensing, from medicine to agriculture, from defense industry to space research; and these methods have achieved tremendous success as compared to traditional methods. Together with substantial growth in available data with high-quality labels and computational resources, these deep neural network architectures and techniques have seen remarkable developments. The major difference between deep learning and classical recognition methods is that deep learning methods consider an end-to-end learning scheme which gives rise to learning features from raw data. Better regularization techniques and robust optimization algorithms introduced with state-of-the-art deep learning models are other factors leading this difference. In this paper, we discuss the remote sensing problems and how deep learning can be used to solve these problems with a special focus on medical and remote sensing applications. In particular, we briefly review outperforming architectures within the deep learning literature and their use cases.

Keywords: Applications in Medicine, Applications in Remote Sensing, Deep Neural Network Architectures, Supervised Learning, Unsupervised Learning,

Introduction

Machine learning approaches called deep artificial neural networks have led to significant developments in many areas related to artificial intelligence, including computer vision in the past ten years. One of the underlying factors in the success of deep models is the end-to-end learning scheme these models have. Unlike the hand-crafted features extracted from raw data and used by traditional machine learning models in learning, these deep models can learn features from the raw data in a hierarchical manner. However, the distinctness of these features and the performance of deep models depend directly on the capacity of these models as well as on the amount of data on which the models are trained.

The companies such as Google, Microsoft, Facebook have used deep learning actively in their applications and services for object recognition, target detection, image segmentation, and classification. Moreover, deep learning toolboxes like Tensorflow from Google, CNTK from Microsoft, Torch from Facebook, Caffe, MXNet, Theano, Deeplearning4j led to a surge of studies on various topics (De Felice, 2017). For example, it has been proven that Convolutional Neural Networks (CNNs) are especially good at extracting medium and high-level features from raw data in a translation-invariant manner (Zhu et al., 2017). In this way, the features learned by CNNs are successfully used in problems such as image recognition, object detection, and semantic segmentation. On the other hand, Recurrent Neural Networks (RNNs) have shown significant success in sequential data analysis such as action recognition, video analysis, and subtitling. In addition to these achievements, the use of deep learning is becoming widespread in remote sensing, medical and defense industries thanks to increased data sizes and computing resources. Medical and remote sensing applications present some new challenges for deep learning. For instance, data in both application areas appear as multi-modal. Data acquired in remote sensing industry using various cameras such as light detection and ranging (LiDAR), radio detection and ranging (RADAR), synthetic aperture radar (SAR), hyperspectral should be collected and processed together. Similarly, it may be necessary to integrate a patient's genetic information and test results (e.g. blood and urine) with information obtained from images such as magnetic resonance imaging (MRI), ultrasound and x-ray. Multi-modality also requires the processing and interpretation of data received at different resolutions, at different times or over a long period of time series.

In this paper, firstly we attempt to answer the question about what deep learning is and what kind of success has been achieved. We briefly review state-of-the-art deep learning models and approaches. Then, we discuss supervised and unsupervised deep learning approaches, architectures and frameworks i.e. deep convolutional neural networks and deep generative networks. Also, we provide two sections for deep learning applications in medicine and remote sensing.

Table 1. Properties of models that won the ImageNet classification challenge. Top-5 error represents the percentage of test samples for which the correct class was not in the first 5 predicted classes.

Model	Year	Key points	# of Layers	# of Parameters	Top-5 Error
AlexNet	2012	ReLU non-linearity Multiple GPUs Dropout Data augmentation Overlapping pooling	8	61 M	16.4 %
VGGNet	2014	Simple architecture model Small receptive fields	19	144 M	7.3 %
GoogleNet	2014	Deeper in parallel path with different receptive field size Inception module	22	6.8 M	6.67 %
ResNet	2015	Shortcut connection Residual mapping	152	60 M	3.57 %
SENet	2017	Channel relationship Squeeze-and-Excitation blocks	SE bloks	145 M	2.251 %

Deep Learning Models

Deep convolutional neural networks used today for the classification of images obtained especially in the visible band have started to become popular with the successful realization of classification of handwriting numbers and letters with the model named LeNet-5 proposed by LeCun and his colleagues (Lecun et al., 1998). Several architectures have been proposed after LeNet. In 2012, Krizhevsky et al. proposed a deep artificial neural network model named AlexNet, whose architecture is similar to LeNet, for the problem of classification of high-resolution visible band images containing 1.2 million datasets with 1000 different classes (Krizhevsky et al., 2012). With its structure and concomitant techniques such as the dropout regularization, AlexNet beat all the competitors of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). After these developments, with the increasing amount of data, GoogLeNet (Szegedy et al., 2015) and VGGNet (Simonyan and Zisserman, 2014) were declared as the two most successful architectures in the ILSVRC 2014 competition. Later, with a revolutionary development, ResNet came to the forefront with its 152-layered structure, and won the ILSVRC 2015 competition (He et al., 2016). ResNet has the deepest network running on ImageNet. Its core idea is in the identity shortcut connection that skips one or more layers. The most important contribution of ResNet is that it can train the network more rapidly and successfully while increasing the depth of the network.

DenseNet (Dense Convolutional Network) is another family of CNN architecture which got best paper award in 2017 CVPR. DenseNet creates short paths from early layers to later layers and unlike short-cuts in ResNet, it combines features by concatenating them. Due to the fact that each layer receives feature maps from all previous layers, DenseNet has fewer parameters to learn, making them easier to train (Huang et al., 2017). Squeeze-and Excitation Network (SENet), the winner of the ILSVRC 2017 classification challenge, is another excellent example of CNN architectures. SENet networks have a mechanism that can recalibrate features for adaptively stating the importance of each channel. A single parameter is added to each channel and a linear criterion is used to show how relevant the added channel is. SE is actually a block that develops inter-channel dependencies for CNNs with virtually no computational costs. The goal of using these blocks in the convolution layers in SENet is to increase the sensitivity of the network against the useful attributes to be used in the network and to suppress the less useful ones. SE blocks are flexible that can be integrated into modern architectures with sophisticated designs (Hu et al., 2018). Table 1 shows some of these popular models with their properties and Figure 1 presents the mainstream architectures of aforementioned top competitors.

ResNet is easy to train, however, it cannot take full advantage of features in sequential convolutional layers with larger receptive fields. DenseNet combines these hierarchical features but the training of a deep dense network with large blocks is difficult. To address these drawbacks, Zhang et al. proposed Residual Dense Networks (RDN) to fully make use of all hierarchical features from the original image (Zhang et al., 2018). It is a fact that as the models developed to be run in the ImageNet dataset become more complex, the performance accuracy has increased but most models are not effective in terms of processing cost and need tedious manual parameter tuning. In recent years, researchers tried to create architectures that will produce more efficient results with smaller models. Unlike the approaches that arbitrarily scale network sizes such as width, depth, and resolution, Tan and Le have proposed a method called EfficientNets, which have scaling factors for each dimension (Tan and Le, 2019). EfficientNet has achieved superiority over the state of the art models by providing 84.4% accuracy with 66M parameter on ImageNet while being 8.4x

smaller and 6.1x faster on inference than the best available model.

AlexNet, VGGNet, GoogLeNet, ResNet and similar architectures are supervised deep learning methods that require labeled data for each class. However, another important issue is to learn deep networks with unsupervised learning. For instance, deep generative network models (Goodfellow et al., 2014; Maaløe et al., 2016) used in unsupervised learning, learn data distributions in high dimensions. Another popular example neural network architecture used in unsupervised learning is Deep AutoEncoders (AEs), in which outputs are the inputs (Deng and Yu, 2014; Goodfellow et al., 2016). AEs take the original data as an input encodes it for compact representation. Then, AEs decode to reconstruct the original input. Both supervised and unsupervised deep learning techniques will be described in more detail in the following sections.

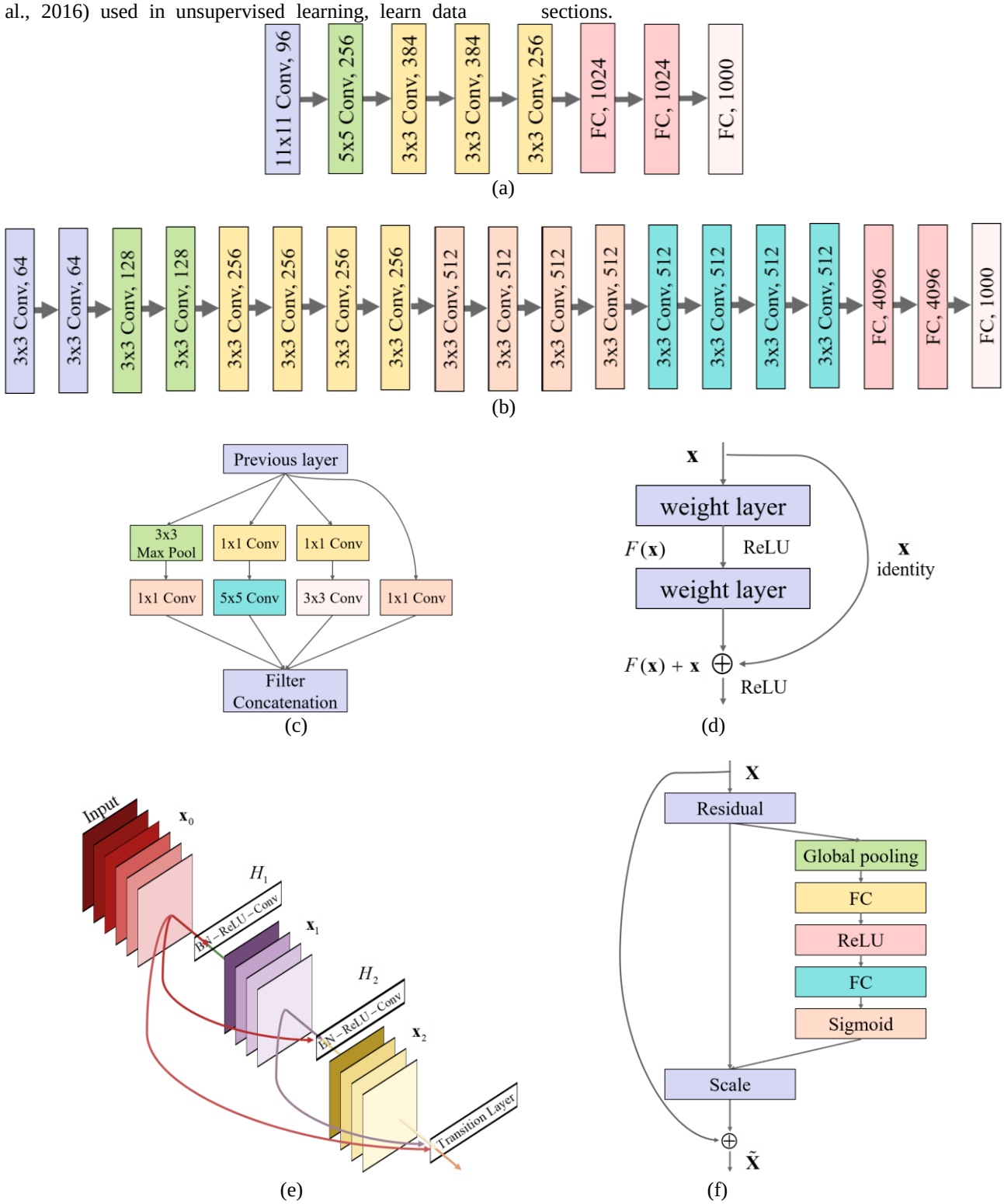

Figure 1. The mainstream architecture of (a) *AlexNet*, (b) *VGGNet*, (c) *GoogleNet*, (d) *ResNet*, (e) *DenseNet*, and (f) *SE-ResNet* module.

Supervised Deep Learning

Most of the deep networks, including *AlexNet*, which is given as an example here, are equipped with the following other important features:

i) Rectified Linear Units (ReLU) have been used in *AlexNet* instead of the traditionally used nonlinear functions such as sigmoid or hyperbolic tangent. ReLU not only reduced training time but also prevented training from becoming saturated.
ii) Overfitting of the data is prevented with the dropout layers.
iii) Data augmentation techniques have been used to artificially increase the size of the training set.
iv) Graphic processing units (GPU) have been used to rapidly incorporate all these techniques into the training.

Convolution, pooling and normalization layers, which are among the fundamental elements that stand out in *AlexNet* and similar architectures, are shown in Figure 2.

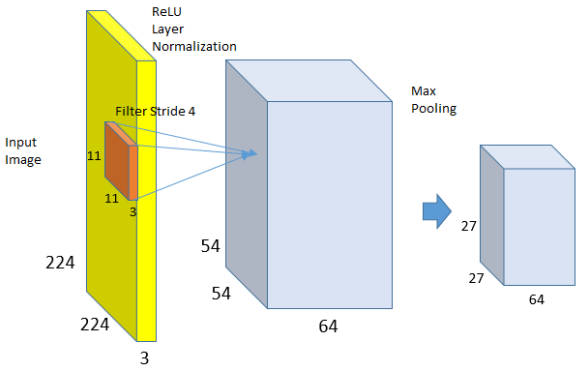

Figure 2. Change of the input data size after the initial convolution and pooling layers of *AlexNet* (Krizhevsky et al., 2012).

Today, architectures such as *AlexNet*, *VGGNet*, and *ResNet* can be used directly. Moreover, the filters learned by these architectures are being used in many studies as a feature extractor. With that, there are three main ways of training deep convolutional neural networks:

Classification with fully-trained networks:
In this type of training, the entire architecture is designed end-to-end and this new deep network is trained with its own data. This method corresponds to designing and training similar architectures such as AlexNet, GoogLeNet, VGGNet. The training of these architectures require millions of data, need GPUs and plenty of time to process them.

Feature extraction using pre-trained networks
In this type of learning, the last layers of networks like *AlexNet* (such as some of the fully connected layers and convolutional layers) are removed, and the remaining layers are used directly for feature extraction. This method is particularly preferred for applications with only limited number of data.

Feature extraction by fine-tuning on pre-trained networks
In this approach, as in the previous learning, the last layers are discarded however, the replaced layers are retrained according to the desired problem. Thus, both the training time is reduced and a deep network specific to the desired problem is learned.

Unsupervised Deep Learning

Exploring the ways of using deep networks in unsupervised learning is an increasingly important topic in deep learning literature, and various deep-generative network models have been proposed (Goodfellow et al., 2014; Gregor et al., 2015; Kingma and Welling, 2014; Larochelle and Murray, 2011; Radford et al., 2016; Rezende et al., 2014; Theis and Bethge, 2015; Van den Oord et al., 2016a; Van den Oord et al., 2016b; Shaham et al., 2019; Choi et al., 2019). These deep generative models aim to learn the probability distributions of samples via high dimensional examples such as digital images. Real-like synthetic samples can be produced by using deep generative networks through these learned distributions after the training procedure. For the case where the samples consist of digital images, the purpose of the generative models is to capture the pattern between the pixels and produce synthetic images that look as realistic as possible. These patterns can have local relationships such as adjacent pixels, which show similar visual characteristics in terms of color, brightness, and texture. It may also be possible to identify a particular object, such as the human face, along with the parts of the pixels. Considering their computational approaches, deep generative networks can be divided into three main groups as shown in Figure 3.

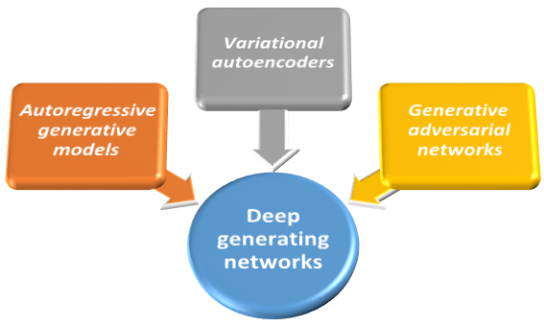

Figure 3. Major parts of deep generative networks.

The first group of the generative models is autoregressive generative models, which represent the fully observed model group among the generative models and they model sample production as a series of predictive problems (Larochelle and Murray, 2011; Theis and Bethge, 2015; Van den Oord et al., 2016a;

Van den Oord et al., 2016b). The best-known examples of models in this group are PixelCNN (Van den Oord et al., 2016b) and PixelRNN (Van den Oord et al., 2016a), which encode relationships between pixels via convolutional and recursive deep networks.

Variational autoencoders, unlike autoregressive generative models, define a hidden variable that can not be directly observed for each observed data point. Thus, they try to encode the distributions of the related examples, which are generally expressed in lower dimensions (Gregor et al., 2015; Kingma and Welling, 2014; Rezende et al., 2014). The most important advantage of this approach is that a deep hierarchical structure can be included in these models with its deep network. Another advantage of hidden variables is to express the structure of the data represented by the samples in a low-dimensional vector, which also provides compression in data representation. On the other hand, variational autoencoders that process digital images as examples typically produce partially blurry samples in the next generation phase and so, the photorealism level of these samples can be low.

The last group, generative adversarial networks (GANs), are transformation-based models and model the process of learning the distribution of real samples as a minimum-maximum game (Goodfellow et al., 2014; Radford et al., 2016). This game consists of two networks that compete with each other which are productive and distinctive. The productive network tries to create synthetic samples from noise vectors that will convince the distinctive network. The distinctive network tries to distinguish between real samples taken from the training data and synthetic samples produced by the productive network. Thus, both the productive network and the distinctive network become experts at the problems they try to solve. As a result, the productive network begins to produce highly realistic examples that are indistinguishable from reality. The training of generative adversarial networks is not easy due to its game theory-based approach, but the quality of the produced synthetic images is generally much sharper and more natural than the ones obtained from auto-encoders.

On the hand, GAN models are vulnerable to training instability. As a solution to this problem, the Spectrally Normalized GANS (SNGANs) propose a weight normalization technique to stabilize the training of the discriminator (Miyato et al., 2018). This normalization technique is computationally efficient and capable of generating high-quality synthetic images compared to previous training stabilization methods such as Wasserstein - GAN (WGAN) (Arjovsky et al., 2017). One of the state-of-the-art models in image synthesis, the BigGAN (Brock et al., 2018), also uses this spectral normalization method. Futher, the first version of GANs could produce high-quality synthetic images; but the size of the produced images often remained relatively small. BigGAN, on the other hand, focuses on scaling up the GAN models for generating both high quality and larger images by configuring the model and the training process.

In the vanilla GANs, although the generated images become more realistic over time, controlling and understanding their output remains one of the main challenges. Style-Based Generator Architecture for GANs (StyleGAN), one of the most famous GAN models recently published, proposes a new model that addresses this problem (Karras et al., 2019). StyleGAN produces synthetic images by using the progressive growing method, i.e. it generates high-resolution images by starting with very low resolution. It focuses on the generator network and uses the adaptive instance normalization module to transfer the encoded information into the generated image via a mapping network. StyleGAN has made great advances in the quality of GANs, but when the images generated are examined, blob-like unnatural artifacts (water droplets) are observed. To tackle this problem, StyleGAN2, an extended version of StyleGAN, has been developed recently (Karras et al., 2020). StyleGAN2 takes advantage of a hierarchical generator by skipping connections instead of progressive growing and uses the normalized weight of CNN instead of adaptive sample normalization statistics to eliminate these water droplets.

Traditional versions of GANs learn representations form multiple real images. However, in SinGAN, a variant of GAN introduced in recent years, image generation is learned from a single training image (Shaham et al., 2019). The model includes a pyramid of fully convolutional GANs in which generator and discriminator are responsible for learning representation at different scales of the image. This architecture allows generating new samples with arbitrary size and aspect ratio. The model is trained in a coarse-to-fine fashion. The pair of generator and discriminator learn coarse features like background at the lowest scale, whereas they learn fine details like edges and corners at high scales. SinGAN achieved impressive results in a variety of image manipulation tasks, thereby winning the best paper award at ICCV (IEEE International Conference on Computer Vision) 2019.

Another compelling application of GANs is the image-to-image translation, which expresses the problem of generating new images based on input images. StarGAN is one of the well-known image-to-image translation models (Choi et al., 2018). It uses a unified model architecture that allows simultaneous training of multiple datasets with different domains within a single network. The superiority of the approach has been tested on a facial attribute transfer and facial expression synthesis tasks. StarGAN v2 is an extended version of StarGAN, which addresses two major challenges in image-to-image translation framework: i) translating an image from one domain to diverse images in target domain, ii) supporting multiple target domain at the same time (Choi et al., 2019).

With this summary of the recent advancements in deep learning, the two sections below survey some of the leading examples of deep learning in medicine and in remote sensing.

Applications of Deep Learning in Medicine

Machine learning in medicine is being developed all the way across the treatment spectrum; for disease prevention, for ordering patients according to their urgency and for diagnosis (Lee et al., 2019; Suzuki, 2017).

A remarkable example of deep learning on radiological diagnostics is the study published by Rajpurkar and his colleagues in 2018 (Rajpurkar et al., 2018). The software named CheXNeXt that has been developed in this study uses a DL based algorithm to detect clinically important pathologies in front-view chest X-rays. Trained on more than 100,0000 chest X-rays from approximately 31,000 patients in the National Institutes of Health (NIH) ChestX-ray8 dataset, this software has been tested on 420 images and the results have been compared with nine different radiologists. The CheXNeXt algorithm has not only achieved a comparable success as compared to the radiologists decisions, but also managed to analyze this data in under 1.5 minutes as opposed to the 240 minutes spent by the radiologists. It is thought that such software may be used in regions where radiologists are absent. It is also thought that CheXNeXt may reduce misdiagnosis due to lack of expertise or fatigue (Saria et al., 2018).

Another noteworthy example of using deep learning as a preventive method is the study of Corey and his colleagues to detect complications that may occur within 30 days after surgery (Corey et al., 2018). With the data obtained from approximately 88,000 patients, a software called Pythia has been developed that takes into account the age, race, gender, medications used, and the history of comorbidity. With this information, Pythia has been found to be able to detect particularly high-risk patients with high sensitivity. The purpose of such preventive support systems is to identify high-risk patients and to minimize the complications that may occur in these patients with precautions taken before or after the operation. Similarly, MedAware software which is a commercially available product, monitors the patient's historical data, detects potential drug errors by machine learning methods and warns them.

There are also several studies that provide the classification of heart and lung sounds with deep convolutional neural networks while preserving the comfort of the specialist and patient without collecting any additional data from the patient (Aykanat et al., 2017; Bardou et al., 2018). Bien et al. presented MRNet, an automated deep learning method for interpreting knee magnetic resonance imaging (MRI), and compared their results with general radiologists (Bien et al., 2018). They examined 1,370 knee MRI images at Stanford University Medical Center. Their findings support the claim that deep learning-based models can improve the performance of clinical specialists during medical imaging interpretation.

Hosny and his colleagues have trained an end to end 3D CNN structure on seven independent datasets across five institutions totaling 1,194 patients to assess the utility of deep learning networks in predicting 2-year overall survival of non-small-cell lung cancer patients from computerized tomography data. They have concluded that learned deep learning features dramatically predominate over existing prognostication methods in post-surgery patients (Hosny et al., 2018).

Pham and his colleagues have used a two-step deep learning approach to detect lung cancer of lymph node metastases from whole slide histopathologic images. In the first step of this study, a CNN based deep learning structure has been used to eliminate misclassified noncancerous regions. In the second step, the cancer cells were detected by a deep learning classifier (Pham et al., 2019).

The study of Gurovich and his colleagues is another conspicuous example of deep learning applications in medicine. They have presented a facial analysis framework for genetic syndrome classification called DeepGestalt (Gurovich et al., 2019). They have analyzed facial images with DeepGestalt, a CNN-based structure trained on a dataset of over 17,000 images representing more than 200 syndromes. Deep Gestalt has produced results that will add significant value to phenotypic evaluations in clinical genetics and precision medicine.

The study of Zheng and his colleagues is one of the newest examples of deep learning applications in medicine (Zheng et al., 2020). They have developed an early-warning-system for high-risk suicide attempt patients. Advanced machine learning algorithms and deep neural network models have been used on the data obtained from the electronic health records of Berkshire Health System. They have calculated a final risk score for each patient and illustrated the probability of a suicide attempt in the next one-year period.

Most recently, deep learning has also found its use in the Coronavirus 2019 disease (COVID-19). COVID-19 is a highly contagious viral respiratory disease characterized by high fever, shortness of breath, and respiratory distress. The devastating effect of COVID-19 on global health continues: as of 25th June, the number of COVID-19 cases has exceeded 9.5 million and the death toll has exceeded 485,000 in more than 213 countries and territories across the world (Worldometers, 2020). Deep learning methods have also been utilized to control the spread of the disease by quickly screening large numbers of suspected cases to take appropriate quarantine and treatment measures. Wang et al. have used a CNN based architecture on the collected 453 computed tomography images to extract certain graphical features of COVID-19 (Wang et al., 2020). Wang and Wong have proposed a deep neural

network model called COVID-Net for the detection of COVID-19 cases from chest-X-ray (CXR) images (Wang and Wong, 2020). They have also introduced COVIDx, an open access dataset over 13,900 images across 13,870 patient. They have concluded that their model has achieved high positive predictive value (98.9%). These two studies show that deep learning methods can be applied to CT or CXR to effectively scan COVID-19. It is hoped that such advancements in the screening technologies would save time in treatment and help prevent the spread of the disease.

Applications of Deep Learning in Remote Sensing

In the remote sensing, deep learning appears both in target detection applications that require image and signal processing and also in logistics. Machine learning in logistics is needed in applications such as dynamic inventory optimization or predictive maintenance and services. For example, the system estimates when the used vehicles will expire or fail and orders accordingly; providing an urgent supply of batteries for a vehicle that is running out of battery (Lennon, 2018).

In target detection applications, the goal is to rapidly detect the targets from the data collected by the satellites, from flying vehicles such as UAVs, aircrafts, drones, or from land vehicles and fixed stations. For this purpose, data can be signals or images obtained from sensors like radar, SAR, LiDAR, hyperspectral, thermal, electro-optic, X-ray, mm-Wave or terahertz. In addition to interpreting these images instantaneously with deep learning techniques, predicting enemy behaviors, collecting weather conditions and coming up with decisions that are independent of weather conditions, and identifying narrow pass points can be listed as some of the many other goals (Shakhatreh et al., 2019).

Vessel/ship detection and tracking in target detection problems stand out as an important military problem both due to the scale difference and the diversity of the background. To detect ships with deep learning methods, Fu et al. and Yang et al. have used Google Earth images obtained from Quick Bird satellite (Fu et al., 2018; Yang et al., 2018), Rainey et al. used Digital Globe satellites from World View-1 and World View-2 images (Rainey et al, 2016); Wang et al. have used SAR images (Wang et al, 2017) and Khellal et al. have used infrared images (Khellal et al., 2018)

As an example of predicting environmental conditions with deep learning, Drönner and his colleagues have made cloud predictions with deep convolutional neural networks to both make weather forecasts and to effectively distribute solar energy (Drönner et al., 2018). Ouala et al. have developed a system to measure the temperature of the sea surface using the seven-year satellite data provided by the UK Met Office (Ouala et al, 2018).

Deep learning techniques also plays an important role in mine, wire and handmade explosive detection, which are some of the most vital target detection problems. Ground Penetrating Radar (GPR) is an extensively used tool for buried target detection. In their related studies, Aydın and Yüksel (Aydın and Yüksel, 2017; Aydın and Yüksel, 2018; Aydın and Yüksel, 2019) have utilized transfer learning as it is very difficult and time-consuming to extract features in the classification of targets with similar electrical conductivity. They have taken the first two layers of AlexNet architecture and trained the last layer for GPR classification. They have also used multitasking learning so that the model learned in one soil type could be trained according to another soil type. In this method, they have developed a deep learning architecture that understands which type of soil (dry, wet, moist) the target is in.

Salman and Yüksel have used a similar architecture for the fusion of the hyperspectral and LiDAR data (Salman and Yüksel, 2018). However, since the dimensions of the hyperspectral and LiDAR data are different from the images used in classical deep learning architectures, they have proposed various pre-processing methods to equalize these sizes. They have shown that the results are much more successful than traditional methods.

In the overall, an ideal remote sensing system should see what cannot be seen, should hear what cannot be heard, and should predict what is coming long before it happens to give the soldiers or decision makers ample time. Through the fusion of several sensors and external/prior information, deep learning methods are showing a promise in extracting the most useful information in each sensor in many applications as discussed above.

Conclusion

In this study, we have provided a brief review of deep learning approaches with a special focus on their use in medicine and remote sensing. A general look at the successful applications indicates what we always knew: that data is the key. However, now more than ever, collecting and understanding data is being more efficiently dealt with due to the advancements in technology and the methods in use. This includes storing patients' data from all sorts of tests and for as long as possible for healthcare; and incorporating all kinds of sensors' data such as weather information, intelligence data and the like for remote sensing. Considering the successes of the existing studies, we believe that deep learning will be used in these areas with increasing momentum in the coming years.

References

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein gan. *arXiv preprint arXiv*:1701.07875.

Aydın, E., Yüksel, S. E. (2017, May). Buried target detection with ground penetrating radar using deep

learning method. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Aydın, E., Yüksel, S. E. (2018, May). Transfer and multitask learning method for buried wire detection via GPR. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Aydın, E., Yüksel, S. E. (2019, May). Transfer and multitask learning using convolutional neural networks for buried wire detection from ground penetrating radar data. In *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIV* (Vol. 11012, p. 110120Y). International Society for Optics and Photonics.

Aykanat, M., Kılıç, Ö., Kurt, B., and Saryal, S. (2017). Classification of lung sounds using convolutional neural networks. *EURASIP Journal on Image and Video Processing*, 2017(1), 65.

Bardou, D., Zhang, K., and Ahmad, S. M. (2018). Lung sounds classification using convolutional neural networks. *Artificial intelligence in medicine*, 88, 58-69.

Bien, N., Rajpurkar, P., Ball, R. L., Irvin, J., Park, A., Jones, E., and Halabi, S. (2018). Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. *PLoS medicine*, 15(11), e1002699.

Brock, A., Donahue, J., Simonyan, K. (2018). Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv*:1809.11096.

Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., Choo, J. (2018). Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8789-8797).

Choi, Y., Uh, Y., Yoo, J., Ha, J. W. (2019). StarGAN v2: Diverse Image Synthesis for Multiple Domains. *arXiv preprint arXiv:1912.01865*.

Corey, K. M., Kashyap, S., Lorenzi, E., Lagoo-Deenadayalan, S. A., Heller, K., Whalen, and K., Sendak, M. (2018). Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia): A retrospective, single-site study. *PLoS medicine*, *15*(11).

De Felice, M. (2017). Which deep learning network is best for you? IDG Communications. [Online]. Available: http:// www.cio.com/ article/ 3193689/ artificial-intelligence/whichdeep-learning-network-is-best-for-you.html

Deng, L., and Yu, D. (2014). Deep learning: methods and applications. *Foundations and trends in signal processing*, 7(3–4), 197-387.

Drönner, J., Korfhage, N., Egli, S., Mühling, M., Thies, B., Bendix, J., Freisleben, and Seeger, B. (2018). Fast cloud segmentation using convolutional neural networks. *Remote Sensing*, *10*(11), 1782.

Fu, K., Li, Y., Sun, H., Yang, X., Xu, G., Li, Y., Sun, X. (2018). A ship rotation detection model in remote sensing images based on feature fusion pyramid network and deep reinforcement learning. *Remote Sensing*, *10*(12), 1922.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning.* MIT press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville and Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).

Gregor, K., Danihelka, I., Graves, A., Rezende, D. J. and Wierstra, D. (2015). Draw: A recurrent neural network for image generation. *arXiv preprint arXiv:1502.04623*.

Gurovich, Y., Hanani, Y., Bar, O., Nadav, G., Fleischer, N., Gelbman, D., and Bird, L. M. (2019). Identifying facial phenotypes of genetic disorders using deep learning. *Nature medicine*, 25(1), 60-64.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hosny, A., Parmar, C., Coroller, T. P., Grossmann, P., Zeleznik, R., Kumar, A. and Aerts, H. J. (2018). Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLoS medicine*, 15(11), e1002711.

Hu, J., Shen, L., and Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).

Karras, T., Laine, S., Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4401-4410).

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T. (2020). Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8110-8119).

Khellal, A., Ma, H., Fei, Q. (2018). Convolutional neural network based on extreme learning machine for maritime ships recognition in infrared images. *Sensors*, *18*(5), 1490.

Kingma, D. P., Welling M. (2014). Auto-encoding Variational Bayes. ICLR.

Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Larochele, H., Murray, I. (2011, June). The neural autoregressive distribution estimator. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* (pp. 29-37).

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324.

scenes based on multiscale rotation dense feature pyramid networks. *Remote Sensing*, *10*(1), 132.

Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y. (2018). Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2472-2481).

Zheng, L., Wang, O., Hao, S. Ye, C., Liu, M., *et al.* (2020). Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health records. *Transl Psychiatry* 10, 72.

Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8-36.